



LIBRO DE ACTAS

XXV Workshop de Investigadores en Ciencias de la Computación

Buenos Aires, abril del 2023
UNNOBA – Red de Universidades
de Carreras de Informática RedUNCI



ITT

Instituto
Transferencia
Tecnología



ESCUELA DE
TECNOLOGÍA



UNNOBA



UNIVERSIDAD NACIONAL
NOROESTE • BUENOS AIRES

Datos de la edición

Escuela de Tecnología

Directora: Sarobe, Mónica

Secretario Académico: Cicerchia, Lucas Benjamin

Instituto de Investigación y Transferencia en Tecnología – ITT

Director: Ramón, Hugo

Investigador: Esnaola, Leonardo

Investigador: Tessore, Juan Pablo

Red de Universidades de Carreras de Informática – RedUNCI

Representante: Russo, Claudia

Edita

CEDi Centro de Edición y Diseño. UNNOBA

Callao 289 3.º piso, CP. 1022

Tel: 54 -11-535-31520. Ciudad Autónoma
de Buenos Aires, Argentina

Diseño y diagramación de tapa e interior:

CEDi | cedi@unnoba.edu.ar

Fecha de catalogación:

Queda hecho el depósito que marca la ley 11723

Reservados todos los derechos

Julio | 2023

Libro de Actas WICC 2023-UNNOBA : XXV Workshop de Investigadores en Ciencias de la
ComputaciónXXV Workshop de Investigadores en Ciencias de la Computación ...

[et al.];

compilación de Mónica Sarobe ... [et al.]. - 1a ed - Junín : Universidad Nacional del
Noroeste de la Provincia de Buenos Aires. UNNOBA, 2023.

Libro digital, PDF

Archivo Digital: descarga

ISBN 978-987-3724-66-4

1. Computación. 2. Educación. 3. Universidades. I. XXV Workshop de Investigadores
en Ciencias de la Computación II. Sarobe, Mónica, comp.

CDD 004.0711



Esta obra está bajo una licencia de Creative Commons Reconocimiento
- NoComercial - SinObraDerivada 4.0 Internacional.

Comité Académico

UBA – Cs. Exactas
Ceria, Santiago

UBA – Ingeniería
Echeverría, Adriana

UN La Plata
Pesado, Patricia

UN Sur
Rueda, Sonia

UN San Luis
Printista, Marcela

UNCPBA
Aciti, Claudio

UN Comahue
Grosso, Guillermo

UN La Matanza
Eterovic, Jorge

UN La Pampa
Alfonso, Hugo

UN Tierra del Fuego
Korembli, Gabriel

UN Salta
Gil, Gustavo

UN Patagonia Austral
Lasso, Marta

UN San Juan
Rodriguez, Nelson

UADER
Mengarelli, José Luis

UN Patagonia SJB
Buckle, Carlos

UN Entre Ríos
Tugnarelli, Mónica

UN Nordeste
Dapozo, Gladys

UN Rosario
Casali, Ana

UN Misiones
Caballero, Sergio

UNNOBA
Russo, Claudia

UN Chilecito
Carmona, Fernanda

UN Lanús
Azcurrea, Diego

UN Santiago del Estero
Figuerola, Liliana

Universidad de la Defensa Nacional
Cipriano, Marcelo

UN Litoral
Loyarte, Horacio

UN Rio IV
Arroyo, Marcelo

UN Córdoba
Fridlender, Daniel

UN Jujuy
Herrera Cagnetta, Analía

UN Rio Negro
Vivas, Luis

UN Villa María
Prato, Laura

UN Lujan
Fernandez, Juan Manuel

UN Catamarca
Poliche, Maria Valeria

UN La Rioja
Molina, Miguel

UN Tres de Febrero
Oliveros, Alejandro

UN Tucumán
Luccioni, Griselda María

UNAJ
Morales, Martín

UN Chaco Austral
Zachman, Patricia

UN del Oeste
Foti, Antonio

UN de Cuyo
Haderne, Marisa Fabiana

UN de Mar del Plata
Ríos, Carlos

UN de Quilmes

UN Hurlingham
Puricelli, Fernando

UNSAAdA
Ramón, Hugo

UNSAM
Estayno, Marcelo

UMET
Quiroga, Flavia

U Morón
Chapperon, Gabriela

UAI
De Vincenzi, Marcelo

U Belgrano
Guerçi, Alberto

U Kennedy
Panizzi, Marisa

U Adventista del Plata
Bournissen, Juan

UCAECE
Malbernath, Lucía

U Palermo
Alvarez, Adriana

UCA Rosario
Grieco, Sebastián

U Salvador
Zanitti, Marcelo

U Aconcagua
Giménez, Rosa

U Gastón Dachary
Ruidías, Hector Javier

UADE
Feijó, Daniel

UCEMA
Guglianone, Ariadna

U Austral
Cosentino, Juan Pablo

U Atlántida Argentina
Rathmann, Liliana

UCA La Plata
Bertone, Rodolfo

ITBA
Bolo, Mario

U Champagnat
Brachetta, Mariana

U Mendoza
Asensio, Mariela

Coordinadores de Áreas

Agentes y Sistemas Inteligentes

Marcelo Falappa (UNS)
Marcelo Errecalde (UNSL)
Daniel Pandolfi (UNPA)

Arquitectura, Redes y Sistemas Operativos

Luis Marrone (UNLP)
Daniel Arias Figueroa (UNSa)
Orlando Micolini (UNC)

Computación Gráfica, Imágenes y Visualización

Martín Larrea (UNS)
María J. Abásolo (UNLP-UNCPBA)
Roberto Guerrero (UNSL)

Ingeniería de Software

Pablo Fillottrani (UNS)
Pablo Thomas (UNLP)
Fernanda Carmona (UNdeC)

Innovación en Sistemas de Software

Monica Tugnarelli (UNER)
Marisa Panizzi (UK)
Guillermo Feierherd (UNTDF)
Gladys Dapozo (UNNE)

Procesamiento Distribuido y Paralelo

Marcelo Naiouf (UNLP)
Marcela Printista (UNSL)
Javier Balladini (UNCOMA)

Tecnología Informática aplicada en Educación

Zulma Cataldi (UBA-UTN)
Alejandra Zangara (UNLP)
Gustavo Gil (UNSa)

Procesamiento de señales y Sistemas de Tiempo Real

Oscar Bría (INVAP)
Fernando Tinetti (UNLP)
Nelson Rodriguez (UNSJ)

Bases de Datos y Minería de Datos

Laura Lanzarini (UNLP)
Claudia Deco (UNR)
Norma Herrera (UNSL)

Innovación en Educación Informática

Claudia Russo (UNNOBA)
Lucía Malbernat (UCAECE)

Seguridad Informática

Paula Venosa (UNLP)
Jorge Eterovic (UNLaM)
Javier Echaiz (UNS)
Antonio Castro Lechtaller (UBA)

Jurado Tesis Doctorado

Marcela Printista (UNSL)
Laura De Giusti (UNLP)
Silvia Castro (UNS)
Alejandra Cechich (UNCOMA)
Horacio Kuna (UNaM)

Índice

ARSO – Arquitectura, Redes y Sistemas Operativos	Pág (12-57)
Desafíos y buenas prácticas para pequeños proveedores de Internet zonales.	Pág (13-16)
Migración controlada de procesos en sistemas distribuidos.	Pág (17-21)
Estudios del aporte de tecnologías como BLOCKCHAIN, NFT y VPN a la seguridad en plataformas de IOT.	Pág (22-26)
IOT Aplicado a la ganadería de precisión	Pág (27-31)
IoT for Vineyard Posts.	Pág (32-36)
IoT-based energy monitoring system, to deploy an AMI.	Pág (37-41)
Utilización Eficiente y Sostenible de Recursos en Clúster de bajo costo.	Pág (42-46)
Tecnologías de Smart IoT para la solución de problemas en el medio socio productivo.	Pág (47-52)
Despliegue de aplicaciones contenerizadas: un caso de implementación basado en crane.	Pág (53-57)
ASI – Agentes y Sistemas Inteligentes	Pág (58-156)
Interfaces de Programación de Aplicaciones orientadas a la Integración de Sistemas Judiciales.	Pág (59-63)
Organizaciones inteligentes asistidas por tecnologías 4.o.	Pág (64-68)
Aprendizaje Automático para Nodos Inteligentes.	Pág (69-71)
Depuración de bases de datos de segunda sección del Boletín oficial de la república argentina mediante Aprendizaje de máquina.	Pág (72-76)
Una herramienta de Inteligencia Artificial para el modelado de distintos escenarios con la determinación del Índice de Calidad del Agua (ICA) para aplicar a ríos urbanos de Argentina.	Pág (77-81)
Movilidad inteligente en una ciudad sostenible.	Pág (82-86)
Asistente virtual con comportamiento definible de consultas referidas a la legislación argentina para personas con discapacidad.	Pág (87-89)
Automatización robótica de procesos en las pymes.	Pág (90-94)
Clasificación de partículas de arena a través de redes neuronales convolucionales.	Pág (95-100)
Resolución de problemas regionales NP-duros usando técnicas de optimización.	Pág (101-105)
Confeción de dataset de patentes de automóviles argentinos para entrenamiento de redes neuronales.	Pág (106-110)
Metaheurísticas aplicadas al alineamiento múltiple de secuencias de ADN y/o de aminoácidos.	Pág (111-114)
Aplicación de inteligencia computacional y computación de alto desempeño en el desarrollo de un modelo de predicción de las condiciones predisponentes al quemado del arroz (<i>pyricularia oryzae</i>).	Pág (115-119)
Lenguaje natural aplicado a los sistemas conversacionales para el aprendizaje de lenguas extranjeras.	Pág (120-124)
Tecnologías Semánticas para el desarrollo de Agentes Inteligentes: Generación de comentarios a partir de código fuente.	Pág (125-128)
Deep Learning para Visión por Computadora.	Pág (129-133)
Inteligencia artificial como herramienta para innovar y dinamizar procesos.	Pág (134-139)
Detección de Ransomware en Blockchains.	Pág (140-142)
Desarrollo de soluciones tecnológicas de bajo costo a problemáticas en el sector hortícola.	Pág (143-147)
Desarrollo de herramientas para generación de material de entrenamiento y análisis de textos estructurados.	Pág (148-152)
Sistema de predicción de incendios forestales.	Pág (153-156)

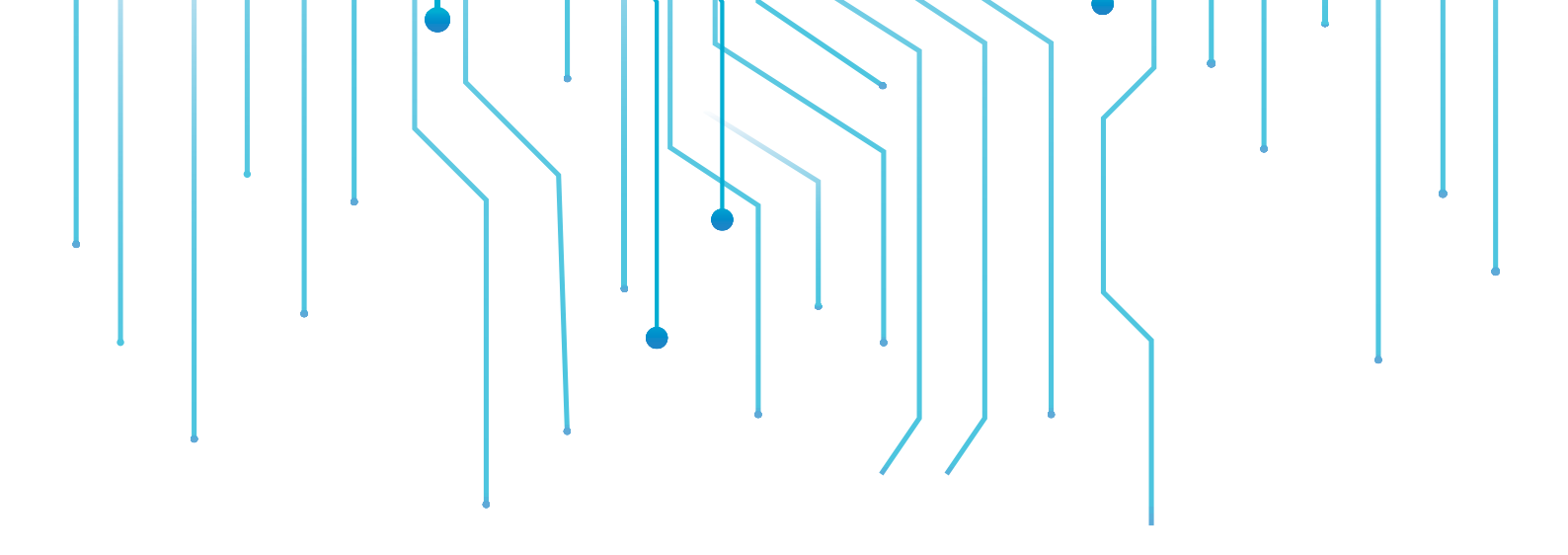
BDDM – Bases de Datos y Minería de Datos	Pág (157-216)
Propuesta de técnicas de validación para la calidad de datos abiertos e identificación de patrones para predicciones con Machine Learning: Segunda Parte.	Pág (158-162)
Knowledge Graphs para accesibilidad y reutilización de activos digitales del Mar Argentino.	Pág (163-166)
Descubrimiento de patrones de comportamiento vinculados al abandono en la Universidad Nacional de Luján mediante la aplicación de técnicas de aprendizaje automático.	Pág (167-173)
Desarrollo de soluciones de inteligencia y analítica de negocios.	Pág (174-177)
Rediseño de la Plataforma Web MultiDB, Cuyo Propósito es el Aprendizaje de Lenguaje SQL, con Acceso a Distintos Motores de Base de Datos.	Pág (178-182)
Sistemas inteligentes en el uso de aplicaciones de bioinformática y sistemas embebidos.	Pág (183-187)
Datos no Estructurados: Indexación, Búsquedas y Aplicaciones	Pág (188-192)
Administración y Recuperación de Información en Bases de Datos Masivas.	Pág (193-197)
Diversas Aplicaciones de Bases de Datos Espaciales y Espacio Temporales.	Pág (198-202)
Bases de Datos no Convencionales.	Pág (203-207)
Comparación de métodos y métricas para la predicción de series temporales aplicados a las ventas masivas de productos minoristas.	Pág (208-212)
Índices de dominio para identificación eficiente mediante código de iris.	Pág (213-216)
CGIV – Computación Gráfica, Imágenes y Visualización	Pág (217-261)
Testing de visualizaciones.	Pág (218-221)
PACS Cloud Open Source.	Pág (222-226)
Análisis visual de movimientos oculares en entornos dinámicos.	Pág (227-231)
Visualización comparativa: estudio y diseño de visualizaciones orientadas a la comparación de conjunto de datos.	Pág (232-236)
Tecnologías inmersivas y visualización situada aplicadas a geociencias.	Pág (237-241)
Visualización de datos multidimensionales procedentes de las geociencias.	Pág (242-246)
Análisis visual guiado para el estudio de datos geoquímicos.	Pág (247-251)
Industrias Inteligentes: Detección de uso de Elementos de Protección Personas en una Planta de Reciclaje.	Pág (252-256)
La importancia de la realidad extendida en el metaverso.	Pág (257-261)
IEI – Innovación en Educación en Informática	Pág (262-331)
Adopción de verilog en materias de diseño digital.	Pág (263-267)
Aplicación con realidad aumentada para el aprendizaje de los conceptos de programación.	Pág (268-272)
Innovación Tecnológica y Metodológica en Educación.	Pág (273-277)
Evaluación de la Calidad de Objetos de Aprendizaje, para la enseñanza de la Informática, desde el Enfoque de los estudiantes. Medición del significado afectivo.	Pág (278-282)
Marco para desarrollar nuevas unidades de proyectos e incorporar docentes y alumnos en actividades de investigación.	Pág (283-287)
Herramientas colaborativas multiplataforma para el aprendizaje basado en competencias en la ingeniería de software.	Pág (288-295)
El portafolio como metodología de trabajo en un grupo de investigación de infraestructura it.	Pág (296-300)
Enseñar programación en la universidad en tiempos de gran demanda de programadores.	Pág (301-305)
Eye tracking para la enseñanza de la programación.	Pág (306-310)
Demanda laboral y competencias en inteligencia artificial dentro de carreras grado en ciencias aplicadas.	Pág (311-315)

Proyecto Esport UNLP Diplomatura en Deportes Electrónicos.	Pág (316-321)
Programación y Autismo: una arquitectura epistémica centrada en Ciencia Ciudadana.	Pág (322-326)
Desarrollo de Recursos Educativos para la Enseñanza de las Ciencias de la Computación en la Escuela Secundaria.	Pág (327-331)
IS – Ingeniería en Software	Pág (332-481)
Principios esenciales de la ingeniería de software en ámbito público.	Pág (333-337)
Catálogo y recomendaciones de recursos (métodos) para facilitar experiencias de Diseño Participativo usando Design Thinking.	Pág (338-342)
Análisis de viabilidad del despliegue de sistemas de software en PyMES.	Pág (343-347)
Métodos, técnicas y herramientas para mejorar y evaluar la calidad de requisitos en proyectos ágiles de software.	Pág (348-352)
Métodos, técnicas y herramientas para mejorar y evaluar la calidad de la capacidad de producción de entregables dentro de un proceso ágil de desarrollo de software	Pág (353-357)
Creación de un observatorio tecnológico de la gestión del conocimiento en la ciudad de buenos aires.	Pág (358-362)
Aprovisionamiento automático en la nube mediante la Ingeniería Dirigida por Modelos.	Pág (363-367)
Hacia la evaluación de la calidad de datos abiertos.	Pág (368-372)
Aplicaciones Móviles, Realidad Virtual y Realidad Aumentada.	Pág (373-376)
Evaluación de las políticas de protección de datos personales en las iniciativas de apertura de datos en el ámbito de la Universidad Nacional de Catamarca.	Pág (377-381)
Desarrollo de Computadoras de a Bordo para Misiones Satelitales de Órbita Baja.	Pág (382-386)
Enfoques y Tendencias en el Desarrollo de Aplicaciones para Dispositivos Móviles.	Pág (387-392)
Construcción de colecciones de proyectos de software para estudios empíricos.	Pág (393-397)
Buenas prácticas de calidad de software en el ámbito público.	Pág (398-402)
Evaluación de lenguajes de modelado gráficos de procesos de negocio colaborativos utilizando un método multicriterio.	Pág (403-407)
Diseño de un marco de trabajo para la gestión del conocimiento de productos de software en evolución.	Pág (408-410)
Aspectos de ingeniería de software, bases de datos relacionales, y bases de datos NOSQL para el desarrollo de sistemas de software híbridos.	Pág (411-417)
Desarrollo de Software Público: Un aporte para la mejora de la calidad.	Pág (418-422)
Cultura, turismo y tecnología. Experiencias turísticas significativas para circuitos históricos de caleta Olivia.	Pág (423-427)
Desarrollo y evaluación de productos web y móvil centrados en la experiencia de usuario.	Pág (428-433)
Evaluación de la Sostenibilidad en un Modelo de Calidad del Software.	Pág (434-438)
CoVaMaT: Modelo conceptual de una herramienta para el soporte a la gestión de variedad en Sistemas Big Data.	Pág (439-443)
Proceso de validación de requerimientos aplicando técnicas de procesamiento de lenguaje natural en un entorno colaborativo.	Pág (444-448)
Computación afectiva aplicada a la interacción con Interfaces Gráficas de Usuario.	Pág (449-453)
Calidad de información en comunidades virtuales.	Pág (454-458)
Accesibilidad web. Estrategias y aplicaciones.	Pág (459-463)
Estimación del tamaño de un sistema mediante el análisis de diagramas UML y la aplicación de métodos basados en puntos de función.	Pág (464-467)
Gobernanza de Sistemas y Tecnologías de la Información (SI/TI) en el ámbito Universitario.	Pág (468-474)
Transformación Digital – Propuestas de solución para las Ciudades Inteligentes Sostenibles, la Gobernanza Digital y las Buenas prácticas y calidad.	Pág (475-480)

ISS – Innovación en Sistemas de Software	Pág (481-573)
Estado actual de la industria 4.0 en las PyMES del partido de Hurlingham.	Pág (482-486)
Sistema de recomendación ubicuo para contribuir a la disminución del estrés académico en estudiantes universitarios.	Pág (487-490)
El avance en la madurez digital de las pymes a través de la transformación digital.	Pág (491-495)
Tecnologías de la información facilitadoras para la interoperabilidad de software en Gobierno Abierto: Segunda Parte.	Pág (496-500)
Desarrollo de una app móvil de Nueva Generación.	Pág (501-504)
Soluciones IoT desde Smart Cities a Blockchain.	Pág (505-510)
Diseñar una solución de identidad auto-gestionada para acceso a servicios de calidad con redes Blockchain multipropósito en la Universidad Nacional de Río Negro.	Pág (511-515)
Aplicación de geofencing para el cuidado de adultos mayores mediante un hardware dedicado.	Pág (516-519)
Programa integrador de proyectos de Software con transferencia a la comunidad.	Pág (520-523)
Innovación en las Ciudades mediante las TIC Análisis de implementaciones en Ciudades Inteligentes.	Pág (524-528)
Aplicación de técnicas de inteligencia artificial en la prevención de problemas de salud mental.	Pág (529-533)
Soluciones de IOT: ámbitos de aplicación y desafíos.	Pág (534-538)
Machine learning y deep learning en la interpretación del lenguaje de señas.	Pág (539-543)
Diseño de una estrategia para la gestión del tiempo en la comunicación de corrección de fallas en el mantenimiento industrial.	Pág (544-548)
Implementación de una BCI mediante el diseño y desarrollo de Serious Games para la rehabilitación de pacientes neurológicos.	Pág (549-553)
La Inteligencia Artificial y su aplicación a los servicios de Internación Domiciliaria.	Pág (554-558)
Abordajes científico tecnológicos basados en el uso de vants.	Pág (559-563)
Innovación tecnológica a través de la hiper automatización.	Pág (564-568)
Tecnologías de Datos Espaciales, Visualización y Realidad Virtual.	Pág (569-573)
PDP – Procesamiento Distribuido y en Paralelo	Pág (574-636)
Paralelización de algoritmos y evaluación de rendimiento en Plataformas de cómputo de altas prestaciones.	Pág (575-580)
Análisis y modelado de sistemas utilizando Cloud Computing, IaaS privados y públicos.	Pág (581-585)
Estrategias de modelado y simulación en sistemas de HPC y salud.	Pág (586-590)
Arquitecturas Multiprocesador en HPC: Software de Base, Modelos y Aplicaciones.	Pág (591-596)
Arquitecturas Edge-Fog-Cloud en Procesamiento Distribuido. Aspectos de Eficiencia y Resiliencia.	Pág (597-602)
Soporte Serverless para Aplicaciones Móviles de Nueva Generación.	Pág (603-607)
Predicción de fenómenos naturales mediante paralelismo, algoritmos evolutivos y búsqueda por novedad.	Pág (608-612)
Simulación y análisis eficiente de grandes volúmenes de datos.	Pág (613-617)
Autoescalado basado en Aprendizaje Profundo por Refuerzo de Workflows Científicos en la Nube.	Pág (618-622)
Análisis de rendimiento de aplicaciones y sistemas complejos mediante computación paralela y técnicas de optimización de recursos.	Pág (623-627)
Técnicas de Recuperación y Procesamiento de Grandes Volúmenes de Datos Sísmicos. Un Repositorio Público Basado en Serverless.	Pág (628-631)
Aplicaciones para la salud y métricas de rendimiento en sistemas paralelos y distribuidos.	Pág (632-636)

PSSTR – Procesamiento de Señales y Sistemas en Tiempo Real	Pág (637-655)
Avances en el desarrollo de un prototipo experimental de segmento terreno Satelital multiplataforma-multimisión UGS UNLaM.	Pág (638-641)
Partición de Visual SLAM para su uso con celulares.	Pág (642-646)
Implementación y Simulación de Sistemas de Tiempo Real.	Pág (647-650)
Captura de variables para alerta temprana de incendios forestales y su almacenamiento para integración a sistemas de predicción mediante utilización de redes inalámbricas de sensores.	Pág (651-655)
RCCI – Redes de Cooperación Científica Internacionales	Pág (656-666)
Maestría en Gestión y Tecnología de Ciudades Inteligentes Resultado del Proyecto CAP4CITY.	Pág (657-660)
Avances del Consorcio para la colaboración en I+D+I en Temas de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET).	Pág (661-666)
SI – Seguridad Informática	Pág (667-726)
Estudio de las propiedades criptológicas de los Tokens No Fungibles Inteligentes para asegurar dispositivos de Internet de las Cosas.	Pág (668-673)
Análisis de las Herramientas para Realizar Pruebas Estáticas de Seguridad de las Aplicaciones.	Pág (674-679)
Tecnología Blockchain Aplicada a la Ciberdefensa.	Pág (680-683)
Aprendizaje por Refuerzo Aplicado al Etiquetado de Tráfico de Red.	Pág (684-688)
Criptografía Liviana y Ciberseguridad aplicados a Sistemas Ciberfísicos.	Pág (689-693)
Software para recolección de evidencias digitales rápidas en sistemas Windows.	Pág (694-698)
Protección de activos de software.	Pág (699-703)
Técnicas para Incrementar la Seguridad en Web Services basados en WSDL.	Pág (704-708)
Avances sobre la Especificación Integral del Sistema OTP-Vote Orientada a su Implementación.	Pág (709-713)
Criptología Maliciosa para la Ciberdefensa.	Pág (714-718)
Métodos, Técnicas y Herramientas para la Protección de Sistemas de Software.	Pág (719-723)
La blockchain y la certificación de documentos en Argentina.	Pág (724-726)
TD – Tesis Doctorales	Pág (727-775)
Generalización del modelado de cadencias de tecleo con contextos finitos para su utilización en ataques de presentación y canal lateral.	Pág (728-737)
Cálculo Científico Distribuido sobre Clientes Móviles Indeterminados en redes MANETs.	Pág (738-747)
DepProMod: Modelo de Proceso de Despliegue de Sistemas de Software.	Pág (748-757)
Tesis Doctoral. Métricas de calidad para validar los conjuntos de datos abiertos públicos gubernamentales.	Pág (758-767)
Modelo de madurez para servicios de gobierno electrónico en el ámbito universitario.	Pág (768-775)
TIAE – Tecnología Informática Aplicada en Educación	Pág (776-897)
Tecno pedagogía en el Proceso Formativo.	Pág (777-781)
Estrategias de accesibilidad para una educación inclusiva en la UNSAdA.	Pág (782-786)
Diseño de herramientas tecnológicas en Modelos educativos híbridos.	Pág (787-791)
Recursos educativos accesibles. Estudio de distintas técnicas y herramientas basadas en inteligencia artificial.	Pág (792-797)
Experiencia de Usuario en Sistemas Interactivos usando reconocimiento de Emociones.	Pág (798-802)

Un bot virtual en Moodle para asistir y complementar actividades académicas en espacios de educación superior.	Pág (803-807)
Interacción Persona Ordenador y Educación. Desarrollos y experiencias.	Pág (808-813)
Realidad aumentada en contextos educativos y su relación con el rendimiento académico universitario.	Pág (814-815)
Tecnologías digitales aplicadas al escenario educativo. Desarrollos y experiencias.	Pág (816-820)
Tecnologías disruptivas y aulas emergentes en el nivel medio y superior en el contexto de la UNRN.	Pág (821-825)
Desarrollo y aplicación de laboratorios virtuales destinados a educación de grado y pregrado.	Pág (826-830)
La mejora del aprendizaje en AVAs aplicando analítica del aprendizaje.	Pág (831-835)
La incorporación de la Ciencia de Datos como estrategia transversal para reforzar la enseñanza de la programación en el aula.	Pág (836-840)
Realidad Virtual y Educación.	Pág (841-845)
Diagnóstico de competencias genéricas universitarias, a través de software business intelligence.	Pág (846-850)
Revisión de alternativas moodle para la extracción de indicadores de autorregulación del aprendizaje.	Pág (851-855)
Realidad virtual para la enseñanza de portugués: Una guía para la implementación de contenido educativo en estudiantes de nivel medio.	Pág (856-860)
Definición de Lenguajes Específicos del Dominio para la Generación de Sitios Web de Instituciones Escolares.	Pág (861-865)
Definición de Lenguajes Específicos del Dominio para la Generación de Sitios Web de Instituciones Escolares.	Pág (866-869)
Minisatélite CANSAT para la enseñanza de las ciencias experimentales en diferentes contextos educativos.	Pág (870-874)
Calidad educativa: instrumentos para alcanzar la mejora continua de la educación superior.	Pág (875-879)
Evaluación del desempeño docente en el nivel superior.	Pág (880-884)
Software para gamificar terapias de rehabilitación AD-HOC.	Pág (885-890)
Tecnologías innovadoras aplicadas en contextos educativos	Pág (891-894)
Alfabetizando: innovación en el proceso de enseñanza aprendizaje en escuelas secundarias usando visualización de datos para desarrollar pensamiento computacional en materias de Humanidades y Ciencias Sociales.	Pág (895-897)
Autores:	Pág (898-901)



ARSO

Arquitectura, Redes
y Sistemas Operativos

Desafíos y buenas prácticas para pequeños proveedores de Internet zonales

Peliza Carlos, Dufour Fernando, Serra Ariel, Micieli Gustavo, Paz Carla

Departamento de Informática e Investigaciones Tecnológicas

Universidad Nacional de La Matanza

Florencio Varela 1908, San Justo

cpeliza@unlam.edu.ar; fdufourf@hotmail.com ; aserra@unlam.edu.ar ;
gmicieli@gmail.com

RESUMEN

Este proyecto de investigación pretende generar una capacitación para alumnos interesados en las redes, asesorar a quienes decidan mejorar o crear redes, conformar un grupo de alumnos con capacidad de actuar como consultores en redes FTTH y generar un informe final en el que entre otros aspectos se detallaran las mejores prácticas para la implementación y mantenimiento de una red para pequeños proveedores de internet, utilizando tecnologías disponibles en el mercado y servicios actuales. Se analizarán para cada elemento de la red al menos 3 proveedores diferentes con un análisis de cada uno de ellos.

El impacto que producirá la investigación será en el entorno de pequeñas empresas que brindan el servicio de internet, tendrán una guía para determinar qué tipos de implementaciones son factibles para mejorar sus servicios actuales de comunicaciones. En el ámbito académico, el impacto de la investigación estará dado por la incorporación de un tema de punta en las cátedras de Redes y Comunicaciones del Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLAM.

La metodología para utilizar será la investigación documental, entrevistas a especialistas del mercado que son gerentes o socios de pequeños proveedores de internet, pruebas de configuración y funcionamiento, análisis de alternativas de configuración de redes y diferentes modelos de despliegue de redes de acceso por fibra óptica (FTTH).

Se buscarán las incompatibilidades y a partir de un debate interno se sacarán las

conclusiones respecto de las mejores prácticas que se contrastarán con pruebas hechas por los integrantes del equipo de investigación en una empresa proveedora de internet.

Palabras Clave: Telecomunicaciones, ISP, FTTH, OLT, ONT.

CONTEXTO

La secretaria de investigaciones tecnológicas, en conjunto con el Departamento de Ingeniería e investigaciones tecnológicas, son quienes dictan las políticas de investigación y las entienden como la base de producción y distribución de conocimiento. Particularmente este grupo de investigación ha venido desarrollando una serie de investigaciones orientadas a las redes de comunicaciones o a las telecomunicaciones en general, como antecedentes se pueden citar: C164 (Carrier- Ethernet), C189 (LTE) y trabajos en diferentes congresos como “Carrier Aggregation”, “Narrow Band IoT”, “VoLTE”.

Temáticas que luego son introducidas en cátedras de comunicaciones de la Universidad Nacional de La Matanza, en carreras de Informática como de Electrónica. Adicionalmente, se brindan charlas de divulgación para la comunidad educativa de acuerdo con las necesidades del centro de estudiantes de dicha institución

1. INTRODUCCION

Las empresas de comunicaciones más grandes del país centran sus despliegues de red y, por consiguiente, las ofertas de conectividad y servicios en las zonas urbanas de rápido retorno de inversión y gran rentabilidad.

Fuera de los grandes centros urbanos tenemos un mercado que en gran parte está

abierto por pequeñas compañías que brindan el servicio de acceso a internet tanto por acceso inalámbrico como también por fibra óptica, ayudado por la mejora en el acceso a las nuevas tecnologías como por ejemplo FTTH (Fiber To The Home).

Actualmente, proyectos nacionales coordinados por ENACOM (Ente Nacional de Comunicaciones) se destacan como motivadores para el desarrollo y creación de estas pequeñas empresas de acceso a la conectividad. Un ejemplo de estos programas es el que brinda capital para la creación de redes de acceso a las instituciones públicas, desarrollado en conjunto con entidades municipales denominado PROGRAMA DE ACCESO A CONECTIVIDAD PARA INSTITUCIONES PÚBLICAS.

Las expectativas del sector de las comunicaciones, desde la visión de las grandes empresas se ven envueltas en un contexto sociopolítico complejo, los costos para mantener los servicios actuales son altos y el aumento del precio de servicio está regulado por el Estado. En resumen, para las grandes Telcos, el costo operativo aumenta en mayor proporción que los ingresos que se generan con lo cual ante este escenario, deciden no invertir más de lo necesario para mantener el servicio actual.

La situación descripta anteriormente, permitiría la creación de ecosistemas zonales de acceso a Internet, donde emprendedores sean capaces de desarrollar pequeñas empresas de servicio de Internet que actúen en comunidad con las instituciones públicas brindando la mencionada conectividad a precios y costos razonables. Sin embargo, la curva de aprendizaje necesaria para crear y operar la red puede resultar en una barrera infranqueable para una empresa que se inicia en redes y llevar al fracaso a cualquier proyecto.

Los sistemas GPON se caracterizan en general por un sistema de terminación de línea óptica (OLT, optical line termination) y una unidad de red óptica (ONU, optical network unit) o una terminación de red óptica (ONT, optical network termination) con una red de

distribución óptica (ODN, optical distribution network) pasiva que los interconecta. Por lo general, existe una relación de tipo uno a muchos entre la OLT y las ONU/ONT respectivamente.

La característica principal de GPON es el uso de divisores pasivos que permiten que un solo conductor óptico de alimentación, tendido desde la ubicación central del proveedor de la red, sirva a múltiples usuarios dentro de sus hogares y empresas. (ITU-T, 2003)

Las redes con GPON son la solución perfecta para entornos con varios nodos/puntos o edificios separados porque la tecnología permite dividir la tasa de alimentación de bits de manera proporcional al deseo de cada cliente y ofrece una red de fibra personalizable de alta capacidad para formas de servicios basados en IP.

La siguiente figura, extraída de la ITU-T G.984.1 muestra una de las configuraciones recomendadas

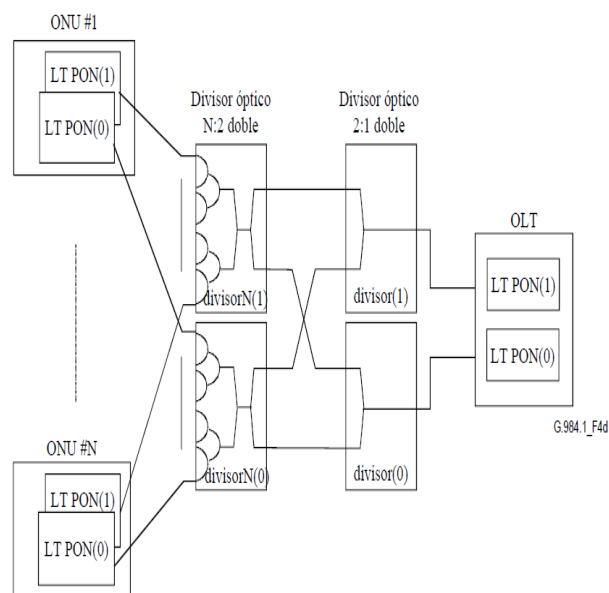


Ilustración 1 - Configuración recomendada con protección de caminos ópticos luego de la primer distribución

En ella, se distingue a la izquierda el componente hogareño (ONT o HGU según las siglas en inglés usadas) y a la derecha la OLT.

El predominio de fabricantes con casa central en China presenta varias ventajas y

desventajas que para el caso del inicio de una empresa de provisión de servicios de conectividad son factores para tener en cuenta, más allá de los obvios posibles problemas y demoras con la importación del producto se puede mencionar.

- Falta de soporte técnico en sitio, ante fallos desconocidos, la necesidad de contar con expertos en la materia tiene la complejidad de la diferencia horaria entre Asia y América.
- La interoperabilidad entre fabricantes siempre es mencionada como un escollo difícil de sortear y en particular cuando problemas de costos ajustados puede dar por tierra con el emprendimiento.

La encuesta realizada por este grupo de investigación es consistente con lo antedicho, puede notarse que quienes fueron consultados estiman un horizonte de crecimiento de sus redes

Tabla 1 Estimación sobre el futuro de la red de acuerdo a la pertenencia del encuestado

Tipo de empresa	Estimación del estado de la red	Porcentaje de respuestas
Empresa grande	Crece	50%
	Se mantiene	50%
Multinacional	Crece	90,91%
	Se mantiene	9,09%
Pyme	Crece	100%
	Se mantiene	0%

2. LINEAS DE INVESTIGACION y DESARROLLO

La problemática que investigar incluye, entre otras cosas, las percepciones de los principales especialistas en estas tecnologías y lo que ocurrirá con la misma durante los próximos 3 años.

Surgirán de los estudios distintas soluciones posibles de implementación, donde cada una de ellas llevará asociadas ventajas y limitaciones. Estas determinarán los servicios que cada caso pueda prestar.

El campo específico de la investigación consiste en trabajar sobre los aspectos de

interoperabilidad y capacidades de los elementos de red de distintos proveedores, determinando su compatibilidad de funcionamiento frente a los estándares, concentrándonos en el comportamiento en este escenario tecnológico.

Además de la investigación tecnológica propiamente dicha, es importante realizar vinculación entre diferentes departamentos universitarios y con la comunidad, por tal razón una de las líneas de investigación se halla ubicada dentro del campo de las ciencias económicas, concretamente la conformación de una empresa/sociedad dedicada a brindar servicios TIC.

Siguiendo con las líneas de investigación y desarrollo, es propuesta de este grupo de trabajo conformar grupos de alumnos con capacidad emprendedora que quieran incursionar en las TIC, para ello propondremos modelos de crecimiento y expansión que sirvan como primeros pasos para una empresa inserta en su comunidad.

3. RESULTADOS OBTENIDOS/ESPERADOS

Resultados en cuanto a la producción de conocimiento:

Obtener Know How sobre el estado de arte real de la tecnología y sus posibilidades de implementación de proveedores zonales en Argentina.

Incentivar la conformación de PYMES que trabajen sobre la tecnología FTTH

Lograr reconocer las implementaciones más convenientes (mejores prácticas) que podrán ser utilizadas por empresa actuales y futuras de la Argentina.

El desarrollo de saberes prácticos permitirá la provisión de servicio de asesoramiento a empresas e incluso la constitución de ellas.

El mercado podrá obtener información de mejores prácticas verificadas en la presente investigación. Las pequeñas empresas que provean internet se verán beneficiados obteniendo información relevante para iniciar u optimizar las redes con un objetivo orientado

a redes de acceso de fibra FTTH que les permita competir con empresas de mayor envergadura.

Los emprendedores que deseen brindar el servicio de FTTH se verán beneficiados con la experiencia de las mejores prácticas para este tipo de negocio.

4. FORMACION DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente del Departamento de Ingeniería de la Universidad de La Matanza, con miembros de las carreras de Electrónica e Informática, todos con gran experiencia en el área de telecomunicaciones.

Adicionalmente, se ha incorporado una alumna de la carrera de Ciencias Económicas para una prueba piloto donde compartir experiencia con carreras no dedicadas a las ingenierías.

Para el segundo año de investigación, se tiene planificado incorporar grupos de alumnos, como parte de un plan para mejorar sus habilidades blandas o soft skills, que incluyan la redacción científica de textos y/o a los que la investigación les permita formalizar su trabajo de fin de carrera. Mientras que, por otro lado, se buscará incentivar el interés emprendedor en los alumnos y que las fuerzas vivas de la comunidad matancera, en principio, colaboren con quienes quieren emprender en la provisión de servicios de internet.

5. BIBLIOGRAFIA

Cámara Argentina de Internet. (22 de Julio de 2022). Obtenido de <https://www.cabase.org.ar/2020-internet-index-2/>

Corning. (3 de 9 de 2022). Obtenido de <https://www.corning.com/optical-communications/worldwide/en/home/products/fiber-optic-cables.html>

ENACOM. (10 de Julio de 2022). *Datos abiertos.* Obtenido de <https://datosabiertos.enacom.gob.ar/daviews/252830/accesos-a-internet-fijo-por-tecnologia-y-localidad/>

Furukawa. (8 de 8 de 2022). Obtenido de <https://www.efurukawa.com/ar/sobre-a-furukawa/>

Grazzini, H. O. (2021). *Fibras ópticas: Conceptos teóricos y aplicaciones prácticas* (COMUNICACIÓN TELECOMUNICACIÓN | SISTEMAS Y TEORIA n° 7) (Spanish Edition). EDITORIAL CIENTÍFICA UNIVERSITARIA DE CÓRDOBA.

Huawei. (1 de 6 de 2022). Obtenido de <https://www.huawei.com/>

Inspeccion General de Justicia. (1 de 12 de 2022). Obtenido de <https://www2.jus.gov.ar/igj-homonimia/Principal.aspx>

ITU-T. (2003). *Redes ópticas pasivas con capacidad de Gigabits: Características generales.*

Kramer, G. (2011). *Present state of standards for Ethernet PON systems.* .

MILLAN ESTELLER, J. M. (2014). *Configuración de infraestructuras de sistemas de telecomunicaciones Tapa blanda.* Ediciones Paraninfo, S.A.

Mordor Intelligence. (6 de 12 de 2022). Obtenido de <https://www.mordorintelligence.com/es/industry-reports/passive-optical-network-pon-equipment-market#:~:text=Se%20espera%20que%20el%20mercado,en%20el%20per%C3%ADodo%20de%20pron%C3%B3stico.>

Peliza, C., & Veloso, J. (2017). *Curso introductorio de redes informáticas.* San Justo: UNLaM.

MIGRACIÓN CONTROLADA DE PROCESOS EN SISTEMAS DISTRIBUIDOS

Marecos Brizuela, Terecio Diosnel^a; Agostini, Federico^b; La Red Martínez, David^b

^a Facultad de Ciencias Aplicadas -Universidad Nacional de Pilar

^b Facultad de Ciencias Exactas y Naturales y Agrimensura – Universidad Nacional del Nordeste
ecosmar097@hotmail.com

RESUMEN

En los sistemas distribuido es necesario gestionar la asignación de recursos a procesos garantizando la exclusión mutua en el acceso a recursos compartidos cuando ello es requerido y distribuyendo la carga de procesamiento y de transmisión de datos, buscando la autorregulación del sistema. Así surge el siguiente interrogante: ¿cuáles son los modelos de decisión y los operadores de agregación que habrá que generar partiendo de los desarrollados en el PI 16F001 de la UNNE (“Modelos de decisión y operadores de agregación para la administración de procesos en sistemas distribuidos”), incorporando la utilización de conjuntos de etiquetas lingüísticas para expresar las variables de estado de los nodos, contemplando la posibilidad de que la información de control proveniente de algún nodo no llegue al nodo coordinador, por lo que habría que aplicar imputación de datos y, considerando la posibilidad de migrar procesos entre nodos con el propósito de autorregular el sistema balanceando su carga de trabajo y de transmisión de datos?. Para ello habrá que considerar la utilización de etiquetas lingüísticas, la necesidad de imputación de datos bajo determinadas condiciones, y la posibilidad de migración de procesos entre nodos. Los modelos de decisión considerarán la posibilidad de utilizar operadores ponderación, incluyendo los de la familia OWA (Ordered Weighted Averaging) (Yager, 1988, 1993), buscando generar operadores de agregación específicos.

Palabras claves: sistemas distribuidos, migración de procesos, balanceo de carga, imputación de datos, etiqueta lingüística.

CONTEXTO

Los modelos de decisión en la actualidad disponibles y generalmente aplicables en los sistemas distribuidos se basan en algoritmos de intercambio de permisos que intentan lograr un

acuerdo de todos los procesos intervinientes para realizar determinadas acciones, como el acceso a un área de memoria compartida o a registros de archivos compartidos, recursos a los que frecuentemente se debe acceder en la modalidad de exclusión mutua.

Se considera especialmente importante estudiar la aplicación de modelos de decisión y operadores de agregación para la asignación de recursos a procesos, considerando no sólo las exigencias de la exclusión mutua cuando ella es requerida, sino también, el estado global del sistema, de tal manera que las decisiones tomadas respecto de recursos y procesos contribuyan a equilibrar la carga de trabajo y de transmisión de datos. El estado global del sistema está integrado por información de control proveniente de los distintos nodos, la que podría expresarse de manera numérica o lingüística, considerando la posibilidad de carencia de información de control de algún nodo y también la necesidad de migrar procesos entre nodos con el propósito de equilibra la carga global del sistema y mejorar así su rendimiento, buscando su autorregulación.

El punto de partida de esta investigación está dado por los modelos de decisión y operadores de agregación desarrollados en el marco del PI 16F001 de la UNNE, a los que se propone adaptar a distintos escenarios, considerando la utilización de información lingüística, necesidad de aplicar imputación de datos, conveniencia de aplicar migración de procesos y combinación de los escenarios anteriores. El marco para estas actividades está dado por el PI 20F005 de la UNNE (“Modelos de decisión para gestión de recursos y procesos en sistemas distribuidos considerando migración de procesos, imputación de datos y lógica difusa en los nuevos operadores de agregación”) y el PI 126-2020 de la UNCAus (“Desarrollo de un simulador para la evaluación de algoritmos clásicos y nuevos para la gestión

de recursos compartidos en sistemas distribuidos contemplando exclusión mutua”).

En los sistemas computacionales de procesamiento distribuido es necesario gestionar de manera eficiente la asignación de recursos a procesos, respetando las restricciones de exclusión mutua, teniendo en cuenta las prioridades de los procesos y el balanceo de la carga de procesamiento y de transmisión de datos, buscando la autorregulación del sistema distribuido considerado como un todo.

Los modelos de decisión y sus operadores de agregación propuestos para resolver las cuestiones indicadas precedentemente deben intercambiar información de control generada por los distintos nodos, deben enfrentar situaciones de falta de determinada información y deben resolver problemas de posibles sobrecargas de procesamiento y/o de transmisión de datos entre los nodos.

Ante la situación mencionada precedentemente surge la necesidad de nuevos modelos de decisión y los operadores de agregación correspondientes incorporando una visión global superadora de los modelos clásicos, para los siguientes tipos de situaciones o escenarios:

1) Que los nodos intercambien información con el nodo coordinador indicando los valores de sus variables de estado y prioridades de manera numérica o lingüística, mediante conjuntos de etiquetas no necesariamente homogéneos, por lo que habrá que aplicar traslación simbólica para que finalmente cada nodo reciba la información acorde a su propio conjunto de etiquetas lingüísticas.

2) Que la información de control proveniente de algunos nodos y necesaria para la operación del nodo coordinador central no esté disponible cuando sea requerida, lo que obligará a aplicar técnicas de imputación de datos faltantes, bajo determinadas condiciones.

3) Que a los efectos de balancear la carga de procesamiento y de transmisión de datos entre los nodos, sea necesario migrar procesos entre nodos, lo cual requerirá establecer métricas que permitan seleccionar el mejor esquema de migración posible en un momento dado.

4) Que los escenarios mencionados precedentemente se presenten concurrentemente como una combinación de dos o más de los escenarios planteados.

INTRODUCCIÓN

La proliferación de sistemas informáticos, muchos de ellos distribuidos, en los cuales existen múltiples procesos que cooperan para el logro de una determinada función, hace necesario disponer de modelos de decisión y operadores de agregación que permitan una adecuada asignación de recursos a procesos, respetando la exclusión mutua en el acceso a recursos compartidos que así lo requieran, buscando lograr un eficiente funcionamiento del sistema distribuido mediante mecanismos de autorregulación y balanceo de carga de procesamiento y de transmisión de datos.

En dichos sistemas distribuidos es deseable considerar soluciones en las cuales los nodos del mismo puedan indicar al nodo coordinador central el valor de sus variables de estado no sólo de manera numérica, sino mediante etiquetas lingüísticas, en cuyo caso cada nodo podría utilizar su propio y diferente conjunto de etiquetas, lo que implica trabajar con conceptos de traslación simbólica.

Otro aspecto esencial a tener en cuenta es la problemática originada por la falta de información de control proveniente de alguno o algunos de los nodos del sistema distribuido, que es necesaria para la operación del modelo de decisión y de los operadores de agregación en el nodo coordinador central, no pudiendo determinarse en dicho coordinador si la información no ha llegado por una sobrecarga de trabajo en los nodos que debieron enviarla a tiempo, o por un problema de conectividad o por una caída definitiva de los nodos problemáticos. Ante esta situación, durante un determinado lapso, el modelo de decisión y sus operadores de agregación podrían seguir operando, reemplazando los datos faltantes por datos imputados por un mecanismo que habrá que determinar, siendo esta solución aplicable hasta que la información llegue al nodo coordinador o hasta que, agotado el tiempo de espera, en nodo coordinador asuma de los nodos “demorados” realmente están fuera de servicio.

Una cuestión adicional que debe ser considerada con el propósito de lograr un adecuado balanceo de la carga de procesamiento y de transmisión de datos, buscando la autorregulación del sistema, es la migración de procesos, que podrán ser trasladados para su ejecución en nodos diferentes de los nodos originarios de los mismos, bajo determinadas condiciones que habrá que analizar, utilizando métricas específicas que se definirán al respecto.

Finalmente, también es deseable buscar soluciones donde los tres escenarios mencionados precedentemente se mezclen de diferentes maneras, tratando de cubrir la mayor cantidad posible de problemáticas que se podrían generar.

Es preciso señalar que la problemática de los sistemas distribuidos y la gestión de recursos en los mismos ha sido tratada ampliamente por numerosos autores, pero generalmente el enfoque ha estado centrado en las prioridades primarias u originales de los procesos y en buscar una forma eficiente de intercambio de mensajes, asegurando la exclusión mutua, sin considerar un enfoque integrador como el que se propone para esta investigación.

Ejemplos de lo mencionado se encuentran en (Tanenbaum, 1996 y 2009), donde se describen los principales algoritmos de sincronización en sistemas distribuidos, en (Agrawal et al., 1991), donde se presenta una solución eficiente y tolerante a fallas para el problema de la exclusión mutua distribuida, en (Ricart et al., 1981), (Cao & Singhal, 2001) y en (Lodha & Kshemkalyani, 2000), donde se presentan unos algoritmos para gestionar la exclusión mutua en redes de computadoras, en (La Red Martínez, 2004), donde se describen los principales algoritmos de sincronización en sistemas distribuidos, en (Stallings, 2005), donde se detallan los principales algoritmos para la gestión distribuida de procesos, los estados globales distribuidos y la exclusión mutua distribuida. Estos temas y otros relacionados también han sido tratados en (Joshi & Holzmann, 2007), (Alagarsamy, 2003), (Fitzek & Katz, 2014), (La Red Martínez, 2017), etc.

LINEA DE INVESTIGACIÓN Y DESARROLLO

Con respecto a la imputación de datos en la gestión de recursos compartidos en sistemas distribuidos, en esta propuesta se considera que hay un nodo central que se encarga de recibir y mantener actualizada la información de control de todos los nodos, y es el encargado de la asignación de recursos en la modalidad de exclusión mutua, asegurando la disponibilidad de los mismos y respetando las prioridades de los procesos. En un determinado ciclo de recolección de información de gestión, necesaria para ejecutar y asegurar lo mencionado anteriormente, el nodo central puede recibir de alguno de los nodos información de control incompleta o inexistente, por ejemplo, de los criterios de evaluación de carga nodal, de las prioridades o preferencias de los procesos, etc. Estos datos faltantes, constituyen un obstáculo importante en la gestión de recursos. Las técnicas de imputación de datos permiten estimarlos utilizando diferentes algoritmos, mediante los cuales se puede imputar una característica importante para una instancia en particular.

Con respecto a la gestión del tráfico en redes de datos utilizando etiquetas lingüísticas y 2-tuplas, en este trabajo de investigación, la propuesta consistió en implementar un nuevo modelo de decisión que permita gestionar el control del tráfico en redes de datos en Internet, siendo la información de control representada con etiquetas lingüísticas y 2-tuplas. En el tráfico de redes es necesario que los nodos deban tomar decisiones basados en acuerdos respecto del acceso a rutas disponibles; las decisiones pueden estar relacionadas con el estado de los nodos de acuerdo con criterios y aspectos de seguridad.

En la propuesta de migración controlada de procesos en sistemas distribuidos se propone un modelo de migración controlada de procesos para la gestión de recursos y procesos en sistemas distribuidos considerando aspectos de balanceo de carga, consumo de energía y requisitos técnicos específicos, que permitirá resolver los inconvenientes relacionados a la sobrecarga que pueden presentarse en los distintos nodos de un sistema distribuido.

RESULTADOS ESPERADOS/OBTENIDOS

En el modelo de decisión propuesto por (Duré Attis, 2021a), considera la imputación de datos

cuando la información sobre las variables que indican el estado de carga de alguno o algunos de los nodos no llega en su totalidad al Runtime central, reemplazando esos valores faltantes por valores estimados, para que finalmente el modelo de decisión establezca un correcto orden de asignación de recursos a los procesos, respetando la exclusión mutua. Es un método general, que permite ser adaptado a diferentes circunstancias y objetivos, de acuerdo con las necesidades de los nodos, cargas de trabajo, prioridades, etc. El método de imputación utilizado es considerado uno de los más fiables y de menor consumo de recursos computacionales. Es importante mencionar que la solución presentada es compatible con la gestión de transacciones globales habitualmente implementada en los sistemas de gestión de base de datos.

En la línea de investigación de (Silva, 2021a), se consideraron diferentes situaciones relacionadas con el hecho de despachar o no los paquetes, con criterios para el control de la congestión del tráfico en redes. Además, se crearon etiquetas lingüísticas que permitieron clasificar la información del estado de los nodos, paquetes, enlaces, etc. Los modelos desarrollados se evaluaron comparando sus características con los modelos habitualmente utilizados. Esto permitió desarrollar mecanismos para optimizar la gestión del tráfico en redes de datos en Internet, utilizando etiquetas lingüísticas y 2-tuplas para la toma de decisiones, dando un enfoque cualitativo a la información.

Se prevé desarrollar un simulador en el que se consideren los distintos escenarios posibles para permitir al sistema predecir, comparar y optimizar el comportamiento de sus procesos simulados en un tiempo muy breve sin el coste ni el riesgo de llevarlos a cabo, haciendo posible la representación de los procesos, recursos y nodos en un modelo dinámico. Con la ayuda del correspondiente soporte informático, el modelo de simulación permitirá la capacidad de considerar complejas tareas interrelacionadas y proyectarlas mediante la realización de muchas combinaciones alternativas en cuestión de segundos. Además, la interacción de los recursos con los procesos se traducirá en un gran número de escenarios y de posibles resultados imposibles

de abarcar y valorar sin la ayuda de un modelo de simulación computarizado.

En (Marecos Brizuela, 2023) se espera generar un modelo de decisión en un Sistema Distribuido agregando la migración controlada de procesos a un operador de agregación, que permita la redistribución de la carga de trabajo entre los nodos del sistema distribuido, para mejorar la utilización de los recursos, el tiempo de respuesta de los procesos, la disminución del consumo de energía y la mejora del rendimiento.

En el marco del proyecto 20F005 se han finalizado dos tesis de maestría (Duré Attis, 2021a) y (Silva Ruiz, 2021a) y se han realizado diferentes publicaciones en revistas y congresos (Forneron Martínez et al., 2021, 2023) (Dure Attis et al., 2021b, 2021c) (Silva Ruiz et al., 2021b).

FORMACION DE RECURSOS HUMANOS

El equipo de trabajo está conformado por un director, un codirector, dos docentes investigadores y tres tesis de posgrado. En este contexto se ha propuesto el desarrollo de tres tesis correspondientes a la Maestría en Informática y Computación, de la Facultad de Ciencias Aplicadas de la Universidad Nacional de Pilar (Paraguay), de las cuales han finalizado dos y una está en la etapa final. Los trabajos finalizados son de Diego David Duré Attis con el tema “Imputación de datos en la gestión de recursos compartidos en sistemas distribuidos”, y César Alberto Silva Ruiz con el tema “Gestión de redes de datos utilizando etiquetas lingüísticas y 2-Tuplas”. La tesis de maestría en curso, de Terecio Diosnel Marecos Brizuela, tiene por título “Migración controlada de procesos en sistemas distribuidos”.

REFERENCIAS

- Agrawal, D., & El Abbadi, A. (1991). “An Efficient and Fault-Tolerant Solution for Distributed Mutual Exclusion.” *ACM Transactions on Computer Systems (TOCS)* 9(1). doi: 10.1145/103727.103728.
- Alagarsamy, K. (2003). “Some Myths about Famous Mutual Exclusion Algorithms.” *ACM SIGACT News* 34(3). doi: 10.1145/945526.945527.
- Cao, G., & Singhal, M. (2001). “A Delay-

- Optimal Quorum-Based Mutual Exclusion Algorithm for Distributed Systems.” IEEE Transactions on Parallel and Distributed Systems 12(12). doi: 10.1109/71.970560.
- Duré Attis, D. D. (2021a). Imputación de datos en la gestión de recursos compartidos en sistemas distribuidos. Maestría en Informática y Computación. Facultad de Ciencias Aplicadas. Universidad Nacional de Pilar.
- Duré Attis, D. D., Agostini, F., & La Red Martínez, D. L. (2021b). “Imputación de información de control faltante en la gestión de recursos y procesos”. Conferencia Ibero Americana WWW/Internet, (CIAWI 2021).
- Duré Attis, D. D., Agostini, F., & La Red Martínez, D. L. (2021c). “Imputación de valores faltantes sobre la información de control en el contexto de los sistemas distribuidos”. 9no. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI)
- Fitzek, F. & Katz, M. (2014). Mobile Clouds: Exploiting Distributed Resources in Wireless, Mobile and Social Networks. L. John Wiley & Sons, ed., New Delhi, India. ISBN: 978-0-470-97389-9.
- Fornerón Martínez, J. T., Agostini, F., & La Red Martínez D. L. (2021). “Modelo de Decisión para Gestión de Procesos y Recursos en Sistemas Distribuidos con Balanceo Dinámico de Carga de Trabajo”. The 2021 International Conference on Information Technology & Systems (ICITS'21). RISTI - Revista Ibérica de Sistemas e Tecnologias da Informação. ISBN N° 1646-9895. Febrero 2021. Libertad city, Ecuador. Virtual Conference.
- Fornerón Martínez, J. T., Agostini, F., & La Red Martínez, D. L. (2023) “Resource and Process Management With a Decision Model Based on Fuzzy Logic”. Aceptado para publicar.
- Joshi, R., & Holzmann, G. J. (2007), A Mini-Challenge: Build a Verifiable Filesystem, Formal Aspects of Computing, Vol.19.
- La Red Martínez, D. L. (2004). Sistemas Operativos. EUDENE. Argentina.
- La Red Martínez, D. L. (2017). “Aggregation Operator for Assignment of Resources in Distributed Systems.” International Journal of Advanced Computer Science and Applications 8(10). doi: 10.14569/ijacsa.2017.081053.
- Lodha, S., & Ajay, K. (2000). “A Fair Distributed Mutual Exclusion Algorithm.” IEEE Transactions on Parallel and Distributed Systems 11(6). doi: 10.1109/71.862205.
- Marecos Brizuela, T. D. (2023). Migración controlada de procesos en sistemas distribuidos. Tesis en curso. Maestría en Informática y Computación. Facultad de Ciencias Aplicadas. Universidad Nacional de Pilar.
- Ricart, G., & Ashok, K. A. (1981). “An Optimal Algorithm for Mutual Exclusion in Computer Networks.” Communications of the ACM 24(1). doi: 10.1145/358527.358537.
- Silva Ruiz, C. A. (2021a). Gestión de redes de datos utilizando etiquetas lingüísticas y 2-Tuplas. Maestría en Informática y Computación. Facultad de Ciencias Aplicadas. Universidad Nacional de Pilar.
- Silva Ruiz, C. A., Agostini, F., & La Red Martínez, D. L. (2021b). “Gestión del Tráfico en Redes de Datos utilizando Etiquetas Lingüísticas y 2-Tuplas”. 9no. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI).
- Stallings, W. (2005). Sistemas Operativos Aspectos Internos y Principios de Diseño. Vol. Quinta Edición.
- Tanenbaum, A. S. (2009). Sistemas Operativos Modernos Tercera Edición.
- Tanenbaum, A. S., Guerrero, G., & Velasco, Ó. A. P. (1996). Sistemas Operativos Distribuidos. México: Prentice-Hall Hispanoamericana S.A.
- Yager, R. (1988). “On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking.” IEEE Transactions on Systems, Man and Cybernetics 18(1):183–90. doi: 10.1109/21.87068.
- Yager, R. (1993). “Families of OWA Operators.” Fuzzy Sets and Systems 59(2). doi: 10.1016/0165-0114(93)90194-M.

ESTUDIOS DEL APOORTE DE TECNOLOGÍAS COMO BLOCKCHAIN, NFT Y VPN A LA SEGURIDAD EN PLATAFORMAS DE IOT

Marianetti, Osvaldo^{1,4}; Godoy, Pablo Daniel^{1,2,3}; Chediak, Ernesto^{1,4}; Fontana, Daniel^{1,4}

¹Universidad Nacional de Cuyo, Facultad de Ingeniería

²Universidad Nacional de Cuyo, Facultad de Ciencias Exactas y Naturales

³Universidad Nacional de Cuyo, ITIC

⁴Universidad de Mendoza

osvaldo.marianetti@ingenieria.uncuyo.edu.ar

RESUMEN

Surgen interrogantes sobre quien tiene la responsabilidad de resolver los problemas de seguridad de los dispositivos de IoT conectados a Internet: ¿le corresponde al fabricante, al vendedor, al usuario?, ¿debe ser regulado por los gobiernos? No siempre queda claro quién es responsable de las decisiones de seguridad. Una compañía diseña un dispositivo, otra provee el software, otra opera la red en la que se lo integra y otra pone a disposición el equipo. No existen normas y estándares aceptados a nivel internacional.

Existe un conjunto amplio de vulnerabilidades en los dispositivos IOT como son: comunicaciones locales y remotas sin encriptar, almacenamiento sin cifrar, acceso remoto por Shell, cuentas ocultas, acceso a través de las interfaces UART y JTAG, entre otros.

Ante este escenario, se hace necesario encontrar alternativas de tecnologías para la solución de esta problemática. En este trabajo se propone investigar tecnologías como Blockchain, NFT (token no fungible) y redes privadas virtuales (VPN). Dichas tecnologías utilizadas con éxito en otras aplicaciones, se constituyen en una posible herramienta para que los entornos de los dispositivos utilizados en sistemas de IoT se beneficien en lo que refiere a la seguridad de los datos que procesan.

Palabras clave: plataformas de IoT, seguridad, vulnerabilidad, Blockchain, NFT, VPN.

CONTEXTO

Desde la Secretaría de Investigación, Internacionales y Posgrado de la UNCUIYO se desarrollan convocatorias a proyectos bienales

que tienen por finalidad la promoción de la investigación científica y tecnológica.

Este proyecto está inserto en la convocatoria Proyectos de investigación SIIP 2021.

Los investigadores que forman parte de estos proyectos son profesores de asignaturas como Arquitectura de Computadoras, Redes de Computadoras, Sistemas Embebidos y Arquitecturas Distribuidas en la Universidad Nacional de Cuyo y la Universidad de Mendoza, en carreras de pregrado, de grado y de posgrado.

1. INTRODUCCIÓN

El presente proyecto es una evolución del proyecto “Dispositivos de lógica programables como alternativa de solución en problemas de seguridad en Internet de las Cosas” presentado en la convocatoria de proyectos SIIP 2019. En dicho proyecto se investigó el aporte de los recursos disponibles en dispositivos programables y reconfigurables como las FPGAs en el proceso de diseño del hardware de sistemas embebidos presenten en los sistemas de aplicaciones de IoT. En este proyecto se pretende extender dicha investigación analizando el aporte de tecnologías como Blockchain, NFT y VPN a la seguridad en plataformas de IoT

La tecnología Blockchain tiene la capacidad de resolver una problemática en el mundo de los archivos digitales que carecen de la posibilidad de dar prueba de autenticidad. Esta característica es el motivo por el cual seguramente podrá aplicarse para muchos otros usos, no sólo para monetizar un activo, sino también para certificar diferentes tipos de trazabilidad. Esta tecnología es hoy una interesante alternativa que permite incorporar una capa de seguridad a determinadas

soluciones de adquisición de datos mediante dispositivos de Internet de las Cosas (IoT). Algunas ya se están explorando, por ejemplo: trazabilidad de medicamentos, de expedientes judiciales, supply chain, agroindustria, cadena de fabricación y otras.

La tecnología de Blockchain puede aportar una nueva perspectiva en este contexto, una visión descentralizada de los procesos y del sentido de propiedad. Esta visión colabora en la creación de nuevos escenarios híbridos, real-virtual, confiables y transparentes, donde los contratos inteligentes (smart-contracts) controlen que contratos de la vida real sean cumplidos. [1]

Desde hace un tiempo, dentro del ámbito blockchain, se hace referencia al concepto NFT (aunque en realidad existen conceptualmente desde el 2012). NFT es el acrónimo de Non Fungible Token (token no fungible). Un “Token” es un activo cuyo valor proviene de aquello que representa, y “no fungible” significa que no puede ser reemplazado por un activo similar del mismo valor, porque tiene características que lo hacen único. Por tanto, un NFT es un token digital único, no duplicable, ni divisible, que contiene información sobre su propietario, y que es registrado a través de contratos inteligentes (básicamente son fragmentos de código que están verificados y registrados con tecnología blockchain). Los NFTs pueden presentarse de diferentes formas dependiendo del estándar en el que se basen. Por ejemplo, ERC721 y ERC1155 para las NFT de Ethereum. Cada uno de estos estándares tiene sus propias ventajas y limitaciones, que pueden afectar a los tipos de NFT que pueden crearse. [2]

Algunos ejemplos para qué pueden llegar a utilizar NFTs son los siguientes:

Activos del juego. Los NFTs pueden representar objetos en el mundo gamer, como armas, potenciadores, vehículos, personajes, etc.

Obras de arte. Los NFTs pueden representar obras de arte individuales que han sido tokenizadas y ahora están representadas por un token único. La propiedad del NFT equivale a la propiedad de la obra de arte subyacente. [3], [4]

Básicamente, los NFTs nos permiten replicar en el mundo digital los conceptos de propiedad, unicidad y escasez del mundo físico. [5]

Un NFT podría usar "rastreadores" del mundo físico para demostrar a un contrato inteligente que estaba en un lugar específico en el mundo físico en un momento específico. El contrato inteligente podría desencadenar que suceda algo en el mundo digital, incluida la acuñación de un nuevo NFT que utilice datos del mundo físico, propiedad exclusiva del titular del dispositivo, como prueba. Este escenario podría ser útil para asegurar un dispositivo de IoT, ya que permitiría capturar y firmar criptográficamente datos del mundo físico, como la ubicación, el clima, el movimiento y la luz, utilizando un elemento seguro. Este concepto podría incluir datos de salud comprobables y seguros de dispositivos portátiles o datos automotrices de vehículos, todos vinculados a contratos inteligentes. [6]

Otra alternativa a considerar para proteger un sistema de los riesgos de seguridad presentes en internet, es utilizar una VPN (red privada virtual) para IoT.

Las VPN existen desde hace mucho tiempo y sólo recientemente se han vuelto disponibles para el sector minorista. Son una combinación de protocolos de seguridad que crean un túnel digital a través del cual los dispositivos pueden conectarse de manera segura a internet. Las conexiones son difíciles de rastrear, apareciendo como una conexión aleatoria y anónima que pasa a través de un ISP hacia algún lugar desconocido y casi imposible de tenerlo como objetivo. Además, todos los datos y toda la información enviada a través de la conexión están encriptados antes de salir del dispositivo. Los beneficios de la VPN son muchos. Los datos se cifran, las conexiones están ocultas y no pueden ser objetivo de malware o virus, y su dirección IP está enmascarada. [7]

Existen distintos factores que contribuyen a que los dispositivos IoT sean vulnerables: mezcla de malos diseños, entornos no regulados o regulados ineficientemente, entre otros. Los atacantes pueden usar las vulnerabilidades de estos dispositivos para

obtener acceso permanente a estos equipos y a la infraestructura de red por la cual están conectados. En muchos desarrollos de sistemas IoT se usan componentes de software de terceros (como librerías), que pueden incorporar vulnerabilidades.

La violación a la seguridad de los dispositivos IoT, puede radicar en:

(1) Los dispositivos IoT como víctimas y (2) los dispositivos IoT como herramientas de ataques.

2. LINEAS DE INVESTIGACION y DESARROLLO

2.1 Hipótesis del trabajo

En este trabajo se propone la siguiente hipótesis: Se puede optimizar la seguridad de los dispositivos utilizados en plataformas de IoT mediante la utilización de Blockchain, NFT (non fungible token) y/o VPN.

Se plantean los siguientes objetivos:

Objetivo general: Estudiar el impacto del uso de las tecnologías de Blockchain, NFT (non fungible token) y VPN en la optimización de la seguridad de los dispositivos de los sistemas de IOT.

Objetivos específicos:

- 1) Implementar una VPN en una aplicación que procese datos de sensores y analizar su comportamiento respecto de la integridad de los datos.
- 2) Estudiar los mecanismos de introducción de datos de en la tecnología Blockchain y NFT.
- 3) Analizar la factibilidad de uso de la capacidad de procesamiento remanente de los procesadores que integran los dispositivos de sensado en tareas inherentes al blockchain.

2.2 Medios y recursos disponibles

Se dispone del equipamiento adquirido mediante proyectos desarrollados. Este equipamiento incluye computadoras 6 Raspberry Pi 3 o Raspberry Pi 4, 15 nodos Arduino Uno y dispositivos sensores. También se cuenta con Routers y modems 3,5 G y software de desarrollo de aplicaciones, sistemas operativos y software de análisis de datos.

Se cuenta con las oficinas del ITIC, las cuales cuentan sala de reuniones y oficina de trabajo.

Los profesores que forman parte del proyecto son profesores de 3 asignaturas de la carrera Licenciatura en Ciencias de la Computación, dictada en la Facultad de Ingeniería de la Universidad Nacional de Cuyo. Dichas asignaturas son Arquitectura de Computadoras, Redes de Computadoras, Arquitecturas Distribuidas y Sistemas Embebidos, que se cursan entre segundo y tercer año de la carrera. Los estudiantes de estas asignaturas poseen los conocimientos técnicos necesarios para hacer aportes significativos al proyecto. Se espera que estudiantes de esta carrera participen del proyecto, por esto se los invitará a participar del mismo, pudiendo encuadrar su participación como proyectos finales de materias o como servicios personales, para lo cual podrá utilizarse parte de los fondos del proyecto para fomentar la participación.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1 Resultados obtenidos

Despliegue de una VPN que permite asignar una IP fija e introducir seguridad.

La VPN desplegada tiene como objetivo la adquisición de datos de radares sensores de tránsito para el procesamiento y el análisis de los mismos.

Los radares sensores de tránsito poseen la capacidad de almacenar estadísticas de tránsito recopiladas durante su funcionamiento. Estos datos son utilizados para realizar estudios de caracterización de tránsito, de las calles, avenidas y rutas, así como información de soporte a otros sistemas de seguridad. En virtud de la importancia para la toma de decisiones de alto costo, que tienen estos datos es necesario plantear esquemas de seguridad que permitan garantizar su integridad. Entre otros datos podemos tener: cantidad de vehículos que circularon, los horarios con mayor circulación, las velocidades máximas y promedios, etc. Para acceder a dicha información los radares pueden ser trabajados off line, recorriendo físicamente cada uno y descargando los archivos. Modalidad on line, en este caso por los lugares de instalación es necesario utilizar sistemas de comunicación

inalámbricos. Dentro de este esquema se decidió utilizar routers 3G con WIFI para conectar los radares a internet. A su vez, armar una VPN que permite asignar una IP fija e introducir seguridad para la integridad de estos datos sensibles.

Los radares se conectan a internet vía 3.5G brindada por un módem USB marca Huawei modelo E173 provisto de una tarjeta SIM con línea activa y datos móviles disponibles. Esta conexión se gestiona con un router marca Tp-link modelo TL-MR3020 versión 3.20 cuyo sistema operativo de fábrica se ha reemplazado por OpenWrt versión 22.03.2. Este sistema operativo permite la conexión como cliente a una red OpenVPN así como la conexión a internet mediante el modem 3.5G antes mencionado, además de otras funcionalidades útiles para esta aplicación. La conexión al servidor OpenVPN le permitirá al radar tener una IP fija conocida con la cual acceder al mismo para descargar la información deseada utilizando el software Houston Radar Stats Analyzer. (Software de análisis de datos de tránsito), compatible con los radares utilizados. La experiencia se realizó con los radares instalados en la comuna de Guaymallén, Mendoza. La actual implementación de esta experiencia cuenta con 9 nodos de adquisición de datos.

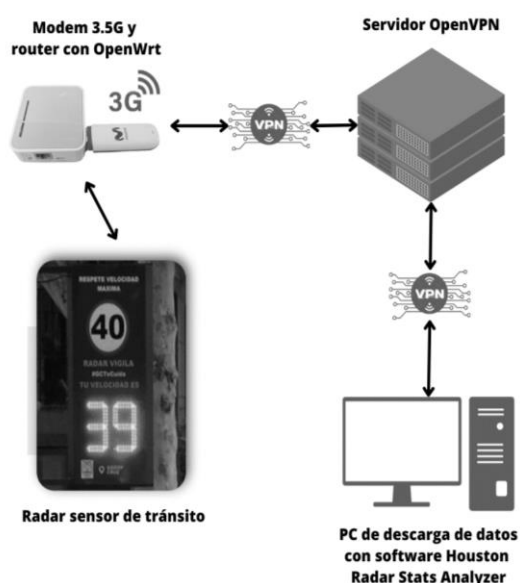


Figura 1: Implementación de la VPN desarrollada.

Como un ejemplo de los datos adquiridos y procesados, la figura 2 representa el número de vehículos promedio por hora y por semana de una autovía de ingreso a la ciudad de Mendoza.

3.2 Resultados esperados

1. Identificar los requerimientos de los recursos necesarios para la implementación operativa de estas soluciones
2. Comprobar los beneficios en relación a la seguridad de los dispositivos utilizados en plataformas de IoT mediante la utilización de Blockchain, NFT (non fungible token) y/o VPN.
3. Redacción y publicación de artículos de acuerdo al grado de avance y los resultados alcanzados en el desarrollo del proyecto

4. FORMACIÓN DE RECURSOS HUMANOS

Serán beneficiarios los estudiantes de carreras como la Licenciatura en Ciencias de la Computación de la Universidad Nacional de Cuyo, otras carreras de la Universidad Nacional de Cuyo en las cuales los integrantes del proyecto se desempeñan como profesores, como carreras dictadas en el Instituto Tecnológico Universitario (ITU) o la Facultad de Ciencias Exactas y Naturales, como carreras de otras unidades académicas.

Por otro lado, se espera contar con la participación de estudiantes de las asignaturas mencionadas durante la realización del proyecto. Estos estudiantes serán beneficiarios por dos motivos, recibir formación en cuanto a su iniciación como investigadores científicos, como también por la realización de prácticas profesionales, ya que los temas a investigar en el proyecto forman parte del programa de estudio de la carrera Licenciatura en Ciencias de la Computación de la Universidad Nacional de Cuyo.

Por lo mencionado en los párrafos anteriores, también serán beneficiarios los desarrolladores de aplicaciones y usuarios en general de plataformas de IoT, tanto los integrantes de la comunidad científica, los profesionales y los recursos humanos del medio local.

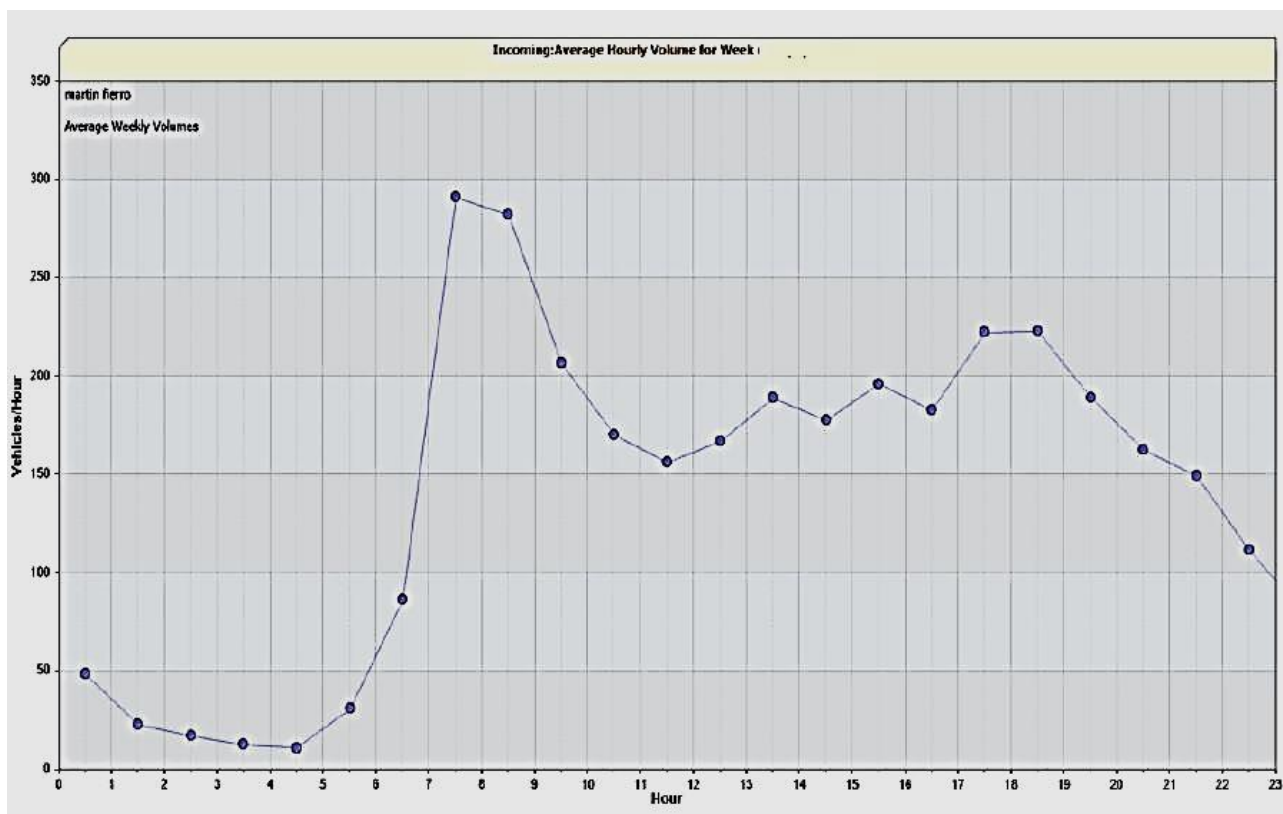


Figura 2. Número de vehículos promedio por hora y por semana

AGRADECIMIENTOS

Los autores agradecen el financiamiento recibido de la Universidad Nacional de Cuyo a través del proyecto B076 “Dispositivos de lógica programables como alternativa de solución en problemas de seguridad en Internet de las Cosas”.

[6] <https://harborresearch.com/more-than-a-meme-nfts-and-the-iot>.

[7] <https://openvpn.net/for/iot-secure-network>

5. BIBLIOGRAFIA

[1] Seok-Won Lee, Irish Singh, Masoud Mohammadian. “Blockchain Technology for IoT Applications”. ISBN: 978-981-33-4122-7. Springer. 2021.

[2] <https://www.ibm.com/topics/blockchain-iot>.

[3] <https://medium.com/@rajnandinibhowmik/benefits-of-using-blockchain-technology-for-iot-applications-and-products-a328199a8f78>.

[4] <https://www.sciencedirect.com/science/article/pii/S187705091830872X>.

[5] <https://www.leewayhertz.com/blockchain-iot-use-cases-real-world-products>.

IOT APLICADO A LA GANADERIA DE PRECISION

Paola I. Beltramini¹, Ivana Lazarte², Marcos D. Aranda^{1,5}, J. Eduardo Cano¹, L. Daniel Villagran¹, Oscar A. Herrera Conegliano^{3,4}, Sergio H. Gallina¹

pbeltramini@tecno.unca.edu.ar, sgallina@tecno.unca.edu.ar

(1) Grupo de Investigación en Internet de las Cosas (GIIoT), Facultad de Tecnología y Ciencias Aplicadas, Universidad Nacional de Catamarca (UNCA).

(2) Laboratorio de Tecnologías de Información y las Comunicaciones, Facultad de Tecnología y Ciencias Aplicadas, Universidad Nacional de Catamarca (UNCA).

(3) Facultad de Ciencias Agrarias, Universidad Nacional de Catamarca (UNCA).

(4) Instituto Nacional de Tecnología Agropecuaria (INTA).

(5) Universidad Siglo 21.

AREA TEMATICA: Arquitectura, Redes y Sistemas Operativos

RESUMEN

La conjugación y aplicación de las TIC, el Internet de las Cosas y las comunicaciones, en el sector agropecuario, posibilitan mejoras considerables de los procesos productivos, contribuyendo al desarrollo, al manejo y a la producción rentable y sostenible de la región. Bajo esta premisa, docentes investigadores de la UNCA estamos trabajando en el diseño e implementación de un nodo IoT de propósitos generales, con tecnología de comunicación LPWAN, que permitirá generar herramientas de base tecnológica para obtener información en tiempo real, con el fin de gestionar recursos de manera eficiente para la toma de decisiones, contribuyendo al desarrollo agroindustrial extensivo en regiones aisladas del noroeste de la República Argentina la finalización del proyecto permitirá dar respuesta a problemáticas del sector público y privado como también al sector científico técnico proponiendo un desarrollo en base a tecnologías altamente integradas y sistemas de comunicación de última generación.

Palabras Claves: IoT, Ganadería, LPWAN

CONTEXTO

El proyecto “IoT aplicado a la ganadería de precisión”, fue presentado en la convocatoria de Proyectos de Investigación y Desarrollo Interdisciplinarios (PIDI) 2022 de la Universidad Nacional de Catamarca (UNCA),

siendo aprobado y financiado por dicha Institución.

El proyecto se ejecuta en la Facultad de Tecnología y Ciencias Aplicadas (FTyCA) de la Universidad Nacional de Catamarca, principalmente en el Laboratorio que posee el grupo de investigación en Internet de las Cosas (GIIoT) y el Laboratorio de TIC's, los cuales aportan los materiales, equipamiento e instrumental necesarios.

Participan además investigadores de la Facultad de Ciencias Agrarias (FCA) de la UNCA y del Instituto Nacional de Tecnología Agropecuaria (INTA-EEA Catamarca), principalmente del grupo de Ganadería de Precisión.

La participación del INTA resulta fundamental ya que permitirá adaptar, evaluar y validar a campo el impacto de la incorporación del sistema desarrollado en el logro de la eficiencia técnico-económica, ambiental y social de los sistemas productivos a diferentes escalas.

Durante el progreso de investigación propuesta se buscará promover y gestionar vínculos y alianzas estratégicas con el INTA, en su función de Organismo vinculado al desarrollo agrotecnológico, y otros Organismos públicos y Asociaciones de productores, para el fortalecimiento de la I+D y la transferencia efectiva de tecnologías surgidas del proyecto.

1. INTRODUCCIÓN

El sector agropecuario tradicionalmente se consideraba lejos de los avances tecnológicos, sin embargo, encontró en las tecnologías blandas la manera de ser más eficiente, rentable,

sostenible y amigable con el medioambiente. La conjugación de la industria del software, el internet de las cosas (IoT) y la ciencia de datos, está brindando a nivel mundial una ventana de oportunidad para toda la gestión agro-ganadera, representando la revolución tecnológica del sector agropecuario, como lo fue el tractor a principios del siglo XIX.

En zonas áridas del Noroeste Argentino (NOA), la ganadería se realiza de manera extensiva, es decir, se emplean amplias superficies donde los animales recorren grandes distancias en busca de alimento y agua, por lo que es una demanda regional actual la necesidad de contar con información de la localización animal, sus patrones de desplazamiento, parámetros biológicos del ganado, entre otros, y herramientas de procesamiento y análisis de estos datos tanto con fines de investigación como con fines productivos.

Para conseguir esta información se utilizan “Nodos”, que tienen como función primaria la recolección y transmisión de datos provenientes de diferentes sensores, a un control centralizado. Los datos recolectados por estos nodos, pueden almacenarse en memorias o enviarse a un servidor de red a través de un Gateway, para su posterior análisis y procesamiento, y, mediante la aplicación de la Ciencia de Datos, generar información de gran valor para la toma de decisiones. Dada las características geográficas de la región, resultan de gran la importancia para la transmisión de la información las redes LPWAN, que son redes inalámbricas de baja potencia y reducido consumo energético, siendo la mejor alternativa ante las actuales tecnologías móviles de comunicación, en aquellas regiones donde 3G / 4G / WiFi, no tienen cobertura.

La incorporación de estas nuevas tecnologías al sector agropecuario ya es una realidad, principalmente aplicaciones de agricultura de precisión, pero no es de uso intensivo en las empresas agropecuarias del noroeste argentino. Existen esfuerzos de diversos actores del sector de ciencia y tecnología y del sector público y privado, que propenden a generar desarrollos de hardware y software para aportar soluciones al sector pecuario. Tal es el caso del Instituto Nacional de Tecnología Agropecuaria (INTA), que está desarrollando herramientas destinadas

a dar respuestas de precisión en diferentes áreas, como, por ejemplo, alimentación, lechería, comportamiento [3], manejo [5] [4], nutrición, consumo residual [1], etc.

En la provincia de Catamarca, estas tecnologías no se han difundido o aplicado en forma masiva en los sectores productivos primarios como son la agricultura y la ganadería debido, en parte, a la falta de dispositivos y componentes a un costo accesible, la falta de operadores y personal de mantenimiento. Además, en esta provincia los sistemas productivos ganaderos tienen en general niveles de productividad bajos y con grandes fluctuaciones entre años, principalmente por sobrepastoreo, avance de frontera agropecuaria, poca tecnificación y mal manejo de suelos.

La implementación de estas tecnologías, además, fomenta un cambio de paradigma para que el trabajo del campo sea más atractivo para las nuevas generaciones, lo cual impacta no solo la producción del campo, sino a los jóvenes que por distintos motivos ya no se encuentran en el sector rural [7].

El fomento de una industria de TIC aplicadas a un sector esencial en Argentina como lo es el conjunto de cadenas productivas agroindustriales, puede mejorar la posición competitiva del país en este segmento a nivel de costos, ganancias de eficiencia y de productividad en el mercado internacional, así como captar rentas tecnológicas adicionales respecto de los servicios que actualmente se proveen al exterior, pudiendo generar una marca país que se potencie a través de las sinergias entre un sector y otro. Por otra parte, la inversión a nivel de infraestructuras, formación de recursos humanos, centros de investigación y los modelos de gestión de la información agropecuaria pueden contribuir como base para desarrollar otros segmentos verticales [2].

Por ello, el impacto a nivel de cadena productiva de la incorporación del desarrollo propuesto, está dado porque se dispondría de hardware y software nacional, ajustado a nuestra realidad y con posibilidades de escalar a otros mercados, mejorando la productividad. Debido a que el factor económico representa la mayor restricción para el desarrollo de estas tecnologías, es imperativo lograr dispositivos de

bajo costo y bajo consumo de energía, como el que se propone en este proyecto.

El desarrollo de estas innovaciones implica además la generación de redes y la articulación público-privada de la región, desde el reconocimiento de demandas, a su validación e implementación, generando recursos humanos altamente capacitados en el uso y aplicación de estas tecnologías.

2. LINEAS DE INVESTIGACION Y DESARROLLO

De las áreas y líneas temática de la convocatoria PIDI 2022, este proyecto se incluye en las temáticas “Energía, materiales y tecnologías” - “Plataformas Tecnológicas”, dentro de los Objetivos de Desarrollo Sostenible el de “Construir infraestructuras resilientes, promover la industrialización inclusiva y sostenible, y fomentar la innovación”.

Se pretende desarrollar el Know-how para el desarrollo y la aplicación de nuevas tecnologías electrónicas, informáticas y de comunicaciones, en la región NOA, que redunden en una mejora de la actividad ganadera, incrementando la producción y rentabilidad, con sostenibilidad del territorio.

Para el desarrollo se deberá lograr el dominio de:

- *Tecnología de sensores de bajo costo y potencia.* Los sensores asequibles y fiables están haciendo posible la tecnología IoT .
- *Conectividad.* Una gran cantidad de protocolos de red para Internet permiten la conexión de sensores a la nube y a otras cosas para lograr una transferencia de datos eficiente
- *Plataformas de informática en la nube.* El aumento en la disponibilidad de plataformas en la nube permite a las empresas y a los consumidores acceder a la infraestructura que necesitan para escalar sin tener que administrarlo todo.
- *Aprendizaje automático y analítica.* Con los avances en aprendizaje automático y analítica, junto con el acceso a cantidades grandes y variadas de datos almacenados en la nube, se puede recopilar información de forma más rápida y fácil.

Asegurar el éxito del desarrollo del mismo requiere la interrelación de al menos tres ingenierías muy relacionadas, la electrónica, la informática y la agronomía.

3. RESULTADOS OBTENIDOS / ESPERADOS

Desde el año 2019, los integrantes del presente proyecto trabajan de manera ininterrumpida en el desarrollo de un nodo prototipo para obtener información acerca de la posición y estado biológico del ganado en áreas extensivas, para atender a una demanda del INTA Catamarca [6]. Derivados de ese trabajo se comenzó el estudio de la transmisión de los datos recolectados por el nodo en redes LPWAN, los protocolos de comunicación y en particular LoRa.

Podemos mencionar algunas de las publicaciones realizadas por el grupo en los últimos años en temáticas afines a la del proyecto: “IoT aplicado a la ganadería de precisión”, 2021; “Tecnología inteligente para la industria ganadera local”, 2021; “Sistema de control de rodeos con nodos inteligentes”, 2016; “Diseño de un nodo inteligente para instalaciones en vivienda”, 2016; “Protocolo de comunicación entre nodos inteligentes de una red domótica”, 2015, entre otras.

En esta misma línea de investigación, el presente proyecto de investigación tiene como finalidad diseñar, modelar, implementar y transferir, un sistema embebido aplicable a la ganadería extensiva de precisión, el cual permita la captura de parámetros ambientales y biométricos, la conexión a una red de datos de baja potencia (LPWAN), la transmisión de la información al usuario final y la posibilidad de manejo de la información recibida. Enmarcado en las nuevas tecnologías IoT.

Se espera, además:

- Desarrollar actividades y productos de I+D para el sector agropecuario, que generen herramientas de base tecnológica para obtener información en tiempo real, tendiendo a la unidad más pequeña de análisis (animal y ambiente) con el fin de gestionar recursos de manera eficiente para la toma de decisiones.
- Estudiar y aplicar las tecnologías de redes inalámbricas de bajo consumo LPWAN,

verificando los alcances reales en terrenos con dificultosa accesibilidad.

- Analizar, seleccionar y aplicar algoritmos basados en aprendizaje automático que permitan mejorar la calidad del desarrollo ganadero de precisión, a través del monitoreo.
- Adaptar, evaluar y validar a campo el impacto de la incorporación del sistema desarrollado en el logro de la eficiencia técnico-económica, ambiental y social de los sistemas productivos a diferentes escalas.
- Promover y gestionar vínculos y alianzas estratégicas con el INTA, en su función de Organismo vinculado al desarrollo agro-tecnológico, y otros Organismos públicos y Asociaciones de productores, para el fortalecimiento de la I+D y la transferencia efectiva de tecnologías surgidas del proyecto.
- Promover espacios de capacitación en el uso de IoT y la Ciencia de Datos con contenidos diferenciados acorde a los grupos objetivo.
- Difundir los resultados obtenidos mediante publicaciones y participaciones en reuniones científicas y académicas, permitiendo complementar la enseñanza en las carreras de Ingeniería Electrónica, Ing. en Informática e Ing. Agronómica de la Universidad Nacional de Catamarca (UNCA).

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está conformado por docentes investigadores de las carreras de Ingeniería Electrónica, Ingeniería Informática e Ingeniería en Agronomía de la UNCA, que vienen trabajando y desarrollando actividades de docencia, investigación y vinculación tecnológica, en temáticas afines al proyecto.

Integrantes del Grupo de Investigación en Internet de las Cosas (GIIoT) de la FTyCA, con experiencia en proyectos de investigación relacionados a los sistemas embebidos, la Domótica y al desarrollo e implementación de nodos y redes inteligentes para propósitos generales y aplicaciones ganaderas. Desde el año 2019, se trabaja de manera conjunta con el INTA, a partir de un Convenio de Vinculación Tecnológica, con el objetivo de promover el desarrollo agropecuario y ganadero, atendiendo de esta manera a los crecientes requerimientos

del sector, principalmente en zonas áridas donde el animal debe recorrer largas distancias diarias para buscar alimento.

Respecto a los integrantes del equipo, continuamente se capacitan en temas afines al proyecto. Uno de los integrantes está realizando el Doctorado en Ingeniería, mención Electrónica, en la Universidad Tecnológica Nacional (UTN) Regional Córdoba, cuyo tema de trabajo es “Aprendizaje Automático aplicado a la calidad del desarrollo en la Ganadería de Precisión”. Por otro lado, poseen formación en docencia universitaria y en enseñanza por competencias, por lo que el aporte desde el punto de vista pedagógico también está previsto ser abordado con ese importante aporte.

Integra el equipo de investigadores un Ing. Agrónomo, investigador del Grupo de Producciones Pecuarias de INTA EEA Catamarca y docente de la Facultad de Ciencias Agrarias de la UNCA. El área laboral de interés y el ámbito de sus publicaciones abarcan la cría de ganado en zonas áridas.

Como actividades complementarias los miembros del equipo participan en actividades de extensión universitaria desde el año 2016, tanto en funciones de dirección e integrantes de proyectos de voluntariado y extensión, como en la participación en ferias y actividades difusión de Ciencia y la Tecnología.

Anualmente se incorporan al equipo alumnos becario en investigación, en esta oportunidad y desde el año 2022 participa una alumna avanzada de la carrera de Ing. Electrónica.

En virtud de las capacidades de los integrantes del grupo, fruto de las actividades de capacitación, investigación y divulgación llevadas a cabo durante sus años de ejercicio profesional, se puede garantizar la factibilidad del proyecto presentado.

5. BIBLIOGRAFIA

- [1] Alende, M.; Garro, R.; Camiletti, M.; Pordomingo, A.J. (2017). El consumo residual en bovinos: base fisiológica. En: Producción bovinos para carne (2013-2017) Programa Nacional de Producción Animal. Alimentación de bovinos para carne. Publicación Técnica No 209, página 291-297.
- [2] Baum, G., Artopoulos, A., Aguerre, C., Albornoz, I., & Robert, V. (2009). Libro blanco

- de la prospectiva tic: proyecto 2020. - 1a ed. - Buenos Aires: Ministerio de Ciencia, Tecnología e Innovación Productiva, Argentina. ISBN 978-987-1632-00-8.
- [3] Bailey, D. W., Gross, J. E., Laca, E. A., Rittenhouse, L. R., Coughenour, M. B., Swift, D. M., & Sims, P. L. (1996). Mechanisms that result in large herbivore grazing distribution patterns.
- [4] Fuhlendorf, S. D., Engle, D. M., Elmore, R. D., Limb, R. F., & Bidwell, T. G. (2012). Conservation of pattern and process: developing an alternative paradigm of rangeland management. *Rangeland Ecology & Management*, 65(6), 579-589.
- [5] Scarnecchia, D. L., & Kothmann, M. M. (1982). A dynamic approach to grazing management terminology. *Rangeland Ecology & Management/Journal of Range Management Archives*, 35(2), 262-264.
- [6] Aranda, M., Beltramini, P., Cano, J., Virragrán, L., Moreno, J., Gallina, S., & Herrera Conegliano, O. A. (2021). IoT aplicado a la ganadería extensiva. *Revista Argentina de Ingeniería*. Año 9, vol. 17. ISSN 2314-0925.
- [7] Soto, J. P. T., Suárez, J. D. L. S. S., Rodríguez, A. B., & Cainaba, G. O. R. (2019). Internet de las cosas aplicado a la agricultura: estado actual. *Lámpsakos*, (22), 86-105.
- Zabaleta Miren "Nodos IoT: qué son y por qué están tan presentes en el Internet de las Cosas". España. 24 de mayo de 2021.
- Vite, H., Vargas, O., Vargas, L. y Vargas, J., (2018) Internet de las cosas aplicado a la producción agropecuaria, Editorial Grupo Compás, Guayaquil, Ecuador.
- Guerrero-Casado, J. (2017). Producción científica latinoamericana indexada en Scopus en el área de las ciencias agropecuarias: análisis del período 1996-2016. *Idesia (Arica)*, 35(4), 27-33.
- Pérez, R., Navajas, S., & Terry, E. (2019). IoT en ALC 2019: tomando el pulso al Internet de las Cosas en América Latina y el Caribe. Banco Interamericano de Desarrollo. Recuperado de: <http://ftp.isdi.co.cu/Biblioteca/BIBLIOTECA%20UNIVERSITARIA%20DEL%20ISDI/COLECCION%20DE%20LIBROS%20ELECTRONICO/S/LE-1966/LE-1966.pdf>
- Quintero Albornoz, J. S. (2020). Dispositivo IoT para el control de identificación y movilidad pecuaria. (Tesis de Maestría). Universidad Nacional de Colombia. Colombia.
- Loja Aguilar, A. F., & Naula Cedacero, E. M. (2022). Diseño de un sistema de monitoreo de posicionamiento y de la temperatura del entorno para ganado bovino utilizando una red de área local (Tesis de Grado). Universidad Politécnica Salesiana. Ecuador.
- Cangrejo Aguirre, D. C., & Hernández López, D. M. (2019). Sistema de apoyo a la prevención del abigeato de ganado bovino utilizando tecnologías IoT y Cloud. (Tesis de Grado). Universidad Santo Tomás, Bogotá. Colombia.
- Gorandi, E., Clemares, N., & Moltoni, A. (2015). Collar con tecnología GPS para monitoreo animal. *Investigación y Desarrollo en Electrónica*. Año 1, Num. 2. ISSN: 2468-9696
- Hermans, J. & Castiaux, A. (2007). Knowledge Creation through University-Industry Collaborative Research Projects. *The Electronic Journal of Knowledge Management*, 5(1), 43-54. <https://researchportal.unamur.be/en/publications/knowledge-creation-through-university-industry-collaborative-rese-2>

IoT for Vineyard Posts

Agustin Giardinelli^{1,2}, Gustavo Mercado¹, Carlos Taffernaberry¹, Ariel Verdejo¹,
Ana Diedrichs¹, Cristian Bernoco¹, Ana Lattuca²

¹gridTICs – Grupo en Tecnologías de la Información y las Comunicaciones
Departamento de Electrónica / Facultad Regional Mendoza / UTN
gmercado@frm.utn.edu.ar

² Cátedra Proyecto Final
Departamento de Electrónica / Facultad Regional Mendoza / UTN
agustingiardinelli@gmail.com

Resumen

El avance de la mecanización evidenció las deficiencias estructurales de los postes de madera impregnados, utilizados en viñedos, los cuales se quiebran y deben ser reemplazados. Se suma a esto la dificultad de su disposición final por su impregnación con CCA (cromo, cobre y arsénico) que es un riesgo ambiental. Una innovación para evitar estos problemas es la utilización de postes metálicos: presentan mayor resistencia mecánica, fabricación estandarizada, ausencia de compuestos tóxicos, son durables y reciclables. La mayoría son importados y si bien existen normativas y reglamentos nacionales para este tipo de material, no existe experiencia en el país ni recomendaciones para su correcto uso. Para desarrollar un manual de buenas prácticas de uso, es necesario generar información de su comportamiento estructural y durabilidad, tanto frente a solicitaciones estáticas y dinámicas, como a los efectos de corrosión que se presentan en condiciones de producción. Con este objetivo, se evaluarán postes metálicos ya instalados en viñedos y las condiciones ambientales y productivas a las que estuvieron sometidos. Para lo que se diseñará y construirá un poste instrumentado para medir, registrar, transmitir, almacenar y realizar análisis, por ciencia de datos, de las tensiones a las cuales está sometido y las variables ambientales que pueden afectar su durabilidad. Esta es la tarea descripta en el presente trabajo.

Palabras Claves: Postes metálicos, Galvanizado, Viticultura, Internet de las Cosas, Cloud Computing, Big Data, Analítica

Contexto

El presente trabajo está inserto en el proyecto de investigación acreditado por la Universidad Tecnológica Nacional código CCECME0009834 denominado “*Contribución a la competitividad y sostenibilidad de la vitivinicultura mediante la innovación en postes metálicos conectados por medio de la implementación de sistema de Internet de las Cosas*” y al Proyecto Final de Grado, denominado “*Medición y Transmisión de parámetros que afectan la durabilidad de postes galvanizados para vitivinicultura (MTpADPGV)*”

El proyecto es llevado adelante por investigadores y alumnos becarios del grupo gridTICs y la Cátedra de Proyecto Final, ambos del Departamento de Electrónica, de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional.

1. INTRODUCCIÓN

La estructura de conducción de la vid es una de las mayores inversiones en el establecimiento de un viñedo y tienen influencia significativa en su productividad y rentabilidad [1]. Se denomina sistema de conducción al método para orientar el crecimiento de los sarmientos y ubicarlos convenientemente. Existe una amplia variedad de materiales que pueden ser usados para el sostén de la viña; la práctica común ha sido seleccionar el material por costo inicial sin considerar la durabilidad y su costo por año de servicio, siguiendo con la tradición de la madera. La tendencia actual de modernización de viñedos y mecanización de la cosecha, requiere de instalaciones que soporten esfuerzos mecánicos adicionales. Y es aquí donde los postes galvanizados son una

alternativa frente a la opción tradicional. Es un producto sencillo de fabricar, fácil de manipular e instalar, más liviano que los postes de madera y reciclable.

Existe una amplia oferta a nivel internacional, con diferentes diseños, calidades de recubrimiento de zinc y resistencia del acero base. Los postes galvanizados surgieron como consecuencia de la preocupación que genera la rápida corrosión del acero en el tiempo. La durabilidad del sistema poste/suelo dependerá del espesor de esa capa y de la composición fisicoquímica de los suelos. Con el recubrimiento de zinc se prolonga la vida útil, a la vez que se preserva la integridad del acero y su resistencia mecánica.

La investigación del comportamiento del zinc en suelos ha recibido poca atención si se compara con los estudios de la corrosión atmosférica del zinc. Existen numerosos antecedentes de estudios electroquímicos realizados en laboratorio con soluciones simuladas de suelos para evaluar el comportamiento de aceros galvanizados [2] y modelar el mecanismo de corrosión [3], que constituyen una referencia para este proyecto. Varios estudios publicados muestran los compuestos químicos que se forman sobre la superficie de las diferentes fases del recubrimiento de zinc cuando se exponen al ambiente, en particular, los que se realizan con microscopio electroquímico para identificar qué ocurre en cada fase [4]. Así, puede conocerse cómo se corroe el zinc y la velocidad de disolución cuando se daña la cobertura por causas mecánicas durante el hincado de un poste.

Actualmente, existe capacidad tecnológica y equipamiento específico, en el país, que permitirá validar los resultados obtenidos.

Hay antecedentes de instalación de sistemas de monitoreo de las tensiones a las que está sometido un sistema de conducción de vid y de frutales, aunque con objetivos diferentes asociados con la evolución del cultivo [5] y de la cosecha [6]. Han demostrado limitaciones cuando se hacen mediciones puntuales en el tiempo, en particular por falta de registro de las cargas dinámicas que afectan a la instalación, como los efectos meteorológicos, por lo que se

han instalado sistemas de adquisición continua de datos. La metodología de trabajo es replicable para evaluar las cargas de tracción que afectan los postes. Por otro lado, el grupo de investigación grid TICs tiene amplia experiencia en I&D en agricultura de precisión con tecnologías IoT, como se indica en [7] y [8].

3. OBJETIVOS Y AVANCES DEL PROYECTO

Objetivo Principal

Para el cumplimiento del objetivo general del proyecto marco PITES, el área de Sensores y Comunicaciones, tiene el objetivo de diseñar, implementar, instalar, validar y operar un sistema de postes de viñedos, en el terreno, instrumentados y comunicados, denominado "IoT for vineyard posts.

Objetivos específicos

- 1- Analizar y desarrollar Sensoramiento, Digitalización y Adquisición de Datos (Sensors and Digital Data Acquisition)
- 2- Diseñar sistemas de pequeños dispositivos de cómputo/comunicación (Motes).
- 3- Diseñar una red Local que comunique al coordinador, con los motes y la nube. (Ad Hoc Net).
- 4- Implementar acciones en Computación en la nube (Cloud Computing).
- 5- Implementar metodologías de análisis de datos (Analytics).

METODOLOGÍA

PROPUESTA

Desarrollar e implementar un sistema, basado en tecnología de IoT, de Sensoramiento, Digitalización, Adquisición, Almacenamiento y Análisis de Datos para aplicar al estudio de las estructuras metálicas de conducción de viñedos, denominado "Poste Instrumentado". En la Figura 1 se muestra un esquema del sistema "IoT for Vineyard Posts".

DESARROLLO

Sensors and Digital Data Acquisition

Los sensores serán el punto de ingreso de los datos para el análisis de comportamiento del sistema en situación real y en tiempo real. Las acciones, se refieren a la selección,

adquisición, adecuación, implementación, instalación y prueba de sensores de toma de información del tipo ambiental, estructural y de corrosión.

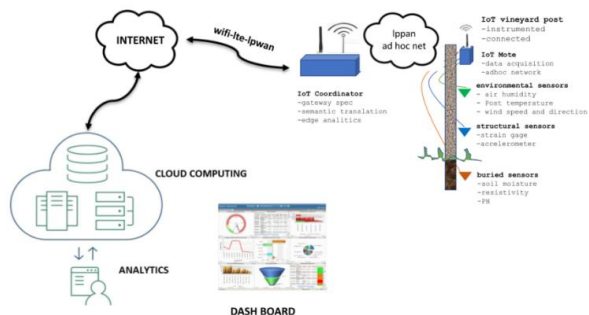


Figura 1: Esquema del sistema "IoT for Vineyard Posts"

Los tipos de sensores a estudiar son:

- Ambientales: Humedad ambiente, velocidad y dirección de viento, temperatura ambiente y temperatura del poste.
- Estructurales: extensómetros (strain gauges), acelerómetro.
- Corrosión: estacas calibradas, humedad de terreno, pHmetro, conductivímetro.

Motes

Los Motes son microcontroladores que reciben la información de los sensores, a través de las interfaces apropiadas, y transmiten la información hacia un coordinador, a través de una red de comunicaciones inalámbrica. Para ello el mote cuenta con subsistema de Transmisión/Recepción que implementa los protocolos necesarios para hacer una "Red de Sensores Inalámbricos" - WSN.

Ad Hoc Net

En este proyecto se propone el uso de protocolos estandarizados y abiertos. Por lo que se usarán, los siguientes protocolos de IoT: IEEE 802.15.4, 6lowpan, rpl, Core/Coap, etc [9]. Los motes envían la información directa o indirectamente a un coordinador. El coordinador, es un microcontrolador cuya función es recibir la información de los motes, a través de una red de comunicaciones inalámbrica, procesar la información y enviarla a un lugar remoto con capacidad de almacenamiento y de procesamiento de la información.

Cloud Computing

La utilización de Cloud Computing permite el almacenamiento de la información y una segunda etapa de procesamiento, así como su monitorización y análisis. Adicionalmente se pueden destacar las siguientes ventajas: Reducción de costos, Pago por necesidad; Alta escalabilidad y disponibilidad y Capacidad para agregar grandes cantidades de datos.

Analytics

En este proyecto haremos uso de la ciencia de datos para analizar los comportamientos de los postes metálicos de viña ante las fenómenos estructurales, de corrosión de materiales y procesos funcionales y operativos.

Field implementation and operation

El sistema se instalará en tres parcelas vitivinícolas, previamente seleccionadas por sus condiciones agrícolas, ambientales y por sus características de suelo. En cada una de ellas se instalarán tres postes conectados de acuerdo a su ubicación en las hileras, eligiendo las siguientes ubicaciones: edge post, middle post y border post, tal como se muestra en la Figura 2. Todos los postes contarán con sus respectivos sistemas de sensores, motes y redes ad-hoc. Cada parcela contará con un coordinador y un sistema de acceso a Internet. El sistema de postes conectados, se prevé, esté operativo al menos dos ciclos agrícolas completos y, por lo tanto deberán poseer un mecanismo de mantenimiento apropiado.

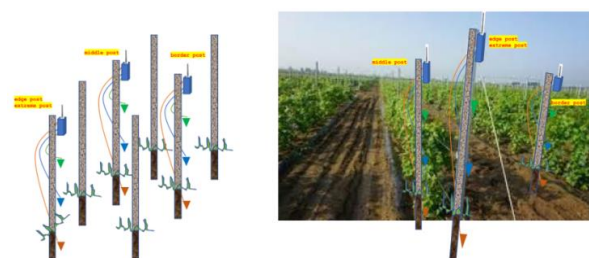


Figura 2: Esquema del montaje de los postes conectados en las parcelas agrícolas

AVANCES DEL PROYECTO

A la fecha se ha logrado diseñar e implementar, en laboratorio, un prototipo funcional "poste instrumentado", con operación "end to end": con se detalla en adelante. ver Figura 3.

Nodo embebido de adquisición de datos y transmisión de datos

Se decidió usar el microcontrolador Openmote CC2538 [10] y su adaptador OpenUSB. El OpenMote-CC2538 posee un núcleo CC2538, que es un SoC de Texas Instruments con un procesador Cortex-M3 de 32 bits y una radio que opera en la banda de 2,4 GHz y cumple con el estándar IEEE802.15.4 [11]. En los nodos Open Mote se instala el Sistema Operativo Contiki [12], el cual tiene funciones que soportan protocolos estándar y de comunicación de bajo consumo, como IPv6/6LoWPAN, RPL y CoAP. Con la capacidad, de los nodos, de comunicación entre sí, se conforma lo denominada “Red de sensores inalámbricos - WSN”

Coordinador

El coordinador actúa como una interfaz entre la red de nodos y la red de internet, transmitiendo la información de los sensores a la nube (cloud). Esto significa, que una parte del coordinador se conforma como un router de borde y la otra como nodo de la red Ad-Hoc, cumpliendo funciones de gatekeeper. El router se implementa con una placa Raspberry Pi Model 3 B [13], que nos permite trabajar en un entorno de Linux. En el nodo (mote) se trabaja con el SO Contiki.

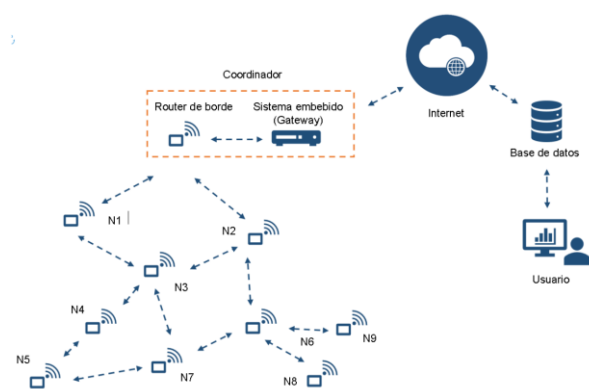


Figura 3: Esquema de implementación end to end

La vinculación física entre el nodo de la red y la placa Raspberry Pi se realiza a través de conexión USB.

Para la comunicación lógica (enlace), Contiki utiliza Tunslip, que es un mecanismo usado para unir el tráfico IP entre un host y otro

elemento de red, generalmente un enrutador de borde, a través de una línea serial. Tunslip crea una interfaz de red virtual (tun) en el lado del npdp y utiliza SLIP (protocolo de Internet de línea serie) para encapsular y pasar tráfico IP hacia y desde el otro lado de gatekeeper.

El siguiente paso es realizar consultas mediante el protocolo CoAP [14], al nodo que actúa como servidor. Para ello, hacemos uso del lenguaje de programación Python, que cuenta con una librería específica llamada “aicoap”.

Se utiliza la herramienta aiocoap-client para enviar una solicitud GET al nodo servidor y acceder a sus recursos.

Base de Datos

En esta etapa se implementó la base de datos, almacenando los datos recolectados de los nodos.

El procedimiento consiste en realizar consultas a los nodos de manera automática. Cuando el nodo envía una respuesta, se almacena en una tabla de la base de datos. Para llevar a cabo este proceso, se utiliza también lenguaje Python, a través de CoAP y se almacenan en una base de datos MySQL [15], utilizando también una librería MySQL.

Se utilizó la herramienta de trabajo MySQL Workbench para crear la base de datos, que consta de una tabla con cinco campos, a saber: identificador de sensor (idsensor), número de dato (numDato), fecha, temperatura y humedad.

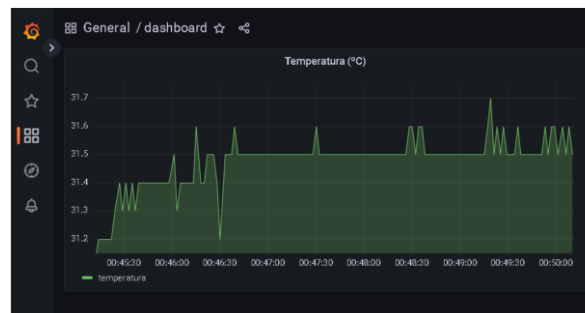


Figura 4: Visualización de datos

Visualización de datos

En esta etapa se implementa la aplicación Grafana [16], que permite crear un panel de visualización de los datos almacenados.

Grafana posee los complementos necesarios para poder vincularse con una base de datos MySQL.

En la Figura 4 se muestra la gráfica de datos:

4. FORMACIÓN DE RECURSOS HUMANOS

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos.

La meta es fortalecer la capacidad para realizar investigación científica, generar conocimientos y facilitar la transferencia de tecnología que permita el desarrollo humano. Para lograr estos objetivos se dispuso del siguiente personal:

Tres Investigadores formados
Un Investigador de apoyo
Dos Becarios alumnos (Beca Manuel Belgrano)
Un Tesista de carrera de grado

5. BIBLIOGRAFÍA

[1] Thomas J. Zabadal, Engineering a Modern Vineyard Trellis, College of Agriculture & Natural Resources, Michigan State University (1997).

[2] C. Soriano, A. Alfantazi, Corrosion behavior of galvanized steel due to typical soil organics, Construction and Building Materials, 102 (2016).

[3] Padilla Pérez, Victor E, Behavior of Hot-Dipped Galvanized Steel in Infrastructure Applications, PhD Thesis, University of British Columbia, Canada (2014).

[4] Manhabosco Sara M. et al, Corrosion behaviour of galvanized steel studied by electrochemical microprobes applied on low-angle cross sections, Corrosion Science, Volume 140, August 2018

[5] Blom PE, Tarara JM, Trellis Tension Monitoring Improves Yield Estimation in Vineyards, Department of Agriculture, American Society of Horticultural Science, HortScience, Vol 44: Issue 3 (2009)

[6] Tarara J. M. et al, Estimation of grapevine crop mass and yield via automated measurements of trellis tension, American Society of Agricultural and Biological Engineers, Transactions of the ASAE, Vol. 47(2): 647-657 (2004)

[7] K. Brun-Laguna, A. L. Diedrichs, J. E. Chaar, D. Dujovne, J. Taffernaberry, G. Mercado, Thomas Watteyne. (2016) A Demo of the PEACH IoT-based Frost Event Prediction System for Precision Agriculture. IEEE International Conference on Sensing, Communication and Networking (SECON), Londres, Inglaterra, 27-30 Junio 2016.

[8] A. Diedrichs, M. Robles, D. Dujovne, F. Bromberg and G. Mercado "Characterization of LQI behavior in WSN for glacier area in Patagonia Argentina", Regular Papers, 2013 Fourth Argentine Symposium and Conference on Embedded Systems (SASE/CASE),

Printed Book IEEE CATALOG CFP1346V-PR, Buenos Aires, August 14th-16th, 2013

[9] MORABITO, Roberto; JIMÉNEZ, Jaime. IETF protocol suite for the Internet of Things: Overview and Recent Advancements. IEEE Communications Standards Magazine, 2020, vol. 4, no 2, p. 41-49.

[10] Vilajosana, X., Tuset, P., Watteyne, T., & Pister, K. (2015). OpenMote: Open-source prototyping platform for the industrial IoT. In Ad Hoc Networks: 7th International Conference, AdHocHets 2015, San Remo, Italy, September 1-2, 2015. Proceedings 7 (pp. 211-222). Springer International Publishing.

[11] Molisch, A. F., Balakrishnan, K., Chong, C. C., Emami, S., Fort, A., Karedal, J., ... & Siwiak, K. (2004). IEEE 802.15. 4a channel model-final report. IEEE P802, 15(04), 0662.

[12] Oikonomou, G., Duquennoy, S., Elsts, A., Eriksson, J., Tanaka, Y., & Tsiftes, N. (2022). The Contiki-NG open source operating system for next generation IoT devices. SoftwareX, 18, 101089.

[13] Raspberry Pi, www.raspberrypi.org/ sitio visitado 15 Feb 2023

[14] Joshi, Manveer, and Bikram Pal Kaur. "Coap protocol for constrained networks." International journal of wireless and microwave technologies 5.6 (2015): 1-10.

[15] Christudas, Binildas, and Binildas Christudas. MySQL. Apress, 2019.

[16] Grafana: The open observability platform | Grafana Labs, grafana.com/ sitio visitado 11 Ene 2023

IoT-based energy monitoring system, to deploy an AMI in a Micro Smart Grid environment.

Gonzalo Di Paola¹, Tomás Suarez¹ Gustavo Mercado², Ana Lattuca¹.

² Cátedra Proyecto Final

Departamento de Electrónica / Facultad Regional Mendoza / UTN
gonzaloopez2206@gmail.com

¹gridTICs – Grupo en Tecnologías de la Información y las Comunicaciones
Departamento de Electrónica / Facultad Regional Mendoza / UTN
gmercado@frm.utn.edu.ar

Resumen

Desde un contexto global, la red eléctrica inteligente (o REI; smart grid en inglés) se puede definir como la integración dinámica de los desarrollos en ingeniería eléctrica, almacenamiento energético y los avances de las tecnologías de la información y comunicación (o TIC), dentro del negocio de la energía eléctrica (generación, transmisión, distribución, almacenamiento y comercialización, incluyendo las energías alternativas)

El concepto de “micro smart grid” se diferencia de la “smart grid” en el concepto del alcance de la administración de la red, mientras una smart grid supone una sistema de distribución de energía con control administrativo público, una micro smart grid, es un sistema cerrado e interno a una entidad privada, que pretende administrar la energía de acuerdo a criterios y normas propias de la institución.

Por otro lado un AMI (Advanced Metering Infrastructure) es un sistema de comunicación bidireccional para recopilar información detallada de medición en todo el sector de servicios de una micro smart grid. La AMI suele estar automatizada y permite realizar consultas en tiempo real y a petición con los puntos finales de medición.

En este proyecto se pretende diseñar e implementar un sistema de registro de energía basado en IoT(Internet of Thing), con el que se pueda desplegar una AMI.

Palabras Claves: Smart Grid, AMI, Internet de las Cosas.

Contexto

El presente trabajo está inserto en el proyecto de investigación acreditado por la UTN código ENUTIME0005424TC “*Smart Micro Grid de Campus Universitario - Desarrollo, implementación y prueba.*” y al Proyecto Final de Grado, denominado “*Micro Smart Grid*”

El proyecto es llevado adelante por el grupo gridTICs y la Cátedra de Proyecto Final,. ambos del Departamento de Electrónica de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional.

1. INTRODUCCIÓN

Desde un contexto global, la red eléctrica inteligente (o REI; smart grid en inglés) [1] se puede definir como la integración dinámica de los desarrollos en ingeniería eléctrica, almacenamiento energético y los avances de las tecnologías de la información y comunicación (o TIC), dentro del negocio de la energía eléctrica (generación, transmisión, distribución, almacenamiento y comercialización, incluyendo las energías alternativas); permitiendo que las áreas de coordinación de protecciones, control, instrumentación, medida, calidad y administración de energía, etc., sean concatenadas en un solo sistema de gestión con el objetivo primordial de realizar un uso eficiente y racional de la energía. El término red inteligente se asocia a menudo con el concepto de medidores inteligentes capaces de ofrecer una facturación detallada por franjas horarias, lo que permitiría a los consumidores no solo el elegir las mejores tarifas de entre las diferentes empresas eléctricas, sino también discernir entre las horas de consumo, lo que a su vez permitiría un mejor uso de la red. Este sistema también permitiría mapear con más

precisión el consumo y anticipar mejor las necesidades futuras a nivel más local. La irrupción de las energías renovables en el panorama energético ha cambiado notablemente los flujos de energía en la red eléctrica: ahora los usuarios no sólo consumen, sino que también producen electricidad a través de la misma red. Por tanto, el flujo de energía es ahora bidireccional. Una red inteligente envía electricidad desde los proveedores a los consumidores y viceversa, usando una tecnología digital para controlar las necesidades del consumidor. El concepto de “micro smart grid” [2] se diferencia de la “smart grid” en el concepto del alcance de la administración de la red, mientras una smart grid supone un sistema de distribución de energía con control administrativo público, una micro smart grid, es un sistema cerrado e interno a una entidad privada, que pretende administrar la energía de acuerdo a criterios y normas propias de la institución.

Por otro lado un AMI (Advanced Metering Infrastructure) [3] es un sistema de comunicación bidireccional para recopilar información detallada de medición en todo el sector de servicios de una micro smart grid. La AMI suele estar automatizada y permite realizar consultas en tiempo real y a petición con los puntos finales de medición.

En este proyecto se pretende diseñar e implementar un sistema de registro de energía basado en IoT(Internet of Thing), con el que se pueda desplegar una AMI.

Antecedentes

Este proyecto es una continuación del proyecto FONARSEC 2013. Denominación: "Red inteligente ciudad Gral San Martín Mendoza", Proyecto FITS Energía 2013 - UREE. financiado por el Fondo sectorial de energía. Participantes: CAPP (EDEST, EMESA y UTN FRM), con vigencia desde el 1/1/2015 [4][5][6].

3. OBJETIVOS Y AVANCES DEL PROYECTO

Objetivo Principal

Se busca, a largo plazo proveer, una herramienta asequible que permita la

escalabilidad de las micro smart grid.

Objetivos específicos

Con este prototipo se busca validar los conceptos:

- La información entregada por el sistema tiene potencial de acoplarse con un sistema de toma de decisiones que permite ahorrar energía en el edificio.
- El servicio tiene como costo una fracción de la energía eléctrica ahorrada.
- Integrar la herramienta AMI a un sistema general de Micro Smart Grid

Avances del Proyecto

Alcance

Se obtendrán datos de dos circuitos eléctricos y se mostrará esa información en una plataforma web interactiva.

Requisitos de Alto Nivel

El siguiente listado presenta los requisitos que el resultado del proyecto (producto o servicio) debe cumplir como condición necesaria.

1. Otorgar datos en tiempo real de al menos un dispositivo o red eléctrica
2. Almacenar datos de consumo históricos de cada edificio.
3. Implementar una interfaz amigable para el usuario donde se visualicen los datos.

Desarrollo

Resumen

El fin del proyecto es mostrar de manera clara en un panel tipo dashboard un resumen de los datos obtenidos por un medidor de energía eléctrica en una fecha determinada.

En la Figura 1 se muestra un esquema con los módulos que conforman al proyecto

MÓDULO HARDWARE

Medidor de Energía Eléctrica

Se adquirió un medidor de la marca SolverBox [7]. Este medidor posee la capacidad de medir: Potencia activa, Corriente RMS, THD, Nivel de armónicos hasta el 7mo y energía acumulada. Estos datos fueron considerados como suficientes para una correcta medición de la **calidad de energía**.

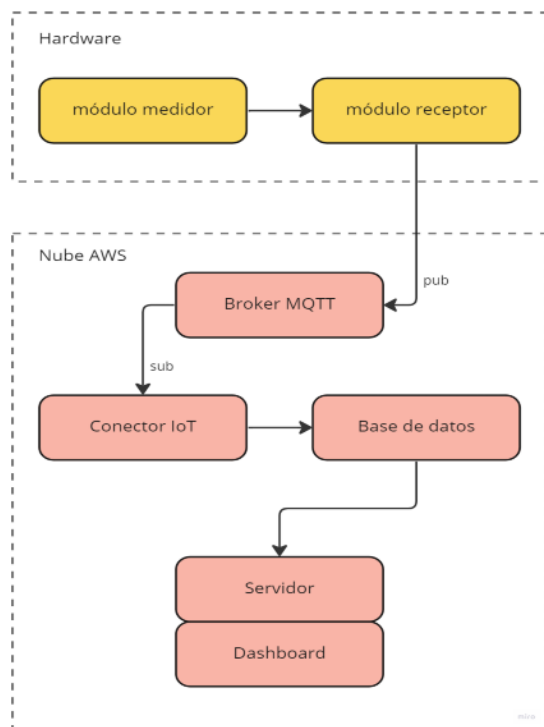


Figura 1: Esquema de AMI

Este dispositivo obtiene datos de corriente/tensión eléctrica a través de pinzas amperométricas no invasivas, que tienen una capacidad de medición de hasta 100 A

Las mediciones obtenidas son enviadas al módulo procesador mediante un puerto de comunicaciones serial (UART), con un formato de una trama tipo String.

Módulo de Adquisición y Transmisión de Datos: NODEMCU 1.0 - ESP8266

NodeMCU 1.0 [8] es una placa de desarrollo basada en el chip ESP8266, con un microcontrolador de bajo costo, bajo consumo de energía y con capacidades Wi-Fi.

NodeMCU 1.0 incluye el chip ESP8266 y una serie de componentes adicionales, memoria flash, EEPROM, puertos USB y GPIO.. Este módulo se programó en C++ utilizando el IDE PlatformIO [9], integrado en el editor de texto Visual Studio code.

La función del nodo es obtener la trama de datos, que es provista por el medidor de energía, procesarla y enviarla a través de la red Wi-Fi hacia un broker MQTT. El procesamiento consiste en identificar y organizar los datos de manera que sea apropiados para enviarlos al broker y adicionalmente agrega la fecha en que se

adquirieron los datos.

MÓDULO SOFTWARE

Criterios, limitaciones y desafíos:

A la hora de elegir un protocolo de comunicación, se contempló la posibilidad de usar CoAP (Constrained Application Protocol) [10]. Es un protocolo de aplicación diseñado específicamente para aplicaciones IoT y dispositivos restringidos. Es muy similar a HTTP y, por lo tanto, es fácil de implementar en dispositivos con limitaciones de recursos. CoAP es eficiente en términos de ancho de banda y es adecuado para aplicaciones en las que la frecuencia de actualización es alta.

Finalmente se eligió usar MQTT [11] debido a la escalabilidad y la facilidad de desarrollo, pues al ser muy utilizado, tiene mucha documentación en internet. MQTT es un protocolo altamente escalable, debido a su arquitectura basada en publicador-suscriptor y su capacidad para manejar una gran cantidad de dispositivos conectados de manera eficiente. El protocolo MQTT permite a los clientes publicar mensajes en un tópico y a los servidores responder a esos mensajes, lo que lo hace adecuado para aplicaciones con grandes cantidades de datos y dispositivos.

También se optó por construir un broker MQTT propio en vez de contratar un broker privado, ya que hacerlo de esta manera, nos brindó libertad y control sobre el broker.

Módulo Broker MQTT

En un sistema de comunicación basado en MQTT, el broker actúa como un intermediario confiable y eficiente entre los dispositivos publicadores y suscriptores.

Los publicadores envían mensajes a un tópico específico, mientras que los suscriptores se registran para recibir mensajes de ese tópico. Para este proyecto se optó por desarrollar un broker MQTT propio. El mismo fue programado en Express.js utilizando la librería “mosca”, que permite correr un broker en un puerto específico.

Express.js está escrito en JavaScript, y utiliza la programación asíncrona y “event-driven” de JavaScript para crear aplicaciones web rápidas y escalables.

Módulo Conector IoT: MQTT to MongoDB

Este módulo es responsable de recibir los mensajes que llegan a un tópico específico en el broker MQTT. El módulo toma los datos, los formatea según un formato especificado, y finalmente los guarda en la base de datos MongoDB.

Está programado en node.js utilizando la librería “mqtt”, la cual tiene amplia documentación y ejemplos. Corre sobre el mismo servidor que el broker MQTT en la nube de AWS.

Módulo Base de Datos

MongoDB [12] es una base de datos NoSQL (sin esquema) de tipo documental. Esto significa que en lugar de almacenar datos en tablas con filas y columnas como lo hacen las bases de datos relacionales, MongoDB almacena datos en documentos en un formato similar a JSON, donde cada documento es un registro independiente con su propio conjunto de campos y valores.

En el caso del proyecto se utilizó la siguiente estructura de datos para ser almacenada:

```
const measureSchema = mongoose.Schema({
  measure: String,
  date: Date
});
```

Esto permite guardar la trama recibida por el suscriptor y la fecha en la cual fue adquirida sin mayores complicaciones a la hora de tener un manejo de datos eficiente.

A continuación se muestra un ejemplo de los datos almacenados en la base de datos:

```
{
  "_id": ObjectId("63e40058fb8c410bfd8799a8"),
  "measure": "r,250.00,27.88,41.00,32.73,2.79,2.37,1.95,1.53,1.12,0.70,0.28,0.13,12...",
  "date": "2023-02-08T17:04:39.000+00:00",
  "__v": 0
}
```

Para el manejo de datos en mongodb se utilizó la librería mongoose.

Mongoose es una biblioteca de modelado de objetos de MongoDB para Node.js. Es un enrutador que actúa como una capa intermediaria entre la aplicación y la base de datos MongoDB, permitiendo interactuar con la base de datos a través de objetos en lugar de

tener que escribir manualmente código de consulta a la base de datos.

Panel DashBoard

En el contexto de sistemas de monitorización de consumo eléctrico, el dashboard es una herramienta que permite visualizar los datos adquiridos por los medidores en tiempo real y de forma organizada. El objetivo principal es proporcionar una representación clara y concisa del consumo eléctrico, con el fin de permitir a los usuarios expertos tomar decisiones informadas basadas en la información proporcionada.

El dashboard recopila, procesa y muestra los datos adquiridos, permitiendo a los usuarios expertos analizar y comparar los patrones de consumo eléctrico a lo largo del tiempo, identificar tendencias y patrones anómalos, y ajustar la gestión de la energía en consecuencia.

BackEnd

El Backend es la parte de una aplicación web que se encarga de las tareas y operaciones que ocurren en el servidor. Este es el encargado de procesar y almacenar datos, y de ofrecer servicios a la parte del cliente (frontend) de la aplicación.

Express [13] es una popular librería de Node.js que se utiliza para crear aplicaciones web y servicios API en el lado del servidor.

Además es común que en el backend se cree una API para comunicar el frontend con la base de datos.

La API desarrollada permite comunicar el Frontend con la base de datos. El Frontend realiza una petición GET el cuál envía a una ruta específica 4 parámetros importantes: Las fechas de inicio y fin, el mes y el piso del cual se quieren obtener los datos.

Con estos datos podemos realizar la consulta a la base de datos para obtener los datos relacionados con las fechas relacionadas. El formato de la fecha recibido es convertido al formato 2023-02-08T17:00:00.000+00:00 según la norma ISO 8601 [14].

Todos estos datos son agregados a un objeto JSON y es enviado al Frontend.

FrontEnd

Del lado del Frontend tenemos la página web desarrollada la cual muestra datos procesados por el servidor. Estos son mostrados en distintas gráficas de forma tal que pueda sacar una rápida conclusión sobre lo que está ocurriendo en el edificio bajo medición.

Los datos presentados son extraídos según las fechas que se indiquen en la página web. Estos incluyen la energía acumulada, el factor de distorsión armónica total (THD), la potencia reactiva, el coseno phi, los armónicos del 1 al 7 con sus respectivos valores en un gráfico de barras.

La página también presenta información sobre la potencia activa, la corriente rms y la corriente pico en tiempo real.

Además de presentar información visual sobre la energía eléctrica, la página web desarrollada en React, también se comunica con un backend desarrollado en Node.js. La comunicación entre la página web y el backend se realiza mediante la utilización de la biblioteca Axios.

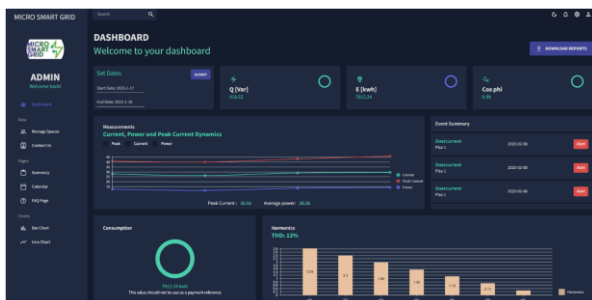


Figura 2: Front End

Axios es un cliente HTTP que permite hacer solicitudes desde el lado del cliente. Esto permite que la página web obtenga los datos necesarios del backend de manera dinámica, actualizando en tiempo real la información presentada en la página. La combinación de React para el desarrollo de la interfaz y Node.js para el backend, junto con Axios para la comunicación, brinda una solución completa y eficiente para la visualización y monitoreo de datos.

4. FORMACIÓN DE RECURSOS HUMANOS

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos.

La meta como investigadores es fortalecer la capacidad para realizar investigación

científica, generar conocimientos y facilitar la transferencia de tecnología que permita el desarrollo humano.

Para lograr estos objetivos se dispuso del siguiente personal:

Un Investigadores formado
Un Investigador de apoyo
Dos alumnos en Proyecto Final de Carrera

5. BIBLIOGRAFÍA

- [1] Shengwei Mei, et al, "Paving the Way to Smart Micro Energy Grid: Concepts, Design Principles, and Engineering Practices" IEEE © 2017 CSEE
- [2] Padmanaban, Sanjeevikumar, et al. Microgrids. CRC Press, 2020.
- [3] A. K. Chakraborty and N. Sharma, "Advanced metering infrastructure: Technology and challenges," 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), Dallas, TX, 2016.
- [4] G. Mercado et al "SG-SM - Smart Grid San Martín. Red de Distribución y Generación de Energía Inteligente en Ciudad Gral San Martín Mendoza" WICC 2015, Salta 16-17, Abril 2015
- [5] G. Mercado, J. Da Peña, G. Vivone, M. Ledda, R. Caceres, C. Taffernaberry, S. Pérez, L. Álvarez, A. Arena, G. Henderson, M. Funes, J. Catalá, "Diseño, Implementación y Análisis de Red de Distribución Inteligente y Generación Distribuida de Energía", Anales de CIIS 2016, May 2016, Universidad de Córdoba, Córdoba.
- [6] A. P. Arena, L. Álvarez, J. Da Peña, R. Stasi, G. López, A. Burlot, L. Barnabo, M. Ledda, R. Caceres, C. Taffernaberry and G. Mercado, "Diseño, Implementación y Análisis de una Red Eléctrica Inteligente (Smart Grid)", Anales del ENIDI 2017, Mendoza 22 al 24 de Noviembre 2017
- [7] Solvers Box Medidores de Energía Eléctrica, <https://solversbox.com/> visitado 10-Noviembre-2022
- [8] Node MCU 1.0 docs.platformio.org/en/stable/boards/espressif8266/nodemcu-v2.html visitado 10-Noviembre-2022
- [9] PlatformIO - Professional collaborative platform for embedded development. platformio.org/ visitado 20-Nov-2022
- [10] Joshi, Manveer, and Bikram Pal Kaur. "Coap protocol for constrained networks." International journal of wireless and microwave technologies 5.6 (2015): 1-10.
- [11] Soni, Dipa, and Ashwin Makwana. "A survey on mqtt: a protocol of internet of things (iot)." International conference on telecom, power analysis and computing (ICTPACT-2017). Vol. 20. 2017.
- [12] MongoDB Atlas, www.mongodb.com/ visitado 12-Oct-2022
- [13] Express Infraestructura web para Node.js <https://expressjs.com/es/> visitado 6-Oct-2022
- [14] Houston, Gary. "ISO 8601: 1988 Date/Time Representations." (1993).

Utilización Eficiente y Sostenible de Recursos en Clúster de Bajo Costo.

Karina Cenci, José Moyano, Andres Salamanca

Laboratorio de Investigación en Sistemas Distribuidos (LISiDi)
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Bahía Blanca, Argentina
e-mails: {kmc, jose.moyano, cas}@cs.uns.edu.ar

Resumen

La computación en la nube permite en forma conveniente y bajo demanda, el acceso a un conjunto de recursos compartidos y personalizables que pueden ser rápidamente suministrados y liberados. La inmersión y omnipresencia de las computadoras en el mundo requiere que consideremos cómo impactan en el mismo. La sostenibilidad es uno de los aspectos primordiales al momento de considerar soluciones, ya que no sólo involucra al medioambiente sino integralmente a un conjunto de dimensiones, como la tecnología.

En la actualidad, la gran mayoría de las aplicaciones, están pensadas y desplegadas para que funcionen en entornos web y móviles. El éxito está relacionado con el tiempo de respuesta, escalabilidad y disponibilidad. Para alcanzar estas metas, es necesario contar con sistema distribuido que gestione el despliegue de los servicios.

A medida que las necesidades de recursos computacionales se trasladan de los dispositivos de los usuarios hacia los grandes *datacenters*, y que el crecimiento de estos centros de cómputo crece en proporción a su demanda, la sociedad cuestiona la sustentabilidad del modelo. El impacto al medioambiente que tienen los centros de cómputo está relacionado a la energía y el espacio requeridos, así como los desperdicios que producen los

repuestos y las actualizaciones tecnológicas.

Palabras clave: Sistemas Distribuidos, Contenedores, Eficiencia, Sostenible

Contexto

Esta línea de investigación y desarrollo se realiza en el ámbito del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur (UNS). En particular, como parte de las tareas que se realizan en el Laboratorio de Investigación en Sistemas Distribuidos (LISiDi). El proyecto se financia parcialmente con fondos asignados por la UNS a proyectos de investigación.

1. Introducción

La omnipresencia y ubicuidad de los dispositivos en la vida cotidiana hacen que la sociedad consuma los servicios que ofrecen los distintos organismos de agencias públicas y privadas. La computación en la nube (Cloud Computing) permite en forma conveniente y bajo demanda, acceso a la red sobre un conjunto de recursos informáticos compartidos y personalizables que pueden ser rápidamente suministrados y liberados con un mínimo esfuerzo de gestión o interacción del proveedor de servicios. Estos sistemas permiten a miles de computadoras distribuidas por Internet ser compartidas por los usuarios,

teniendo en cuenta aspectos de escalabilidad, seguridad, disponibilidad, tolerancia a fallas, sistemas operativos, apoyo a la programación paralela, servicios de búsqueda, administración de sistemas, reserva de recursos, heterogeneidad y muchos más.

La inmersión de las computadoras en el mundo requiere que consideremos el impacto que tiene en el mundo. Uno de los aspectos en los cuales es importante trabajar es en la sostenibilidad, que no sólo involucra al medioambiente sino integralmente a un conjunto de dimensiones. La sostenibilidad se la define como un concepto sistémico y debe entenderse en un conjunto de dimensiones, que incluyen social, individual, ambiental, económica y técnica. La tecnología es parte del dilema para el desarrollo de soluciones sostenibles para mejorar el estado actual y futuro de nuestro mundo.

En la actualidad, la gran mayoría de las aplicaciones, especialmente aplicaciones web y móviles, centran su funcionamiento en una arquitectura cliente-servidor, donde la aplicación web o *app* móvil es el cliente de un sistema preparado para servir desde miles hasta millones de clientes. El éxito de estas aplicaciones depende en igual proporción de satisfacer tanto requerimientos funcionales como no funcionales. Entre los requisitos no funcionales, resultan fundamentales el tiempo de respuesta, la escalabilidad y la disponibilidad. Para lograr este conjunto de requerimientos no funcionales, la aplicación se despliega sobre un sistema distribuido, construido como un clúster, manteniendo réplicas del servicio. Estas réplicas mejoran el tiempo de respuesta al balancear las solicitudes entre cada una de ellas; permiten una alta disponibilidad, al contar con múltiples copias que pueden reemplazar una en caso de fallo; y facilita la escalabilidad con la

incorporación de recursos físicos adicionales al clúster, para aumentar las réplicas y la capacidad.

A medida que las necesidades de recursos computacionales se trasladan de los dispositivos de los usuarios hacia los grandes *datacenters*, y que el crecimiento de estos centros de cómputo crece en proporción a su demanda, la sociedad cuestiona la sustentabilidad del modelo. El impacto al medio ambiente que tienen los centros de cómputo está relacionado a la energía y el espacio que requieren para funcionar, así como los desperdicios que producen los repuestos y las actualizaciones tecnológicas. Una de las posibles formas de abordar esta problemática se encuentra en mejorar la sustentabilidad de hardware. Esta mejora se propone a partir de la migración de las plataformas físicas convencionales hacia *System-on-Chip* (SoC) y *single-board computers*. Estas dos opciones mejoran significativamente el desempeño de la plataforma en relación a la cantidad de energía consumida, requieren mucho menos espacio para operar, y reducen la cantidad de desperdicios. Sin embargo, este enfoque presenta nuevos desafíos técnicos que deben ser abordados.

El resto de este artículo se organiza de la siguiente manera: la sección 2 explica los objetivos de investigación, la sección 3 los resultados esperados/obtenidos y, por último, la sección 4 discute la formación de recursos humanos.

2. Líneas de Investigación y Desarrollo

El estudio de los sistemas distribuidos y paralelos, los paradigmas de computación en la nube, *peer-to-peer*, y movilidad han adquirido gran importancia en distintos ambientes como son los sistemas orientados a la

información y comunicación, los sistemas de censado del ambiente como la temperatura y humedad, para mejorar la producción y calidad de vida. En este proyecto se van a seguir dos líneas de investigación:

- Utilización de contenedores para la extensibilidad y escalabilidad de servicios en una nube privada sobre un clúster con dispositivos de pequeña escala. El estudio, análisis y desarrollo de un *framework* que soporte el diseño y la implementación de servicios que puedan replicarse haciendo uso eficiente de los recursos de manera que balancee las distintas cargas como, por ejemplo: procesamiento, transferencia de datos, consumo energético.
- Automatización para la recuperación y tolerancia a fallas en un clúster con dispositivos de pequeña escala. El mantenimiento y actualización de clúster requiere la utilización de herramientas que faciliten el control y verificación del estado de los componentes. La estabilidad de la infraestructura es esencial para la disponibilidad de los servicios y para el uso eficiente de los mismos.

El contexto de estos desarrollos se puede aplicar en organizaciones, empresas de pequeña y mediana escala para mejorar los procesos del negocio. La utilización de contenedores e infraestructura de bajo costo facilita el testeo y pruebas en entornos similares a los de producción. Además, la facilidad de brindar servicios que estén distribuidos y replicados promueve el uso eficiente y sostenible de los recursos.

3. Resultados Obtenidos/Esperados

El diseño e implementación de la plataforma del clúster es el primer paso realizado para el desarrollo de las actividades [7]. El proceso de construcción de esta plataforma conlleva varias etapas: *inicio*, *en funcionamiento* y *automatización*. Las metas requeridas son:

- **Sustentabilidad:** Hardware de bajo costo de adquisición, costo bajo de mantenimiento, bajo consumo energético, instalable en espacios reducidos o compartidos.
- **Ambiente distribuido:** la selección de la plataforma de software que provea propiedades de un ambiente distribuido.
- **Disponibilidad y Facilidad en el Mantenimiento:** diseño de procesos repetibles, robustos y autodocumentados para la instalación y monitoreo del sistema; soporte de telemetría del sistema; características preparadas para el cambio continuo.

El primer ciclo en el diseño de la infraestructura se encuentra en el diseño de testeos para verificar la automatización y la deuda de hardware, algunos de los inconvenientes que presentan este tipo de dispositivos es la fragilidad y la limitada robustez de los mismos, que en algunos casos está asociado al bajo costo. La estabilidad y funcionamiento de la infraestructura permite avanzar en el diseño de la capa de software como middleware soporte para la alcanzar disponibilidad y tolerancia a fallas.

Los dispositivos utilizados tienen restricciones para la ejecución de aplicaciones, los primeros despliegues para el análisis de la replicación de servicios se están realizando con contenedores, específicamente con Docker swarm.

4. Formación de Recursos Humanos

En relación con la formación de los recursos humanos, el trabajo a desarrollar permite una capacitación práctica adecuada de los profesionales involucrados, permitiendo plasmar con logros concretos las investigaciones realizadas. Además, esta línea de investigación permitirá la dirección de tesis de licenciatura y trabajos finales de ingeniería afines. Asimismo, servirán para generar courseware para materias optativas a dictarse en el Departamento de Ciencias e Ingeniería de la Computación de la UNS.

5. Referencias

- [1.] Bashari Rad, Babak & Bhatti, Harrison & Ahmadi, Mohammad. (2017). An Introduction to Docker and Analysis of its Performance. *IJCSNS International Journal of Computer Science and Network Security*. 173. 8.
- [2.] W. Felter, A. Ferreira, R. Rajamony and J. Rubio, "An updated performance comparison of virtual machines and Linux containers," *2015 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, Philadelphia, PA, USA, 2015, pp. 171-172, doi: 10.1109/ISPASS.2015.7095802.
- [3.] Sheng Li, Kevin Lim, Paolo Faraboschi, Jichuan Chang, Parthasarathy Ranganathan, and Norman P. Jouppi. 2011. System-level integrated server architectures for scale-out datacenters. In *Proceedings of the 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-44)*. Association for Computing Machinery, New York, NY, USA, 260–271. <https://doi.org/10.1145/2155620.2155651>
- [4.] R. Morabito, "Power Consumption of Virtualization Technologies: An Empirical Investigation," *2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC)*, Limassol, Cyprus, 2015, pp. 522-527, doi: 10.1109/UCC.2015.93.
- [5.] Amit M Potdar, Narayan D G, Shivaraj Kengond, Mohammed Moin Mulla, Performance Evaluation of Docker Container and Virtual Machine, *Procedia Computer Science*, Volume 171, 2020, Pages 1419-1428, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2020.04.152>.
- [6.] Shuja, Junaid & Bilal, Kashif & Madani, Sajjad & Othman, Mazliza & Ranjan, R. & Balaji, Pavan & Khan, Samee. (2016). Survey of Techniques and Architectures for Designing Energy-Efficient Data Centers. *IEEE Systems Journal*. 10. 507-519. 10.1109/JSYST.2014.2315823.
- [7.] J. Moyano, C. A. Salamanca, K. Cenci, J. Ardenghi. Evaluación de Docker en un Cluster de Bajo Costo y Alta Disponibilidad, en 10mo Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI). AJEA, (17), 1–2021.
- [8.] R. Buyya, C. Yeo, and S. Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities," in *High Performance Computing and Communications, 2008. HPCC'08. 10th IEEE International Conference on*. IEEE, 2008, pp. 5–13.
- [9.] Imtiaz Ahmad, Mohammad Gh. AlFailakawi, Asayel AlMutawa, Latifa Alsalman, Container scheduling techniques: A Survey and assessment, *Journal of King Saud University - Computer and Information Sciences*, Volume 34, Issue 7, 2022, Pages 394-3947, ISSN 1319-1578, <https://doi.org/10.1016/j.jksuci.2021.03.002>
- [10.] An Antonio Jose Alves Neto, Jose Aprigio Carneiro Neto, and Edward David Moreno Ordonez. 2022. Low-cost clusters on big data - A systematic study. In *11 Euro American Conference on Telematics and Information Systems (EATIS2022)*, June 1–3, 2022, Aveiro, Portugal. ACM, New York, NY, USA 7 Pages. <https://doi.org/10.1145/3544538.3544635>.
- [11.] Bibal Benifa J.V., Dejeay Dharma, HAS: Hybrid auto-scaler for resource scaling in cloud environment, *Journal of Parallel and Distributed Computing*, Volume 120, 2018, Pages 1-15, ISSN 0743-7315, <https://doi.org/10.1016/j.jpdc.2018.04.016>.
- [12.] Luiz F. Bittencourt, Alfredo Goldman, Edmundo R.M. Madeira, Nelson L.S. da Fonseca, Rizos Sakellariou, Scheduling in distributed systems: A cloud

- computing perspective, *Computer Science Review*, Volume 30, 2018, Pages 31-54.
- [13.] Botma, Eugene and Eduan Kotzé. "Feasibility of a low-cost computing cluster in comparison to a high-performance computing cluster: A developing country perspective." *CONF-IRM* (2016).
- [14.] Carlos de Alfonso, Amanda Calatrava, Germán Moltó, Container-based virtual elastic clusters, *Journal of Systems and Software*, Volume 127, 2017, Pages 1-11, ISSN 0164-1212.
- [15.] Divya. V, Leena Sri. R, Docker based Intelligent Fall Detection using Edge-Fog Infrastructure, *IFAC-PapersOnLine*, Volume 53, Issue 1, 2020, Pages 45-50, ISSN 2405-8963.
- [16.] Rekha Gangula, G. Jagadeesh Kumar, Sesham Anand, A.V. Krishna Prasad, Integration of dynamic Docker containers and kubernetes with advanced cloud and Internet of Things, *Materials Today: Proceedings*, 2021, ISSN 2214-7853, <https://doi.org/10.1016/j.matpr.2021.07.276>
- [17.] Monika Gill, Dinesh Singh, ACO Based Container Placement for CaaS in Fog Computing, *Procedia Computer Science*, Volume 167, 2020, Pages 760-768, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2020.03.406>.
- [18.] Dominik Göddeke, Dimitri Komatitsch, Markus Geveler, Dirk Ribbrock, Nikola Rajovic, Nikola Puzovic, Alex Ramirez, Energy efficiency vs. performance of the numerical solution of PDEs: An application study on a low-power ARM-based cluster. *Journal of Computational Physics*, Volume 237, 2013, Pages 132-150, ISSN 0021-9991, <https://doi.org/10.1016/j.jcp.2012.11.031>.
- [19.] H. Jiang, K. Wang, Y. Wang, M. Gao and Y. Zhang, "Energy big data: A survey," in *IEEE Access*, vol. 4, pp. 3844-3861, 2016, doi: 10.1109/ACCESS.2016.2580581.
- [20.] Kumari, Priti and Parmeet Kaur. "A survey of fault tolerance in cloud computing." *J. King Saud Univ. Comput. Inf. Sci.* 33 (2021): 1159-1176.
- [21.] A. Martin, S. Raponi, T. Combe, R. Di Pietro, Docker ecosystem – Vulnerability Analysis, *Computer Communications*, Volume 122, 2018, Pages 30-43, ISSN 0140-3664, <https://doi.org/10.1016/j.comcom.2018.03.011>.
- [22.] Partha Pratim Pande and Amlan Ganguly. 2012. Introduction to the special issue on sustainable and green computing systems. *J. Emerg. Technol. Comput. Syst.* 8, 4, Article 26 (October 2012), 3 pages. <https://doi.org/10.1145/2367736.2367737>

Tecnologías de Smart IoT para la solución de problemas en el medio socio productivo

Jorge Osio^{1,2}, Juan Salvatore¹, Mauro Salina¹, Marcelo Cappelletti^{1,2}, Diego Montezanti^{1,3}, Lucas Olivera^{1,4}, Martín Paez¹, Franco Palacio¹, Matías Busum Fradera¹, Facundo Chazarreta¹, María Florencia Ayala¹, Diego Encinas^{1,3}, Martín Morales^{1,5}

¹ Programa TICAPPS, Univ. Nac. Arturo Jauretche, Florencio Varela (1888), Argentina.

² Grupo de Control Aplicado, Instituto LEICI (UNLP-CONICET), La Plata (1900), Argentina.

³ Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC.

⁴ Becario de la Comisión de Investigaciones Científicas, Pcia. Bs. As. (CICPBA), La Plata (1900), Argentina.

⁵ Centro UTN CODAPLI-FRLP, La Plata (1900), Argentina.

{josio, mcappelletti, dmontezanti, jsalvatore, mdsalina, dencinas, martin.morales}@unaj.edu.ar

Resumen

El presente proyecto se basa en la utilización de internet de las cosas (IoT) como herramienta fundamental para proveer soluciones a problemáticas del medio socio productivo, como son el cuidado del medioambiente y la innovación en el sector productivo, focalizando la investigación en las técnicas de Smart IoT.

Entre los temas de investigación que se desarrollan en esta línea, se incluye el diseño e implementación de técnicas de visión por computadora con el objeto de agregar funcionalidades a dispositivos robóticos. Además, mediante el agregado de control y supervisión remota (con herramientas de IoT), se busca proveerles autonomía para la realización de determinadas tareas. En esta línea, también se están implementando técnicas de visión por computadora para la clasificación de residuos reciclables mediante algoritmos de aprendizaje automático. Además, se utilizan técnicas de aprendizaje profundo y visión por computadora para la clasificación de diferentes condiciones del cielo, con el objetivo de optimizar sistemas que aprovechan la energía solar.

En la misma línea de Smart IoT, se está trabajando en el desarrollo de una plataforma de servicios IoT que contempla los servicios más importantes. Esta plataforma permite la obtención de datos provenientes de una red de sensores, lo que posibilitará, mediante

técnicas de aprendizaje automático, la implementación de un sistema de ayuda a la toma de decisiones para optimizar y mejorar el aprovechamiento de los recursos.

Por último, actualmente se están investigando los aspectos de seguridad en cada una de las capas del sistema, debido a que éste se ha vuelto un aspecto de gran relevancia en los sistemas de IoT.

Palabras clave: *Aprendizaje automático, procesamiento de imágenes, Smart IoT, Visión por computadora.*

Contexto

Las líneas de I/D presentadas en este trabajo pertenecen al grupo SIoT&IA y están incluidas dentro del Programa TICAPPS (TIC en aplicaciones de interés social) de la Universidad Nacional Arturo Jauretche (UNAJ), Resolución N° 064/17, bajo la dirección del Dr. Ing. Martín Morales.

1. Introducción

En los últimos años ha surgido el concepto de Smart IoT (SIoT), favorecido por las mejoras de los procesadores para sistemas embebidos que los hacen aptos para la ejecución de algoritmos de aprendizaje automático [1]. La gran ventaja que aportan es que posibilitan la resiliencia del sistema cuando se pierde

conexión con el servidor permitiendo lograr una alta disponibilidad y por otro lado los dispositivos almacenan y analizan sus propios datos, de manera que las decisiones se pueden tomar localmente y luego comunicar las acciones al servidor para mantener la integración del sistema.

Las tecnologías de internet de las cosas son necesarias para el almacenamiento y uso de grandes cantidades de datos en la nube, por otro lado, el aprendizaje automático permite el entrenamiento de máquinas para la realización de tareas y la visión por computadora permite la manipulación de objetos para automatizar actividades y procesos.

Dentro de la Visión por Computador, la detección de objetos es uno de los temas más candentes. Para entender el problema hay que pensar en cómo queda codificada una imagen digital. En base a esto, cuando una máquina busca un objeto dentro de una imagen lo que realmente hace es buscar patrones que se correspondan con el objeto en particular.

Si dotamos a estas herramientas inteligentes con tecnologías como visión 3D, imágenes multiespectrales y lo combinamos con aprendizaje automático, para un análisis de datos más profundo y una predicción de modelos más precisa, conseguiremos una inteligencia artificial que dote de mayor flexibilidad a los equipos electrónicos para la toma de decisiones. De esta manera, se desarrolla la capacidad de configurar un robot a través del aprendizaje automático. El concepto de internet de las cosas ha adquirido gran relevancia en los últimos años, debido a la posibilidad que ofrece de interconectar objetos entre sí y la conectividad a internet que provee a las redes de objetos [2]. En los últimos años ha surgido el concepto de Internet de Objetos Inteligentes (SIoT), que promete importantes avances en diferentes áreas de aplicación. Para este tipo de aplicaciones se dispone de un conjunto de protocolos de comunicación e interfaces que se pueden implementar mediante herramientas de sistemas embebidos como las desarrolladas en [3-6]. Las investigaciones realizadas mediante smart IoT se podrán aplicar en el

automatismo y monitoreo de procesos, almacenamiento masivo de datos de sensores y dispositivos de cómputo (Big Data).

Respecto al aprendizaje automático [6], una de las áreas en donde se avanzó notablemente es en la de detección de objetos y clasificación de imágenes. Esto se debe en su mayor parte al desarrollo de nuevas técnicas de Machine Learning (Aprendizaje Automático) como el Deep Learning (Aprendizaje Profundo) [7], además de las innovaciones en el manejo de datos a gran escala y la amplia variedad de servicios que se generan mediante cloud computing (computación en la nube) para el análisis y procesamiento de información. Algunos ejemplos de dichos avances en el área se aplican en: control de vehículos autónomos, detección de rostros, diagnóstico de enfermedades, realidad aumentada, etc.

Machine Learning es un subcampo de la IA [8]. Los algoritmos de Machine learning permiten resolver problemas de forma intuitiva y automatizada, sin que el mecanismo de elección se encuentre previamente programado. Por su parte, Deep Learning es un subcampo de Machine Learning, pero existen técnicas de Machine Learning que no utilizan Deep Learning. Este último es utilizado para realizar procesos de Machine Learning empleando redes neuronales artificiales compuestas por varios niveles jerárquicos [7]. Básicamente es como implementar niveles de aprendizaje partiendo de lo más obvio para luego combinarlo posibilitando lograr obtención de características más complejas. Las técnicas de machine learning y deep learning proveen gran soporte para el diseño de aplicaciones de visión por computadora, es por eso que son parte fundamental de la propuesta.

Por último, la importancia de disponer un servidor de servicios IoT propio, permite no solo obtener grandes cantidades de datos, sino realizar un análisis y procesarlos sin depender de servicios externos, esto no solo permite ganar en eficiencia sino también incrementar la calidad del servicio [9]. En esta instancia es de suma importancia considerar los aspectos de seguridad en las plataformas IoT, es por

esto que se considera fundamental la inclusión de esta nueva línea como parte de todo sistema IoT.

Aplicaciones

Entre las aplicaciones que se están desarrollando mediante IoT, procesamiento de datos, imágenes y machine learning se tiene; la detección automática de residuos reciclables, la detección de malezas, técnicas de eficiencia energética, calidad del aire, obtención de magnitudes climatológicas, radiación solar, visión artificial aplicada a robótica y la clasificación de diferentes condiciones de cielo [10].

El reciclado de desechos es considerado de suma importancia para el cuidado del medioambiente, debido a que, supone la reutilización de elementos u objetos ya utilizados, los que de otro modo serían desechados contribuyendo al incremento de la basura y al daño ambiental permanente. Gran parte del problema radica en el esfuerzo que requiere clasificar y separar los residuos inorgánicos, es por eso que en esta línea de investigación se propone desarrollar un sistema basado en redes neuronales artificiales que permita detectar y clasificar los objetos reciclables más comunes como papel, cartón, botellas, latas, etc, y los materiales con los que están hechos (plástico, vidrio, metal, papel) [11].

Respecto a la visión por computadora, se está llevando adelante la implementación de un sistema robótico que sea capaz de detectar obstáculos e identificar señales, mediante procesamiento de imágenes, para dotarlo con capacidades autónomas [12]. Además de un brazo robótico que permite seleccionar objetos previamente identificados mediante visión artificial [12]. En este sentido se incluye la detección automática de gestos como una forma de comunicación, generando comandos y acciones a distancia sin necesidad de emitir sonido ni acción física (pensado para personas). Adicionalmente, se ha desarrollado un sistema de control para el sensado de gases y el control automático de iluminación y ventilación, de tal forma de

contribuir con la eficiencia energética y calidad del aire y calidad del aire en ambientes cerrados [13].

Por otro lado, se implementó una red de sensores IoT que permiten obtener valores de más de 10 variables climáticas, como presión atmosférica, humedad del suelo, humedad ambiente, temperatura, intensidad lumínica, niveles de dióxido de carbono, entre otras. Lo que permitirá la implementación de un sistema inteligente que logre optimizar las plantaciones de cultivos en la zona.

Finalmente, la cobertura de nubes es un factor que atenúa y ocasiona intermitencia en los valores de la energía proveniente del sol [10]. La aparición de grandes nubes que tapan el disco solar produce cambios repentinos en los valores de la radiación solar. Esta variabilidad provoca que la producción de energía que pueden generar los sistemas que aprovechen esta fuente como recurso energético no sea constante ni totalmente predecible en el tiempo. Sobre este punto, se ha desarrollado un sistema capaz de clasificar, mediante técnicas de aprendizaje profundo y visión por computadora, diferentes condiciones de cielo como consecuencia de la cobertura de nubes, lo cual será de suma utilidad para sistemas que utilicen la energía solar.

2. Líneas de Investigación y Desarrollo

El grupo de investigación que se ha constituido en la UNAJ es multidisciplinario, y sus miembros cuentan con experiencia en sistemas de cómputo de altas prestaciones, tolerancia a fallo, procesadores embebidos, IoT, aprendizaje automático, y robótica.

En cuanto a la robótica [12], esta línea está en pleno desarrollo y se buscan aplicaciones innovadoras como la detección de objetos para reciclado y la capacidad de detección de señales y obstáculos para el funcionamiento autónomo mediante procesamiento de imágenes [14].

Las líneas de I/D que se presentan en este trabajo están basadas en el estudio y desarrollo de herramientas alternativas para el

procesamiento de imágenes a partir de algoritmos de procesamiento, aprendizaje automático y visión por computador. Respecto a los sistemas smart IoT se pretende en una primera instancia evaluar las herramientas tecnológicas disponibles para determinar la limitación y las posibles aplicaciones, adicionalmente se considera de vital importancia contemplar los aspectos de seguridad en todo sistema IoT, para evitar ataques y limitaciones en el servicio.

Temas de Estudio e Investigación

Los datos recopilados de la robótica y de los sensores relacionados con el cultivo pueden suponer grandes cantidades de información dependiendo de los sensores de los que disponga y de la integración con otros sistemas. Esto permitirá conseguir una mayor precisión en los sistemas de Machine Learning y, cuantos más sean los datos útiles con los que alimentemos el modelo, mejor será su aprendizaje y, por ende, sus predicciones.

Los aportes originales no solo tienen que ver con el manejo de la información, sino también con las aplicaciones innovadoras que se le puede dar al aprendizaje automático en visión por computadora y robótica, donde el ejemplo más claro se muestra en la posibilidad de disponer de un sistema de reciclaje automático. También se puede innovar respecto a la visión por computador para conseguir la realización de tareas específicas mediante dispositivos robóticos en procesos particulares de la industria local. Por último, la detección automática de gestos puede ser una forma de comunicación para personas con problemas motrices y del habla, pero también puede ser aplicable en lugares inhóspitos y de reducida movilidad.

En cuanto al análisis de los datos climáticos en zonas de cultivos, permitirá estudiar cuáles de las técnicas de aprendizaje automático logra adecuarse mejor para determinar posibles problemas, y por ende soluciones en cultivos futuros. Por otro lado, se están logrando avances en la determinación de la atenuación de la radiación solar global, a

partir de la clasificación de diferentes condiciones de cielo (despejado, parcial o totalmente nublado) y de variables meteorológicas de sencilla adquisición.

En relación a los aspectos de seguridad en sistemas IoT, esta se considera un aspecto fundamental en dichos sistemas, es por eso que se deberá evaluar:

- La seguridad propuesta por el protocolo mqtt basada en SSL que permite el encriptado entre broker y dispositivo.
- Si el servidor es de un tercero se deberán evaluar los canales de encriptado punto a punto (de extremo a extremo) para evitar el acceso a la información crítica.
- Las formas de garantizar la privacidad de los datos del cliente.
- Evaluar la utilidad de una blockchain, (ejemplo BFA)
- Evaluar la seguridad de los dispositivos finales y cómo se pueden comunicar de forma segura con el servidor en cuanto a integridad (algoritmo de Hash), confidencialidad (encriptado simétrico - AES) y autenticidad.

3. Resultados y Objetivos

Resultados alcanzados:

La detección y clasificación de objetos reciclables cuenta con la implementación de los primeros modelos basados en la arquitectura de redes neuronales convolucionales (CNN) con las que se alcanzaron tasas de acierto cercanas al 75%. Respecto a las pruebas desarrollando modelos que hacen uso de las técnicas de aprendizaje por transferencia (Transfer Learning), se logró una mejora en los modelos iniciales elevando la tasa de acierto a un 80% para los modelos entrenados desde cero y casi un 92% en modelos basados en Transfer Learning. Por otro lado, en aplicaciones de detección de objetos defectuosos se obtuvieron buenos resultados mediante el método `get_avg_precision_at_iou`, la performance del modelo entrenado fue de mAP 0.76, es decir, una precisión del 76%. Luego, se realizó un segundo entrenamiento con el modelo Faster

R-CNN Mobilenetv3 Large FPN para reducir la potencia computacional a cambio de la pérdida de precisión (ideal para sistemas de cómputo de bajas prestaciones, como por ejemplo la Raspberry Pi 3 Model B+, en donde se ejecutaron las aplicaciones [10].

Adicionalmente, se están realizando las primeras pruebas de estos modelos utilizando arquitecturas más simplificadas como MobileNet en microcontroladores Esp32-cam, inclusive se espera realizar pruebas de detección de varios objetos en la imagen o video, pudiendo realizar una detección en tiempo real.

En cuanto a las aplicaciones en robótica, actualmente se está trabajando con un robot móvil con cámara y un brazo robótico con esp32 cam. Respecto al robot móvil, se consiguió implementar el guiado autónomo mediante detección de líneas en imágenes usando procesamiento en tiempo real, en donde se está perfeccionando el algoritmo en las partes con curvas pronunciadas, para esto se está utilizando otro algoritmo de control. En cuanto al brazo, este pudo ser controlado desde el esp32 cam de forma remota, accediendo a una página web embebida con los botones de movimiento disponibles, el siguiente paso es mapear la grilla de la imagen con las coordenadas que se le pasan al brazo para que cada vez que se detecta un objeto el brazo lo pueda capturar y trasladar.

Por otra parte, se logró implementar una red sensores basada en IoT que permite obtener información de magnitudes climatológicas, del estado del suelo y del agua de arroyos con el fin de a partir de la información obtenida generar alertas tempranas que serán enviadas a los usuarios. Esta información permitirá perfeccionar las técnicas de cultivo, optimizar la cosecha y detectar el vertido de desechos en arroyos. En esta línea de trabajo se están realizando mejoras mediante la migración de los servicios a la nueva plataforma de servicios IoT desarrollada, conjuntamente con las mejoras en el HW respecto al bajo consumo y la alimentación a batería.

Objetivos esperados:

- Promover la generación de conocimiento y el desarrollo de instrumentos que puedan aplicarse en la formación de recursos humanos para investigación.

- Fortalecer la actividad de investigación y vinculación que contribuya a remediar las problemáticas del territorio de influencia de la UNAJ.

- El objetivo principal de esta línea de investigación es desarrollar nuevas técnicas y obtener resultados favorables respecto a la detección de objetos y características del ambiente aplicando técnicas de visión por computadora y aprendizaje automático.

- Se espera contribuir con el cuidado del medio ambiente y la producción industrial, a través de la clasificación automática de objetos reciclables, la detección de productos defectuosos y la mejora de la producción agrícola mediante el análisis de variables climáticas a través de diferentes técnicas de Smart IoT.

- Teniendo desarrollada nuestra propia plataforma de servicios IoT, se buscará contribuir en todos los aspectos de seguridad desde el dispositivo final hasta el almacenamiento de datos en la nube y el acceso a los diferentes servicios por parte del usuario final. En este sentido, se evaluarán los métodos actuales y se pensarán alternativas para casos particulares.

4. Formación de Recursos Humanos

Uno de los principales objetivos del Programa TICAPPS, dentro de la temática de las líneas de I/D presentadas en este trabajo, es la formación de recursos humanos, tanto de docentes investigadores como de estudiantes.

Dentro de la temática de la línea de I+D, todos los miembros del proyecto participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

En este proyecto existe cooperación a nivel Nacional. Hay dos Doctores en Ingeniería, un Magister, un Especialista, cuatro Doctorandos y un integrante realizando su Maestría en temas relacionados.

Además, hay 4 estudiantes avanzados realizando las PPS de final de carrera en la temática. Actualmente, se encuentran en curso tres Tesis de Doctorado y tres becas de Estímulo a las Vocaciones Científicas del Consejo Interuniversitario Nacional (Becas EVC – CIN) y 2 becas UNAJ de investigación relacionadas directamente con las líneas de I/D presentadas.

5. Bibliografía

[1] M. Taha Yazici, S. Basurra and M. Medhat Gaber, "Edge Machine Learning: Enabling Smart Internet of Things Applications". School of Computing and Digital Technology, Birmingham City University, Birmingham, UK

[2] L. Fatmasari Rahman, "Choosing your IoT Programming Framework: Architectural Aspects", 2016 IEEE 4th International Conference on Future Internet of Things and Cloud

[3] J. Osio, M. Cappelletti, G. Suárez, L. Navarro, F. Ayala, J. Salvatore, D. Alonso, D. Encinas, M. Morales, "Diseño de aplicaciones de IoT para la solución de problemas en el medio socio productivo", UNSJ, San Juan, WICC 2019.

[4] J. Osio, J. Salvatore, M. Salina, M. Angel Cappelletti, D. Montezanti, N. Denon, S. Doti, L. Olivera, C. Botta, M. Busum Fradera, F. Chazarreta, D. Terceros Quiroz, D. Encinas, M. Morales. "Tecnologías de Smart IoT y aprendizaje automático para la solución de problemas en el medio productivo". XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Godoy Cruz, Mendoza), 2022.

[5] J. Osio, J. Salvatore, M. Salina, D. Montezanti, N. Denon, S. Doti, L. Olivera, M. Busum Fradera, D. Alonso, M. Angel Cappelletti, D. Encinas, M. Morales. "Tecnologías de IoT y aprendizaje automático para la solución de problemas en el medio productivo y el cuidado del medioambiente". XXIII Workshop de Investigadores en

Ciencias de la Computación (WICC 2021, Chilecito, La Rioja), 2021.

[6] J. Osio, J. Salvatore, D. Alonso, V. Guarepi, M. Cappelletti, M. Joselevich, M. Morales, "Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo", UNNE, Ciudad de Corrientes, WICC 2018.

[7] N. Buduma, "Fundamentals of Deep Learning", Editorial O'reilly. 2017

[8] A. Gerón, "Hands-On Machine Learning with scikit Learn & TensorFlow", Editorial O'reilly, 2017

[9] F. Chazarreta, M. Busum Fradera, J. R. Osio, J. E. Salvatore, Morales D. M. "Diseño de una plataforma de servicios IoT para aplicaciones de interés social", IX CONAIISI, 2022.

[10] M Salina, JR Osio, MA Cappelletti, M Morales. *Aprendizaje automático aplicado al procesamiento de imágenes para la clasificación de objetos reciclables*. XXVII Congreso Argentino de Ciencias de la Computación (CACIC). 2021

[11] L Olivera, J Atia, L Amet, J Osio, M Morales, M Cappelletti, Uso de redes neuronales artificiales para la estimación de la radiación solar horaria bajo diferentes condiciones de cielo. Avances en Energías Renovables y Medio Ambiente 24, 232-243, 2020/12/30

[12] M. Jamshidi y P. J. Eicker, Robotics and remote system for hazardous environments, Prentice Hall, New Jersey, Estados Unidos (1993)

[13] Lo E., Fain F., Osio J., Cappelletti M., Aróztegui W., "Machine learning aplicada a variables ambientales", VII CONAIISI, 2020.

[14] W. Burger & M. J. Burge, "Digital Image Processing", Second Edition (2008).

Despliegue de aplicaciones contenerizadas: un caso de implementación basado en Crane

José Miguel Silva Pavon¹, Franco Bellino¹, Patricia Bazán³, Alejandra B. Lliteras^{2,4}, Jose Arcidiacono¹, Nicolás del Río¹

UNLP, Facultad de Informática¹, LINTI² UNLP, Facultad de Informática, LIFIA³ UNLP, Facultad de Informática, LINTI,⁴ CICIPBA

js.silva.010@gmail.com, fran85bellino@gmail.com, pbaz@info.unlp.edu.ar,
alejandra.lliteras@lifia.info.unlp.edu.ar, jarcidiacono@linti.unlp.edu.ar,
ndelrio@info.unlp.edu.ar

RESUMEN

Crane es una herramienta para el despliegue local de aplicaciones en contenedores, liviana, de propósito general y con capacidades de escalado automático, lo que la diferencia, por ejemplo, de la herramienta Minikube, la que permite algunas pruebas locales de la API de Kubernetes y se utiliza principalmente para el desarrollo de nuevas funcionalidades para K8s. El diseño de Crane presentado en [Arcidiacono et al., 2022] contempla la implementación de la herramienta con un componente de back-end para la creación y despliegue de contenedores y un componente de front-end orientado a un usuario final que debe desempeñar el rol de devop dentro de un proyecto [Httermann M, 2012].

El objetivo de este trabajo es proporcionar un servicio REST [Fielding & Taylor, 2002] para crear y desplegar servicios docker¹ con sus correspondientes herramientas de medición, definición de políticas de escalado y gestión de alertas para la toma de decisiones.

CONTEXTO

Este proyecto surge con el fin de implementar una herramienta para el despliegue local de aplicaciones en contenedores y que sea liviana, de propósito general y con capacidades de escalado automático.

DEHIA es un administrador de flujo de trabajo para la recopilación de datos con intervención humana [Arcidiacono et al., 2020]. Su arquitectura se basa en microservicios. En un primer intento de uso de DEHIA para la entrega en un servidor local, resultó ser complejo y difícil de replicar debido a sus diversos componentes y tecnologías.

Una posible solución a este problema es la contenerización [Bullington-McGuire et. al, 2020]. Para comenzar, se eligió el componente más simple, una puerta de enlace. Este componente no tiene dependencias funcionales con los otros componentes y no tiene estado interno.

El componente de puerta de enlace sólo necesita un puerto abierto (para recibir las solicitudes que tiene que redirigir) y espera un pequeño conjunto de parámetros. Por eso, era viable implementarlo automáticamente. Para ello se decidió desarrollar una herramienta de despliegue automático de Docker (denominada “Crane”), con el añadido de escalar la aplicación bajo demanda, creando nuevas instancias.

1- INTRODUCCIÓN

En la actualidad la separación entre la infraestructura y el desarrollo se ha reducido considerablemente con el uso de servicios

¹ <https://www.docker.com/>

PaaS [Rani & Ranjan, 2014] los cuales son entornos completos de desarrollo y despliegue en la nube, con recursos que permiten construir y escalar aplicaciones sin necesidad de preocuparse por la infraestructura.

Esta posibilidad hizo que muchos desarrolladores optaran por construir y desplegar aplicaciones directamente en la nube.

Un problema que surge con el despliegue en la nube, es que los costos que genera el alto rendimiento de estos servicios alojados en un PaaS imposibilitan a los desarrolladores hacer pruebas de rendimiento en etapas tempranas, construir y probar aplicaciones escalables o analizar el comportamiento de la misma como servicio distribuido.

Los desarrolladores se terminan “chocando” contra una gran pared al intentar simular esta infraestructura en un entorno local. Ya sea por el hecho de no poseer recursos para alojar los servicios o por el tiempo y la complejidad que lleva toda la implementación.

A su vez un despliegue local tiene dependencias, como por ejemplo versiones de librerías y sistemas operativos donde puede funcionar, esto supone una grave limitación a la hora de migrar el sistema a otro host, ya que se tendría que adaptar una solución local a cada uno de los entornos donde se vaya a ejecutar.

Con el uso de un ambiente de virtualización como Docker [Rad et al., 2017] podemos desplegar la solución en contenedores los cuales se podrán migrar a otro host independientemente de sus características, siempre y cuando tenga Docker instalado.

Una vez realizado el despliegue, surge la necesidad de orquestar, medir y escalar estos contenedores bajo demanda. Existen soluciones como Kubernetes [Burns et al., 2022] pero requiere una gran cantidad de recursos para funcionar, por lo que se hace

necesario que además de ser portable, use pocos recursos.

También se encuentra otro inconveniente y es que muchas veces para configurar contenedores es necesario crear previamente archivos de configuración estáticos, que son usados a la hora de desplegar los mismos y que dependen de rutas locales que no son iguales en todos los hosts.

Es por esto que se toma como base lo propuesto por [Arcidiacono et al., 2022] donde se habla sobre Crane una herramienta que permite optimizar el proceso de despliegue y control de servicios, implementado una API REST que siga esa pauta y permita interactuar y administrar aplicaciones completas de una manera flexible, simple y rápida, pudiendo así disminuir la distancia entre el desarrollo y el despliegue.

2- LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación presentada apunta a la creación de aplicaciones seguras, con reglas claramente definidas y en donde toda la complejidad del despliegue sea transparente al desarrollador, proponiendo una serie de configuraciones básicas, las cuales permitirán ganar tiempo y seguridad a la hora de crear entornos de desarrollo.

Como base de este proyecto, se ha tomado lo propuesto en [Arcidiacono et al., 2022] donde se habla sobre Crane, una herramienta que permite optimizar el proceso de despliegue y control de servicios, implementando una API REST para interactuar con el usuario.

El desarrollador enviará en el cuerpo de la solicitud los nombres de las imágenes del repositorio oficial de docker que desea utilizar y Crane API se encargará de crear, almacenar y desplegar un stack de servicios a partir de un archivo docker compose [Reis, D. et al., 2021].

El mismo será generado y ejecutado utilizando herramientas y librerías que proveen el lenguaje de programación utilizado. Luego se guardarán en memoria las aplicaciones creadas por el usuario con el fin de que puedan ser controladas, escaladas y monitorizadas.

En paralelo se le permitirá, al desarrollador, definir reglas y políticas para la toma de decisiones pudiendo así ajustar el consumo que realiza el servicio, sin desperdiciar/sobrecargar la infraestructura cuando el rendimiento baja/sube.

El proyecto incluye abordar la creación de contenedores mediante un servicio REST, la implementación de un Proxy reverso [Sommerlad et al., 2003] la integración con un sistema de gestión de métricas, la integración con un sistema de gestión de alertas, la integración con un sistema de gestión de políticas, la creación de archivos de configuración dinámicos y estáticos para despliegue de contenedores, la administración de las configuraciones de seguridad de los contenedores, la implementación de base de datos para la gestión de aplicaciones y el modelo RBAC [Ferraiolo et al., 1995] para la autenticación y autorización sobre el sistema

3- RESULTADOS ESPERADOS/OBTENIDOS

Se espera obtener un servicio que permita a los usuarios crear aplicaciones de forma fácil y rápida, manteniendo las mismas configuraciones y políticas en cualquier equipo donde se ejecute. El usuario mediante una API REST podrá crear contenedores que expondrán métricas y ajustar las reglas de escalado dependiendo de su propio análisis y/o reglas previamente definidas. Tendrá el control de los puertos expuestos de la aplicación,

tráfico de los contenedores y podrá clonar las configuraciones de aplicaciones para generar nuevas instancias de forma más rápida.

Además, al ser portable y de fácil acceso, el desarrollador podrá desplegarlo en cualquier servidor ejecutando un comando.

A su vez el usuario tendrá a su disposición los archivos de docker compose generados para poder llevarlos a cualquier lugar sin necesidad de tener que sacar a Crane del entorno local de desarrollo.

En la Figura 1 se representa el flujo de trabajo completo para el despliegue y monitoreo de las aplicaciones contenerizadas creadas a través de la API de Crane. Se observa que se usa Traefik² para el balanceo de carga entre las instancias del servicio, Prometheus³ para recopilar y almacenar métricas sobre el rendimiento y comportamiento. Y luego, la información obtenida será evaluada por Alert Manager y Open Policy Agent⁴ para tomar medidas en caso de que se cumpla alguna regla previamente definida.

4- FORMACIÓN DE RECURSOS HUMANOS

La diversidad de líneas de investigación enunciadas, así como la gran variedad de aplicación de los resultados esperados, propone un foco de investigación claro y concreto para la formación de recursos humanos, de una manera directa en el desarrollo de conocimientos en Docker y el manejo de contenedores, así como, conceptos de seguridad, comunicación entre contenedores o simplemente ser capaz de

² <https://traefik.io/>

³ <https://prometheus.io/>

⁴ <https://www.openpolicyagent.org/>

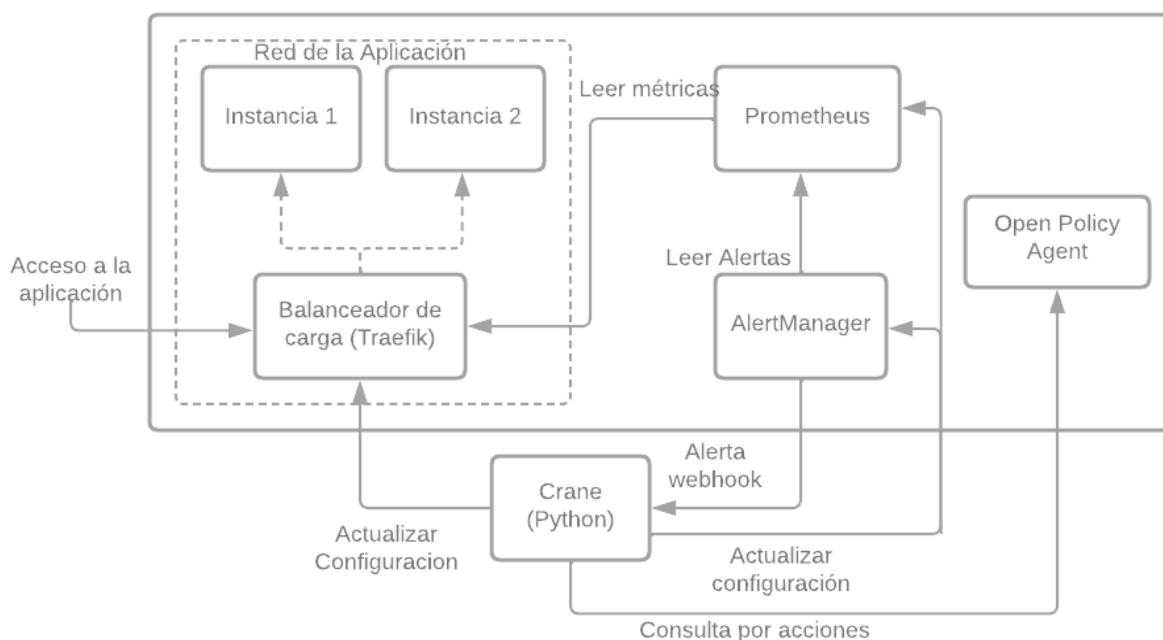


Fig. 1.- Arquitectura instanciada en Crane

analizar las métricas obtenidas por un determinado contenedor y tomar decisiones fundamentadas.

Debido a que este trabajo está cimentado sobre un proxy reverso, esto brinda muchos beneficios en la gestión y optimización de la infraestructura de la aplicación, tales como: balanceos de carga para optimizar al máximo los recursos disponibles, gestión de DNS, escalabilidad y también monitoreo.

También, al tener integrado un sistema de alertas, el recurso podrá recopilar métricas de rendimiento de los sistemas y configurar manualmente alertas personalizadas que, en caso de ser activadas, responderán dependiendo de unas políticas definidas.

Todo esto aporta conocimientos amplios sobre cómo se crean contenedores, cómo se comportan en cada momento y qué se puede hacer para mejorar su rendimiento o qué hacer en caso de fallo.

De una manera indirecta, es posible apoyarse en CRANE para el desarrollo de aplicaciones, que facilita un despliegue rápido del entorno

mediante una simple petición http, lo cual permitirá a los recursos humanos involucrados, formarse de una manera más rápida y efectiva, sin necesidad de pasar tiempo configurando un entorno que, en general, nada tiene que ver con el fin de su trabajo específico de desarrollador.

5- BIBLIOGRAFÍA

[Arcidiacono et al., 2022] Arcidiacono, J., Bazán, P., del Río, N., & Lliteras, A. B. (2022). Crane: A Local Deployment Tool for Containerized Applications. In Conference on Cloud Computing, Big Data & Emerging Topics (pp. 58-71). Springer, Cham.

[Arcidiacono et al., 2020] Arcidiacono, J., Lliteras, A. B., & Bazán, P. A. (2020). DEHIA, una plataforma para la generación y ejecución de actividades de recolección de datos con intervención humana aplicada en el Programa E-Basura. In VI Simposio Argentino de Ciencia de Datos y GRANdes DATos (AGRANDA 2020)-JAIIO 49 (Modalidad virtual).

[Bullington et al., 2020] Bullington-McGuire, R. and Dennis, A.K. and Schwartz, M. (2020). Docker for Developers: Develop and run your application with Docker containers using DevOps tools for continuous delivery. Packt Publishing.

[Burns et al., 2022] Burns, B., Beda, J., Hightower, K., & Evenson, L. (2022). Kubernetes: up and running. " O'Reilly Media, Inc."

[Ferraiolo et al.,1995] Ferraiolo, D., Cugini, J., & Kuhn, D. R. (1995, December). Role-based access control (RBAC): Features and motivations. In Proceedings of 11th annual computer security application conference (pp. 241-48).

[Fielding & Taylor, 2002] Fielding, R. T., & Taylor, R. N. (2002). Principled design of the modern web architecture. ACM Transactions on Internet Technology (TOIT), 2(2), 115-150.

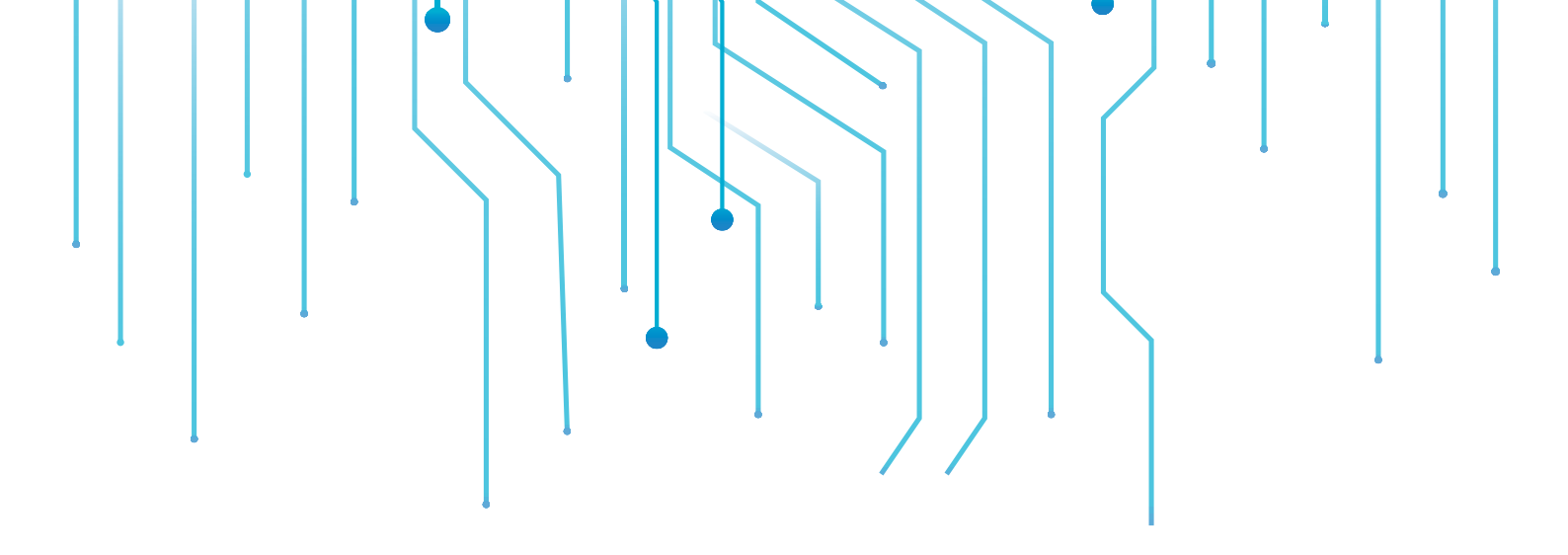
[Httermann, 2012] Httermann, M. (2012). DevOps for developers. Apress: delivers a practical, thorough introduction to approaches, processes and tools to foster collaboration between software development and operations

[Rad et al., 2017] Rad, B. B., Bhatti, H. J., & Ahmadi, M. (2017). An introduction to docker and analysis of its performance. International Journal of Computer Science and Network Security (IJCSNS), 17(3), 228.

[Rani & Ranjan, 2014] Rani, D., & Ranjan, R. K. (2014). A comparative study of SaaS, PaaS and IaaS in cloud computing. International Journal of Advanced Research in Computer Science and Software Engineering, 4(6).

[Reis et al., 2021] Reis, D., Piedade, B., Correia, F. F., Dias, J. P., & Aguiar, A. (2021). Developing docker and docker-compose specifications: A developers' survey. IEEE Access, 10, 2318-2329.

[Sommerlad et al., 2003] Sommerlad, P. (2003, June). Reverse Proxy Patterns. In EuroPLoP (pp. 431-458).



ASI

Agentes y Sistemas Inteligentes

Interfaces de Programación de Aplicaciones orientadas a la Integración de Sistemas Judiciales

Osvaldo Spositto¹, Julio Bossero¹, Viviana Ledesma¹, Gastón Procopio¹, Lorena Matteo¹, Cecilia Gargano¹, Victoria Saizar¹, Patricio Macias¹, Juan Ojeda¹, Fabio Quintana¹, Laura Conti², Sergio García³ y Gustavo Pérez Villar⁴

¹ Universidad Nacional de La Matanza. Departamento de Ingeniería e Investigación Tecnológicas.
{spositto, jbossero, vledesma, gprocopio, lmatteo, cgargano, vsaizar, pmacias, fquintana, jmojeda}@unlam.edu.ar

² Universidad Nacional de La Matanza. Departamento Derecho y Ciencia Política.
lconti@unlam.edu.ar

³ Palacio de Tribunales. Departamento Judicial de Morón.
sergiogabriel.garcia@pjba.gov.ar

⁴ Subsecretaría de Tecnología Informática del Poder Judicial de la Provincia de Buenos Aires.
gperez@scba.gov.ar

RESUMEN

Con la aparición de las API (por sus siglas en inglés Application Programming Interface), o interfaces de programación de aplicaciones, el intercambio de datos entre diferentes sistemas informáticos se ha convertido en una necesidad hoy en día. Los gobiernos ven esta transformación digital en una oportunidad para mejorar los entornos digitales existente y centran su atención en realizar mayores esfuerzos en la coordinación del intercambio de datos que garanticen una integración sinérgica entre diferentes aplicaciones.

En el campo de la Inteligencia Artificial, se entiende por Sistema Experto a un sistema informático que emula la capacidad de tomar de decisiones de un especialista humano.

Se presenta una modificación efectuada a un prototipo de Sistema Experto judicial, el cual tiene por objeto asistir a los distintos agentes de la justicia del fuero penal en el proceso de despacho judicial. En este trabajo se describe la creación de una API que permite crear una comunicación sincrónica entre sistemas que, hasta ahora, se realizaba en forma asincrónica.

Palabras clave: *Inteligencia Artificial, Sistemas Expertos, Interface de Programación de Aplicaciones.*

CONTEXTO

La línea de investigación aquí presentada es la continuación de los proyectos “Diseño e Implementación de un Sistema Experto como Apoyo al Proceso de Despacho de Trámites de un Organismo Judicial parte del proyecto” (2020-2021) e “Inteligencia Artificial Jurídica: la Evolución de Experticia hacia un Modelo de Justicia Predictiva” (2022-2023), ambos pertenecientes al Programa de Incentivos para Docentes Investigadores (PROINCE) de la Secretaría de Políticas Universitarias del Ministerio de Educación de la Nación. Los integrantes de este trabajo son investigadores y docentes dependientes de las siguientes Unidades Académicas de la Universidad Nacional de La Matanza (UNLaM): el Departamento de Ingeniería e Investigaciones Tecnológicas (DIIT), y el Departamento Derecho y Ciencia Política, además, colaboran personal técnico de la Subsecretaría de Tecnología Informática del Poder Judicial de la Provincia de Buenos Aires.

1. INTRODUCCIÓN

El concepto de Inteligencia Artificial (IA) surge en la década de 1950, y puede describirse, de manera simple, como una nueva forma de construir sistemas que tienen la capacidad de trabajar con datos y aprender de ellos para realizar tareas, simular el comportamiento humano y generar conocimiento [1]. Dentro de este concepto se encuadran los Sistemas Expertos (SE) que son sistemas informáticos que emulan la capacidad de toma de decisiones de un especialista humano. Estos sistemas están diseñados para resolver problemas complejos basados en ciertos razonamientos producto del conocimiento humano. Este conocimiento es un recopilado de la experiencia de expertos, un repositorio de reglas, que constituye una base de conocimiento, que representa hechos y reglas concretas con una fórmula determinista (si A entonces B).

En [2] se encuentra una descripción de cada uno de los componentes de un SE:

- Subsistema de adquisición de conocimiento. Es la acumulación, transferencia y transformación de la experiencia para resolver problemas de una fuente de conocimiento a un programa con la finalidad de construir o expandir la base de conocimiento. El estado del arte actual requiere un ingeniero en conocimiento que interactúe con uno o más expertos humanos para construir la base de conocimiento.
- Base de conocimiento - Contiene el conocimiento necesario para comprender, formular y resolver problemas. Incluye dos elementos básicos: heurística especial y reglas que dirigen el uso del conocimiento para resolver problemas específicos en un dominio particular.
- Base de hechos - Contiene los datos propios correspondientes a los problemas que se desean tratar.
- Motor de inferencia – ofrece la lógica

que permite pasar de las premisas a las conclusiones. Contiene la mayor cantidad de razonamientos válidos en ese dominio. Este componente provee direcciones sobre cómo usar el conocimiento del sistema para resolver el problema cuando se realiza una consulta. Cabe aclarar que el motor de inferencia es un sistema que debe ser capaz de tomar dos o más informaciones de la base de conocimiento y obtener una conclusión lógica.

Un prototipo de un SE judicial, denominado Experticia, se comenzó a diseñar en el año 2019, por este equipo de trabajo. El desarrollo de este trabajo se presentó en varios congresos en los años siguientes [3-6]. En la Figura 1 se muestra la estructura del prototipo desarrollado. Como se puede observar en dicha figura, los datos de los hechos y las reglas se encuentran ubicados en forma centralizada en los equipos que se encuentran en los juzgados.

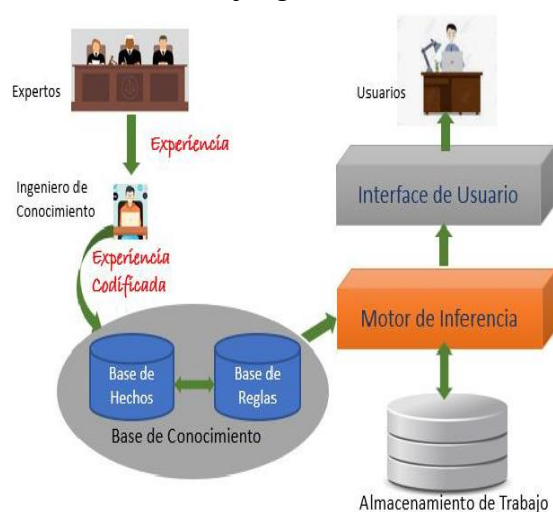


Figura. 1. Estructura básica de un SE.

El objetivo de Experticia se centra en el estudio e implementación de un SE que ayude con la resolución de expedientes judiciales en el Poder Judicial de la Provincia de Buenos Aires. Este organismo utiliza el Sistema Informático de Gestión Asistida Multi-fuero y Multi-Instancia (GAM), más conocido como

Augusta¹. Este es utilizado en toda la provincia, como única herramienta informática, para la asistencia integral en la gestión de causas de los organismos jurisdiccionales de diferentes instancias y fueros.

El prototipo se encuentra en ejecución en el Juzgado de Ejecución Penal Nro. 2 de Morón. En una primera etapa, para comenzar a procesar un expediente, los integrantes de un juzgado se conectan a Augusta para obtener los datos del mismo, y luego cargarlos en Experticia. Una vez que se procesan los mismos, y obteniendo o no, el dictamen final, el operador judicial debe conectarse, nuevamente, al sistema centralizado, para volcar en él, todas las actualizaciones efectuadas en el expediente. En conclusión, este proceso entre Experticia y Augusta se lleva a cabo en forma asincrónica.

Durante el año 2021 se trabajó conjuntamente con los responsables del Poder Judicial, principalmente, en llevar el proceso de intercambio de datos asincrónico, entre los dos sistemas, a otro que pueda actuar en forma sincrónica.

Para resolver el intercambio de datos, se recurrió a la creación de una API. Las API son un tipo de Servicio Web (SW) (Web Service o Web Services), que es una tecnología que utiliza un conjunto de protocolos y estándares que sirven para intercambiar datos entre aplicaciones [7].

La arquitectura propuesta para desarrollar el SW, fue REST, este término, es un acrónimo, cuya traducción al español significa *Transferencia de Estado Representacional*. Esta arquitectura puede ser utilizada en cualquier cliente HTTP [8]. La principal ventaja de esta arquitectura es que ha aportado a la web una mayor escalabilidad, es decir, da soporte a un mayor número de componentes y las interacciones entre ellos [9].

Las implementaciones de la REST también dependen de la noción de un conjunto de operaciones limitadas que tanto el cliente como el servidor entienden totalmente desde el comienzo. En el protocolo HTTP, las principales operaciones utilizadas son las siguientes [10]:

- GET: devuelve la información que se haya identificado mediante el URI² de solicitud.
- PUT: solicita que la entidad adjunta se almacene en el URI de solicitud suministrado.
- POST: solicita que el servidor de origen acepte la entidad adjunta en la solicitud como un nuevo subordinado del recurso identificado por el URI de solicitud.
- DELETE: solicita que el servidor de origen elimine el recurso identificado por el URI de solicitud.

Las primeras tres operaciones son de solo lectura, mientras que las últimas tres son operaciones de escritura [10].

Como se comentó, una API es un conjunto de funcionalidades o recursos que propone que sistema pueda interactuar con otro sistema, sin importar ni los lenguajes de programación en que cada uno de ellos esté escrito ni la tecnología que usan.

Para mantener la comunicación coherente entre estos diferentes dispositivos y plataformas, los datos deben formatearse utilizando un formato estándar, como JSON (JavaScript Object Notation en español Notación de Objetos JavaScript) o XML (Lenguaje de marcado extensible), entre otros [11].

El formato elegido, para este trabajo fue JSON, al igual que la mayoría de los lenguajes, estos son independientes del lenguaje programación que lo utilice, de hecho, sus convenciones y estructura se

¹ <https://www.scba.gov.ar/paginas.asp?id=39889>

² Universal Resource Identifier, o identificador universal de recursos.

asignan sin esfuerzo a lenguajes populares orientados a objetos como JavaScript, Ruby, Java y Python [12].

En la Figura 2, se representa el modelo propuesto utilizando JSON para conectarse con Augusta.

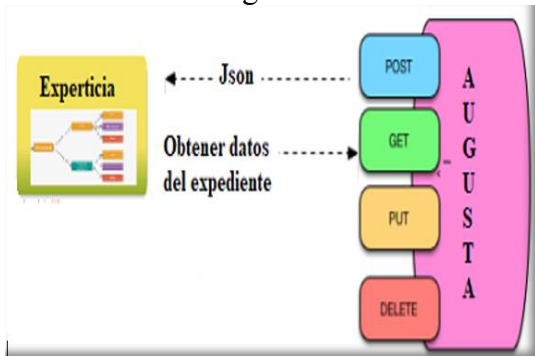


Figura 2. Estructura del SE Experticia Web.

La creación de los objetos JSON implica escribir datos, para ello:

- Los datos están separados por comas.
- Los datos se escriben en pares, siendo primero el nombre o atributo del mismo y luego el valor del dato.
- Los objetos JSON están rodeados por llaves “{}”.
- Llaves cuadradas “[]” guardan arreglos, incluyendo otros objetos.

A continuación, en la Figura 3, se muestra un fragmento de los datos en formato Json usados en Experticia:

```

1  {
2  ..... "idExpediente": 112298,
3  ..... "idOrganismo": 1862,
4  ..... "prefijo": "LC",
5  ..... "numero": 9609,
6  ..... "sufijo": "1",
7  ..... "letraReceptoria": null,
8  ..... "numeroReceptoria": null,
9  ..... "anioReceptoria": null,
10 ..... "caratula": "DE ARMAS BAQUERO, EDISON ALEJANDRO S/ 3
11 ..... "fechaInicio": "2021-03-18T18:32:14.93",
12 ..... "fechaRadicaion": "2021-07-01T00:00:00"
13 }

```

Figura 3. Estructura de datos con Json.

Una ventaja que tiene JSON respecto a XML es que el código resultante es más liviano. Para guardar la misma información utilizando JSON reduce el tamaño ya que no produce redundancia de datos y esto repercute en una mayor velocidad a la hora de transmitir la información [13].

En la nueva versión, luego de seleccionar la causa, Experticia se comunica con Augusta

para solicitar las últimas actualizaciones de la misma. Se lleva a cabo el procesamiento que ahora ya se almacena en Augusta a través de la API desarrollada.

2. LÍNEAS de INVESTIGACIÓN y DESARROLLO

Entre las líneas de investigación a considerar en este proyecto se pueden mencionar:

- El modelado de datos JSON.
- La arquitectura REST.
- El Patrón de diseño REACT (también llamada React.js o ReactJS)

3. RESULTADOS OBTENIDOS/ESPERADOS

Se presentó este trabajo en el año 2022, en el XXVIII Congreso Argentino de Ciencias de la Computación (CACIC), que se llevó a cabo en la Provincia de La Rioja, con el título “*Desarrollo de Interfaces de Programación de Aplicaciones aplicadas en un Sistema Experto Jurídico*”.

Actualmente, esta nueva versión de Experticia se haya en etapa de pruebas. En la Tabla 1 se muestra los tiempos insumidos por el sistema para la consulta inicial de una causa, cuyos detalles se desglosan en la misma tabla.

Tabla 1. Detalle de las acciones junto con sus respectivos tiempos (Elaboración propias).

Proceso	Tiempo
Pedirle a Augusta información de una causa	1202 ms
Traer de Augusta los datos de los involucrados en la causa	143 ms
Actualizar los datos de la causa en Experticia a tratar y enviar tal información al operador	15 ms
Tiempo total insumido para agregar la nueva causa a Experticia con las verificaciones pertinentes	1360 ms

Estas son pruebas muy preliminares los tiempos surgen del promedio de 20 causas consultadas. Como se puede observar el

mayor tiempo insumido corresponde a la petición de la causa, que ha demandado aproximadamente un 88% del tiempo total, esto a primera vista podría parecer excesivo. Sin embargo, cuando se tiene en cuenta que no se trata de la petición de un dato aislado, sino que en realidad está trayendo el histórico de una causa, podría considerarse un tiempo totalmente aceptable.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación la lleva adelante un equipo de 15 integrantes provenientes de dos departamentos de la UNLaM, el DIIT y el Departamento de Derecho y Ciencia Política, además se cuenta con 2 asesores especialistas externos, del Poder Judicial de la Pcia. de Bs. As. y un secretario de juzgado.

Un alumno de grado parte de este equipo, se graduó en la carrera de Ingeniería de Informática, recibiendo su título en 2022.

Dos integrantes han realizado el “Curso Superior en Derecho. Inteligencia Artificial y Derecho”. Organizado por la Fundación General de la Universidad de Salamanca (España), y avalado por la Comunidad Europea.

Por otra parte, la línea de investigación presentada es parte directa de la tesis doctoral titulada “Implementación de la Inteligencia Artificial y su Regulación en los Procesos de Gestión en la Ejecución Penal en la Provincia de Buenos Aires”, que está siendo desarrollada por la Dra. Laura Conti en UNLaM,

BIBLIOGRAFÍA

- [1] Hidalgo Luis Amador. (1996) Inteligencia Artificial y Sistemas Expertos. Universidad de Córdoba. ISBN: 8478013466
- [2] Badaró, S., Ibañez, L., & Agüero, M. (2013). Sistemas Expertos: Fundamentos, Metodologías y Aplicaciones. Ciencia y Tecnología, 13, pp. 349-364. Universidad de Palermo, Buenos Aires, Argentina. https://www.palermo.edu/ingenieria/pdf2014/13/CyT_13_24.pdf
- [3] Spositto, O. y otros. (2020). Experticia, un sistema experto para dar apoyo al despacho de trámites asociados al expediente judicial. XIV Simposio de Informática en el Estado (SIE 2020) - JAIIO 49 (Modalidad virtual) ISSN: 2451-7534, pp. 17-29. oai:sedici.unlp.edu.ar:10915/121991. <http://sedici.unlp.edu.ar/handle/10915/12199>
- [4] Spositto, O. y otros. (2020). Inteligencia Artificial aplicada al Poder Judicial, XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020), Universidad Nacional de la Patagonia Austral (UNPA), pp. 7-11, ISBN: 978-987-3714-82-5.
- [5] Spositto, O. y Otros. (2020) Metodología para evaluar un modelo de Justicia Predictiva. 8vo. Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2020). UTN- Facultad Regional San Francisco. ISBN 978-950-42-0202-8, pp. 527-535.
- [6] Spositto, O. y Otros. (2021). Experticia. Un modelo de sistema experto aplicado al Poder Judicial. XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja). oai:sedici.unlp.edu.ar:10915/119936. ISBN: 978-987-24611-3-3; 978-987-24611-4-0. Páginas:113-118. <http://sedici.unlp.edu.ar/handle/10915/119936>
- [7] Servicios web. IBM. (2022). <https://www.ibm.com/docs/es/was/9.0.5?topic=services-web>. Última actualización: 2022-12-13.
- [8] Beltran, C, (2019). Diferencia entre API y Servicio Web. Disponible en: <https://medium.com/beltranc/diferencia-entre-api-y-servicio-web-5f204af3aedb>
- [9] Amodeo, E. (2013) Principios de diseño de APIs REST (desmitificando REST). Disponible en: <https://qdoc.tips/introduccionapisrestpdf-pdf-free.html>
- [10] Diseño de API RESTful. (2021). Disponible en: <https://www.ibm.com/docs/es/zos-connect/zosconnect/3.0?topic=apis-designing-restful>
- [11] Breje, A. R., y otros. (2018). Comparative study of data sending methods for xml and json models. International Journal of Advanced Computer Science and Applications, 9(12), 198-204.
- [12] D. Florescu and G. Fourny, "JSONiq: The History of a Query Language," in IEEE Internet Computing, vol. 17, no. 5, pp. 86-90, Sept.-Oct. 2013, doi: 10.1109/MIC.2013.97.
- [13] Paiva, R. (2021). Cómo transferir archivos a través de REST para almacenar en una propiedad. Parte I. Disponible en: <https://es.community.intersystems.com/post/cómo-transferir-archivos-través-de-rest-para-almacenar-en-una-propiedad-parte-1>

ORGANIZACIONES INTELIGENTES ASISTIDAS POR TECNOLOGÍAS 4.0

Romagnano, M.^{1,2,3}, Ganga, L.^{1,3}, Pantano, J. C.^{1,2}, Herrera, M.^{1,3}, Lépez, H.¹, Becerra, M.¹, Aballay, A.¹, Sarmiento, A.¹, Jason, S.², Furlani, F.², García, A.¹

¹Departamento de Informática, FCEFN, Universidad Nacional de San Juan

²Instituto de Informática, FCEFN, Universidad Nacional de San Juan

³Departamento de Matemática, Física y Química, FCEFN, Universidad Nacional de San Juan

{maritaroma,mherrera}@iinfo.unsj.edu.ar, juancruz871@hotmail.com, {leonelganga,mcbecerra2008,prof.alicia,sofijason3108,fedefurlani14,garcia.exe}@gmail.com, {adriva2005,lepezhr}@yahoo.com.ar

RESUMEN

A medida que avanza el siglo XXI, las organizaciones, cualquiera sea su tipología, se enfrentan a nuevos y diversos desafíos en la manipulación de sus datos. Deben almacenar y gestionar grandes cantidades de datos que se generan de sus transacciones diarias, adecuarse rápidamente al medio, tomar decisiones de forma casi inmediata, etc. Por lo tanto, tienen que dedicar una importante cantidad de recursos y de tiempo laboral para obtener datos de valor, los que luego serán procesados y convertidos en información oportuna. Además, la conversión de datos en información y de información en conocimiento, requiere la sincronía de muchos especialistas y de especialidades dentro de la organización. Este conocimiento adquirido, y su posterior gestión para la toma de decisiones, puede ser la principal ventaja con la que cuenta una organización. Diversas herramientas, técnicas y metodologías, actualmente conocidas como Tecnologías 4.0, ofrecen ventajas competitivas en cuanto a la trazabilidad y análisis de los datos. Permiten conocer el estado actual y pronosticar el futuro comportamiento de las organizaciones y del entorno, y realizar acciones proactivas basadas en el análisis predictivo y prescriptivo.

Sin embargo, en Argentina, la mayor parte de las organizaciones aún se encuentran operando con tecnologías anticuadas. Muchas no tienen en claro los beneficios que pueden percibir por

adoptar Tecnologías 4.0, por lo que todavía se encuentran reacias a asumir el tiempo y los costos de la adaptación.

Por lo tanto, atendiendo a estas necesidades de gestión de la información por la cual transitan actualmente las organizaciones, el presente trabajo propone identificar un adecuado conjunto de Tecnologías 4.0 a implementar en organizaciones regionales, para asistirles en la toma de certeras y oportunas decisiones; convirtiéndolas así en organizaciones de vanguardia, denominadas “organizaciones inteligentes”.

Palabras Claves: Analítica de Negocios, Inteligencia de Negocios, Tecnologías 4.0, Toma de Decisiones, Organización Inteligente.

CONTEXTO

El presente trabajo se encuentra enmarcado en el proyecto “Tecnologías 4.0 para Asistir a la Toma de Decisiones en Organizaciones Regionales”, presentado en la convocatoria del Consejo de Investigaciones Científicas y Técnicas y de Creación Artística (CICITCA) de la Universidad Nacional de San Juan, para ser desarrollado durante el período comprendido entre 01/01/2023 al 31/12/2024.

Las tareas de investigación se desarrollan en el Laboratorio de Sistemas de Información, en el ámbito del Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.

El grupo de investigación se encuentra conformado por profesionales y estudiantes avanzados de grado y postgrado, de distintas disciplinas, tales como: Informática, Bioingeniería, Estadística, Administración de Empresas y Abogacía. En su mayoría, cuentan con una experiencia de más de 20 años en la línea de investigación que han adoptado, trabajando en temáticas relacionadas con inteligencia de negocios, analítica de negocios, análisis y preservación de datos, sistemas de información y recuperación de información integrando técnicas de aprendizaje automático de la inteligencia artificial. Además, a partir del 2021 se comenzó a incursionar en economía de plataformas, modelos de negocios, fintech, responsabilidad social y revolución 4.0; en distintos dominios de aplicación.

1. INTRODUCCIÓN

En los tiempos que corren, la necesidad de las organizaciones por lograr una mejora continua en sus procesos de negocio, juega un rol protagónico y decisivo. A su vez, aquellas organizaciones categorizadas como empresas tienen que enfrentarse a nuevos modelos económicos, obligadas a realizar frecuentes cambios, tratando de acercarse todo lo posible a sus consumidores. Deben dedicar gran parte de su tiempo y de sus recursos, económicos y humanos, a la obtención, procesamiento, aplicación y proyección de la información.

La idea de Organización Inteligente es relativamente novedosa, sin embargo, ya en 1990 Peter Senge introdujo el concepto al sostener que toda agrupación tiene la posibilidad (y debería) aprender de sus propias experiencias, recopilando y analizando datos, para poder generar lo que se conoce como mejora continua. Entonces, las organizaciones inteligentes se definen como aquellas capaces de aprender permitiendo así expandir sus posibilidades de crecimiento. Estas organizaciones se diferencian de las convencionales en cuanto a que: tienen una visión integral, se apoyan en tecnologías, se basan en datos, valoran el feedback, y retienen

los talentos. El factor determinante en una organización de este tipo es la tecnología. La capacidad de digitalizar o directamente automatizar procesos es condición primordial. Además, deben contar con una infraestructura que facilite la recolección, recopilación y análisis de datos macro. Big Data, la Inteligencia Artificial y el Machine Learning son los principales mecanismos que entran en acción en estas organizaciones (León M., Tejada G., Yataco T., 2003). Por lo tanto, estas organizaciones trabajan bajo la concepción de Paradigma 4.0 al introducir una nueva forma de producir con tecnologías de última generación, Tecnologías 4.0, para automatizar sus procesos productivos y transformar los modelos de negocio, optimizando los recursos (Ministerio de Desarrollo Productivo, 2021).

Estas tecnologías emergen a partir de la Revolución Digital, donde algunas ciencias y disciplinas como la Ciencia de los Datos (DS), Inteligencia Artificial (IA), Inteligencia de Negocios (BI), Minería de Datos (DM), Big Data (BD), Analítica de Negocios (BA), Ingeniería del Conocimiento (KE) y Sistemas de Información (IS) han fusionado sus esfuerzos y saberes.

Dentro del paquete de Tecnologías 4.0 se identifica la categoría de tecnologías destinadas a los sistemas de gestión organizacional (Ministerio de Desarrollo Productivo, 2021). Así, la Inteligencia de Negocios, o simplemente BI, considerada como una especificidad de la Ciencia de los Datos, puntualmente como una especificación de la Inteligencia Artificial para el Negocio, viene a ofrecer esta gestión del conocimiento. Se presenta como un conjunto de procesos, aplicaciones y tecnologías que facilitan la obtención rápida y sencilla de datos en grandes bases de datos (Big Data), provenientes de distintos Sistemas de Información Organizacionales. Por lo tanto, los datos analizados e interpretados, se transforman en conocimiento apropiado para la toma de decisiones. Las posibilidades de procesamiento que ofrece la Inteligencia de Negocios hacen que las organizaciones puedan desempeñarse y responder dinámicamente en

un entorno universal altamente impetuoso, el cual exige una constante toma de decisiones para mantenerse en el medio. Como complemento a la Inteligencia de Negocios, las organizaciones deben tener en cuenta a la Analítica de Negocios, considerada como el conjunto de técnicas y procesos tales como análisis cuantitativo, pronósticos, análisis predictivos, optimización, entre otros, que mantienen y sustentan la performance de negocios, explorando grandes volúmenes de datos para asistir a los ejecutivos en la toma de decisiones (Orrillo, 2018).

Dir&Ge (2020), pronosticaron que en el 2021 el mercado de BI movería alrededor de los 23.100 millones de dólares y crecería a ritmos anuales del 7,6% entre 2020 y 2025.

No obstante, una de las principales restricciones para su implementación es la disponibilidad de personal capacitado. Existe amplio consenso, tanto desde el sector público como privado, respecto de la necesidad de promover la formación en habilidades en este nuevo paradigma 4.0.

La tarea de sensibilización es necesaria, y más en un contexto en el que no abundan los ejemplos exitosos que impulsen al resto; siendo el primer paso para encarar los desafíos tecnológicos.

Por lo tanto, a través de este estudio, se propone impartir conocimiento y proveer de un conjunto de técnicas, metodologías y software libre y de código abierto que ayude a las entidades regionales en su transición tecnológica y cultural, tendiendo a convertirlas en organizaciones inteligentes, que tomen certeras y oportunas decisiones.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Para enfrentar el mercado fluctuante y dinámico, las organizaciones deben actuar con prudencia al optimizar sus recursos y al tomar decisiones estratégicas. El conocimiento logrado, y su posterior gestión puede ser su principal ventaja competitiva. Heavin, Daly y Adam planteaban que la habilidad de una

organización para administrar el conocimiento es esencial en términos de su desarrollo y como un activo estratégico (Heavin, Daly y Adam, 2014).

Los autores de este artículo e integrantes del proyecto se han dividido en subgrupos, de acuerdo a su formación y especialización, y se encuentran trabajando en las siguientes líneas de investigación y desarrollo:

- Revolución 4.0 y Tecnologías emergentes.
- Inteligencia y analítica del negocio.
- Análisis y procesamiento estadístico de los datos.
- Preservación y legislación de los datos.
- Sistemas de información organizacionales.

Consecuentemente, el equipo de investigación, considera la temática de relevancia social y organizacional. Posee implicancias prácticas, es viable y propicia puesto que, en abril del 2021 se lanzó el Plan de Desarrollo Productivo Argentina 4.0. Por medio de este plan se establecen estrategias para impulsar el paradigma 4.0 y promover el desarrollo de soluciones tecnológicas 4.0 (Ministerio de Desarrollo Productivo, 2021). También, se suma el plan de transformación digital (Ministerio de Obras Públicas, 2022). Además, desde el año 2020, Argentina cuenta con políticas de estado tales como Argentina Innovadora 2020 y 2030, Ley de Promoción de la Economía del Conocimiento, Industria 4.0, Plan de Inteligencia Artificial, entre otras (Ministerio de Ciencia, Tecnología e Innovación, 2020; Ministerio de Desarrollo Productivo, 2021; ARGENIA, 2020; Ministerio de Economía, 2020). Precisamente, San Juan, la provincia donde residen los investigadores, posee polos tecnológicos como Casetic San Juan, Servicios Mineros y San Juan TEC, lo cual puede tomarse como puntapié inicial para trabajar, conjuntamente con el Estado y/o con organizaciones privadas, en la sustentabilidad regional, aportando conocimiento y casos de estudio desde la academia. Igualmente, Argentina puede aspirar a posicionarse como un líder regional en la generación de soluciones específicas

basadas en tecnologías 4.0, que puedan aplicarse en cadenas como la agrícola, la alimenticia y la petroquímica, entre otros (eLAC, 2022).

3. RESULTADOS ESPERADOS

Se espera definir un adecuado conjunto de Tecnologías 4.0 a implementar en organizaciones regionales, para asistirles en la toma de certeras y oportunas decisiones.

Para lograr este propósito se ha propuesto:

- Especificar las distintas técnicas, metodologías, software libre y de código abierto existentes, que puedan ayudar a las organizaciones regionales en su migración, tecnológica y cultural, hacia entidades inteligentes.
- Describir la integración de distintas Tecnologías 4.0 que ayudarán a las organizaciones en su proceso de toma de decisiones.
- Identificar los beneficios de incorporar Tecnologías 4.0 a las organizaciones regionales.

Los resultados del proyecto tienen una inmediata transferencia al medio regional, principalmente a la alta gerencia de organizaciones regionales de cualquier tipo, y a la comunidad científica.

Al mismo tiempo, se trabajará en experimentación y difusión de resultados con investigadores de la Facultad de Ingeniería de la Universidad de Quindío, Colombia, y de la Escuela de Ingeniería en Transporte, Pontificia Universidad Católica de Valparaíso, Chile.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del grupo de investigación se identifican integrantes formándose en:

- Doctorado en Ingeniería, Facultad de Ingeniería, UNCuyo.

- Doctorado en Demografía, Escuela de graduados de la Facultad de Ciencias Económicas, UNC.
- Maestría en Estadística Aplicada, Escuela de graduados de la Facultad de Ciencias Económicas, UNC.
- Licenciatura en Ciencias de la Computación, Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.
- Licenciatura en Sistemas de Información, Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.
- Becarios CICITCA, categoría estudiante avanzado.
- Becaria FunDatos.

5. BIBLIOGRAFÍA

ARGENIA (2020). Plan Nacional de Inteligencia Artificial, <https://ia-latam.com/wp-content/uploads/2020/09/Plan-Nacional-de-Inteligencia-Artificial.pdf>, recuperado septiembre, 2022.

Dir&Ge (2020). El mercado de business intelligence crecerá a ritmos anuales del 7,6% entre 2020 y 2025, <https://directivosygerentes.es/innovacion/mercado-business-intelligence-crecimiento>, recuperado enero de 2021.

eLAC (2022). “Tecnologías Digitales para un Nuevo Futuro”, https://repositorio.cepal.org/bitstream/handle/11362/46816/1/S2000961_es.pdf, recuperado febrero de 2023.

Heavin, C., Daly, M., y Adam, F. (2014). Small data to Big Data: The Information Systems (IS) continuum. KMIS 2014 - Proceedings of the International Conference on Knowledge Management and Information Sharing. 289-297.

León M., Tejada G., Yataco T. (2003). Las Organizaciones Inteligentes. Notas Científicas, Vol. (6) 2: pp. 82-87, https://sisbib.unmsm.edu.pe/bibvirtualdata/publicacion/es/indata/Vol6_n2/pdf/organizaciones.pdf, recuperado diciembre, 2022.

Ministerio de Ciencia, Tecnología e Innovación (2020). Argentina Innovadora 2020. Principales Ejes de Políticas Públicas en CTI, <https://www.argentina.gob.ar/ciencia/argentina->

innovadora-2030/plan-argentina-innovadora-2020, recuperado septiembre, 2022.

Ministerio de Desarrollo Productivo (2021). Plan de Desarrollo Productivo. Argentina 4.0. https://www.argentina.gob.ar/sites/default/files/plan_de_desarrollo_productivo_argentina_4.0.vf__1.pdf, recuperado enero, 2023.

Ministerio de Economía (2020). Régimen de Promoción de la Economía del Conocimiento, <https://www.argentina.gob.ar/economia/igualdadygenero/regimen-de-promocion-de-la-economia-del-conocimiento#:~:text=7%2F10%2F2020>, recuperado septiembre, 2022.

Ministerio de Obras Públicas (2022). Plan Estratégico de Transformación Digital para la Gestión de la Obra Pública, <https://www.argentina.gob.ar/obras-publicas/secretaria-gestion/transformacion-digital#:~:text=El%20Plan%20Estrat%C3%A9gico%20de%20Transformaci%C3%B3n,eficiente%20la%20Obra%20P%C3%ABlica%20Nacional>, recuperado octubre, 2022.

Orrillo, J. (2018). Analítica de negocios: Una estrategia basada en la Información, <https://www.esan.edu.pe/conexion/actualidad/2018/03/22/analitica-negocios-informacion/>, recuperado julio, 2022.

Aprendizaje Automático para Nodos Inteligentes

Marcos D. Aranda^{1,2}, Eduardo Piray¹, Paola Beltramini^{1,2}, Elian Gonzalez¹, Gustavo Futo¹
marcos.aranda@ues21.edu.ar

(1) Universidad Siglo 21. Decanato de Ciencias Aplicadas.

(2) Universidad Nacional de Catamarca. Facultad de Tecnología y Ciencias Aplicadas.

AREA TEMATICA: Agentes y Sistemas Inteligentes

RESUMEN

El presente trabajo presenta con el Big Data y el aprendizaje automático puede ser utilizado en el ámbito de Ganadería de Precisión con el fin de mejorar la calidad en la producción animal en la República Argentina. Para ello se espera analizar e identificar algunos algoritmos y herramientas de código abierto utilizados para el aprendizaje automático que nos permitan indagar cuales son las más utilizadas y de mejor rendimiento en el campo de la inteligencia artificial.

Palabras Claves: Inteligencia Artificial, Aprendizaje Automático, IoT.

CONTEXTO

Nuestra propuesta de investigación surge como línea de trabajo debido a la convergencia de avances en algunos campos de conocimiento en particular en el Big Data y la inteligencia artificial, junto con el desarrollo de modernos dispositivos electrónicos que, a partir de sensores, aplican herbicidas únicamente en lugares donde se identifiquen malezas, en lugar de rociar todo el campo. La puesta a disposición de nuevas fuentes de información y el enorme incremento en la capacidad de procesamiento y almacenamiento de datos posibilitado por las Tecnologías de la Información y las Comunicaciones (TICS) dio lugar al surgimiento de un nuevo paradigma tecnoproductivo aplicado a las producciones de base biológica: la llamada Agricultura de Precisión (AP) y Ganadería de Precisión (GP) [1]. Este proyecto inicio en el año 2022 y se enfoca en profundizar los conocimientos sobre métodos actuales de aprendizaje automático como herramienta para el descubrimiento de

conocimiento en problemas de Big Data aplicado a la Ganadería de Precisión.

El principal organismo financiero y beneficiario resulta ser la Universidad Siglo 21 siendo la carrera de grado de la Licenciatura en Informática perteneciente al Área de Ciencias Aplicadas los espacios donde se realizan los estudios.

1. INTRODUCCIÓN

Los dispositivos que cuentan con inteligencia artificial pueden ejecutar distintos procesos análogos al comportamiento humano, como la devolución de una respuesta por cada entrada (similar a los reflejos de los seres vivos), la búsqueda de un estado entre todos los posibles según una acción o la resolución de problema mediante una lógica formal [1]. Machine Learning, es una rama de la Inteligencia Artificial que se encarga de generar algoritmos que tienen la capacidad de aprender y no tener que programarlos de manera explícita. Hay dos tipos de aprendizajes: el supervisado es cuando entrenamos un algoritmo de Machine Learning dándole las preguntas (características) y las respuestas (etiquetas). Así en un futuro el algoritmo pueda hacer una predicción conociendo las características [2]. Además, en los últimos años, ha tomado mucho impulso una rama del Machine Learning conocida como Aprendizaje Profundo o Deep Learning, la cual combinada con técnicas de Visión Artificial permiten clasificar objetos en imágenes o videos, detectar su posición, seguirlos en una escena, contar la cantidad o detectar fallas o defectos en los mismos.

La AP y GP se basan en la utilización de equipos, agropartes y servicios orientados a optimizar los procesos productivos de base biológica. Estas tecnologías buscan adecuar las

prácticas productivas a cada “microambiente”, definido para el caso del agro por ciertas condiciones ambientales, edafológicas –por ejemplo, a partir de la identificación de la composición de nutrientes del suelo– geográficas y climáticas presentes en todo entorno natural [3]. A su vez, la aplicación de tecnologías de AP y GP posibilita un control permanente de los cultivos y/o de los animales. A partir de dispositivos específicos, como drones o imágenes satelitales en agro o instrumentos ligados a tecnologías de Internet of Things (IoT) para las producciones pecuarias, el usuario de estas tecnologías puede mantener un monitoreo completo y sistemático de su producción [3]. Esto le permite desde la detección temprana de malezas, plagas o enfermedades hasta la identificación de la evolución de los rendimientos.

En la República Argentina se han incorporado en la agricultura muchas tecnologías de precisión en los últimos años. El ejemplo más emblemático es el mapa de rendimiento del lote, que permite luego aplicar fertilización y densidad de siembra variables al año siguiente, lo que significa hacer un manejo diferencial por ambientes [4]. Herramientas como estas y otras han permitido aumentar ininterrumpidamente los rendimientos de producción agrícolas. De manera paralela, la ganadería ha avanzado en el proceso de intensificación, pero no en igual medida que la agricultura en nuestro país.

La GP puede ser aplicada siempre y cuando se construya en cada empresa un Proyecto Ganadero, a partir del cual se desarrollen todas las actividades productivas e inversiones. Un productor ganadero puede elaborar “SU” propio Proyecto, que puede ser muy diferente, aunque esté en la misma zona (clima y suelo) y tenga la misma actividad ganadera (cría, engorde o ciclo completo) [4]. La clave de una correcta Ganadería de Precisión es que cada productor tenga “SU” propio Proyecto que consta de una serie de componentes:

- Objetivos y metas a corto y mediano plazo para buscar altas ganancias de peso con la mayor carga animal posible.
- Planificación y manejo de recursos forrajeros, concentrados, para ganar peso por estación climática.

- Controles de monitoreo con báscula.

Este conjunto de componentes para el monitoreo y control en distintas etapas del animal (cría, engorde o ciclo completo) requieren disponer de información en tiempo real, ya procesada y analizada, que permita acceder a un panorama preciso de las condiciones actuales del ganado, definir las prioridades del día y así optimizar su manejo del tiempo, contar con información detallada sobre los movimientos de los animales, alimentación y consumo de agua, niveles inusuales de actividad o aumento de los niveles normales de tos, etc [5]. El aprendizaje automático permite analizar información compleja de todas esas fuentes para lograr detectar tendencias, signos o información precisa para apoyar las decisiones diarias en la GP.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Este proyecto tiene como principal objetivo profundizar el conocimiento sobre métodos actuales de aprendizaje automático como herramienta para el descubrimiento de conocimiento en problemas de Big Data aplicado a la Ganadería de Precisión.

Como objetivos específicos este proyecto se plantea:

- Identificar factores de comportamiento y/o posturas que se relacionan con patologías conocidas o situaciones que por su rareza requieren ser atendidas, considerando diferentes razas y especies animales.

- Analizar e identificar algoritmos y herramientas de aprendizaje automático de código abierto que puedan ser utilizados en la Ganadería de Precisión.

Además de analizar e identificar herramientas utilizadas para aprendizaje automáticos, se intentará utilizar un enfoque de código abierto, utilizando librerías como TensorFlow el cual es útil para Machine Learning a gran escala [6]. Es una mezcla de aprendizaje automático y modelos de redes neuronales que utiliza como lenguaje intérprete Python [7]. La característica más destacada de TensorFlow es que se ejecuta tanto en CPU (Unidad Central de

Procesamiento) como GPU (Unidad de Procesamiento Gráfica) [8].

3. RESULTADOS OBTENIDOS / ESPERADOS

En el presente proyecto buscará realizar una contribución sobre aspectos que aún no se encuentren resueltos en la Ganadería de Precisión, utilizando aprendizaje automático para ofrecer identificación inteligente y trazabilidad de los animales: percepción y detección inteligentes para registrar información. Se plantea profundizar el conocimiento en algoritmos basados en aprendizaje automático que permitan mejorar la calidad del desarrollo ganadero de precisión a través del monitoreo y control.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por 6 integrantes, 3 docentes de la carrera Licenciatura en Informática pertenecientes al Área de Ciencias Aplicadas de la Universidad Siglo 21, 1 docentes de la carrera de Ingeniería Electrónica de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca y además en esta oportunidad se incorporaron 2 alumnos pertenecientes a la Universidad Siglo 21, con el fin de poder iniciarlos en su formación de investigadores.

Este proyecto de investigación posee una diversidad notable, ya que hay docentes de diferentes asignaturas de las mencionadas carreras, en las áreas de arquitectura de computadoras, sistemas operativos, sistemas operativos avanzados y física, por lo que se cuenta con un conocimiento en el área de investigación.

En cuanto a la formación de los integrantes, hay un docente realizando su Tesis de Posgrado el cual es un Doctorando en Ingeniería en formación en temas de desarrollo tecnológico con microprocesadores de última generación e inteligencia artificial y otro en área de educación

en formación que poseen experiencia de divulgación.

5. BIBLIOGRAFIA

- [1] Julián Pérez Porto y Ana Gardey. «Definición de inteligencia artificial - Qué es, Significado y Concepto». [En línea]. Disponible en: <https://definicion.de/inteligencia-artificial/>. [Accedido: 28 - marzo - 2022]
- [2] «Machine Learning Made Faster, More Efficient». [En línea]. Disponible en: <https://www.arm.com/solutions/artificial-intelligence/machine-learning/> [Accedido: 28 - marzo - 2022]
- [3] Jeremías Lachman; Andrés López. «Innovación, habilidades y nuevas áreas de conocimiento en sectores tecnológicos emergentes: el caso de la Agricultura y Ganadería de Precisión». Revista Pymes, Innovación y Desarrollo 2018. Vol. 6, No. 3, pp. 60-85.
- [4] Dr. (Ing. Agr. M.Sc. PhD) Aníbal Fernández Mayer. «¿Ganadería de precisión, es posible?» Nutricionista de INTA Bordenave (Buenos Aires) Centro Regional Buenos Aires Sur (CERBAS). https://www.vetcomunicaciones.com.ar/uploads/archivos/ganader__a_de_precisi__n__es_posible..pdf [Accedido: 22 - marzo - 2022]
- [5] «Inteligencia artificial: el futuro de las granjas». [En línea]. Disponible en: <https://www.sudamerica.boehringer-ingelheim.com/salud-animal/animales-de-produccion/inteligencia-artificial-el-futuro-de-las-granjasehringer-ingelheim.com> [Accedido: 22 - marzo - 2022]
- [6] «Plataforma de extremo a extremo de código abierto para el aprendizaje automático». [En línea]. Disponible en: <https://www.tensorflow.org/> [Accedido: 28 - marzo - 2022]
- [7] Michael Bowles. Published by John Wiley & Sons, Inc. «Machine Learning in Python® : Essential Techniques for Predictive Analysis». Published by John Wiley & Sons, Inc. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781brary>. [Accedido: 28 - marzo - 2022]
- [8] Jason Sanders, Edward Kandrot. «An introduction to general-purpose GPU programming». Published by Addison-Wesley. 2011

DEPURACIÓN DE BASES DE DATOS DE SEGUNDA SECCIÓN DEL BOLETÍN OFICIAL DE LA REPÚBLICA ARGENTINA MEDIANTE APRENDIZAJE DE MAQUINA

Nestor A. Balich, Franco A. Balich, Hugo Fraga

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana. Informática (UAI)
Montes de Oca 745. Ciudad Autónoma de Buenos Aires, Argentina.

{nestor.balich, francoadrian.balich}@uai.edu.ar

hfraga@boletinoficial.gob.ar

RESUMEN

El Boletín Oficial de la República Argentina es una fuente importante de información para empresas, instituciones y particulares, pero la gran cantidad de información publicada diariamente hace que la depuración de bases de datos sea un proceso complejo y costoso. Se hipotetiza que el uso de un modelo de inteligencia artificial basado en aprendizaje de maquina (ML) puede aprender a clasificar y catalogar los avisos comerciales de manera eficiente y autónoma. La aplicación de técnicas de ML en la clasificación de textos ha evolucionado significativamente, especialmente en el uso de modelos de aprendizaje profundo en áreas como la identificación de noticias falsas y la detección de spam. También se han realizado estudios sobre el uso de técnicas de ML en la identificación y corrección de errores en bases de datos, incluyendo la corrección de errores en imágenes médicas y la identificación de avisos comerciales ilegales en la web. Se propone el desarrollo de un modelo de aprendizaje de máquina para la clasificación y catalogación de los avisos comerciales de la segunda sección del Boletín Oficial de la República

Argentina, lo que podría representar un avance significativo en la depuración de bases de datos en este ámbito.

Palabras clave: Depuración, Base de datos, Aprendizaje de Maquina, Inteligencia Artificial, Aprendizaje Supervisado, Clasificación

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Centro de Altos Estudios en Tecnología informática (CAETI) de la Universidad Abierta Interamericana (UAI). En este proyecto participan docentes, alumnos e investigadores enmarcado en los proyectos de transferencia tecnología del laboratorio de robótica física (LRF). Las líneas de investigación sobre inteligencia dentro de los proyectos con financiamiento y duración a 2 años.

1. INTRODUCCIÓN

La segunda sección del Boletín Oficial de la República Argentina es un importante medio de difusión de información para empresas, instituciones

y particulares. Sin embargo, la cantidad de información que se publica diariamente en esta sección hace que la depuración de bases de datos sea un proceso complejo y costoso en términos de tiempo y recursos. En este contexto, surge la hipótesis de que es posible desarrollar un modelo de inteligencia artificial basado en ML capaz de aprender a clasificar los avisos comerciales y luego catalogarlos para obtener los avisos de manera eficiente y de forma totalmente autónoma.

La aplicación de estas técnicas en la clasificación de textos es un área de investigación en constante evolución, y su aplicación en la depuración de bases de datos no es una excepción. En particular, la utilización de modelos de aprendizaje profundo ha permitido mejorar significativamente la capacidad de clasificación de textos en diversas áreas, como la identificación de noticias falsas o la detección de spam en correos electrónicos.

En el ámbito de la depuración de bases de datos, también se han realizado estudios sobre la utilización de técnicas de ML para la identificación y corrección de errores *“el uso de Machine Learning para mejorar la eficiencia y la precisión de la limpieza de datos y la consideración de los efectos de la limpieza de datos en análisis estadístico”* Chu [1]

Los problemas de calidad de los datos son complejos de resolver y existente gran cantidad de técnicas y herramientas para lograrlo Gudivada [2] y que debimos evaluar al seleccionar los modelos a utilizar en el presente trabajo.

Otro estudio interesante donde se desarrolló un modelo para la detección

de publicidad ilegal en vallas publicitarias basado en el aprendizaje automático (FIBAD) Huxiao[3]. Permitió demostrar que los resultados obtenidos por el modelo eran capaces de identificar avisos ilegales con una alta precisión y eficiencia.

En este trabajo, se propone el desarrollo y validación de un modelo de inteligencia artificial basado en ML con aprendizaje supervisado para la clasificación y catalogación de los avisos comerciales de la segunda sección del Boletín Oficial de la República Argentina. Se espera que este modelo permita obtener los avisos de manera eficiente y autónoma, lo cual representaría un avance significativo en la depuración de bases de datos en este ámbito. *“Los datos son uno de los desafíos perennes en el análisis de datos, y no hacerlo puede resultar en análisis inexactos y decisiones poco confiables”* Khayyat [4].

Luego de realizar la (1) adquisición y depuración de datos se realizó el (2) proceso de análisis exploratorio de datos (EDA), donde los datos se examinan con estadísticas descriptivas simples para determinar si existen problemas de datos, como podemos observar detalladamente en el libro *“Data Science in the Database: Using SQL for Data Preparation”* Badia [5]. Continuando con la identificación de los datos representativos obtenidos de cinco tablas (id del tipo de aviso, texto del aviso, texto del tema, id de registro, fecha de ingreso, fecha de publicación, número de boletín) generando un archivo avisos2da.csv con un total de 770.000 registros. En el siguiente paso consistió en la (3) normalización de los datos eliminando los registros con

repubicaciones figura 1, pues el mismo aviso debe ser republicado varios días y con mismo texto. Para fines del entrenamiento solo necesitamos una sola muestra, reduciéndose el dataset a 550.000 registros.

idAviso	idRubro	RubroNombre	TextoXHTML	TextoXHTML_norm
0	89292	2300 AVISOS COMERCIALES	<Aviso>/in <NumeroTramite>#4344178# </NumeroTramite> % <NumeroPagDesde>11</NumeroPagDesde> %/in<Titulo>CENTRAL S.R.L./in</Titulo> <Texto>Por Reunión de Socios de fecha 27/03/2012, se resolvió por unanimidad aceptar la renuncia del Sr. Marcelo Eduardo José Costa a su cargo de gerente titular y designar al Sr. Marcos Pérez como gerente titular y al Sr. Gonzalo Jorge Mateos como gerente suplente. Por lo tanto, la gerencia queda integrada de la siguiente manera: Gerente Titular: Marcos Pérez; Gerente Suplente: Gonzalo Jorge Mateos. Los gerentes constituyen domicilio especial en Supacha 1111, Piso 18, Capital Federal; Luciana Zucatzosta autorizada por Acta de Reunión de Socios de fecha 27/03/2012./inAbogada - Luciana V. Zucatzosta/in</Texto> <Empieza>Vence>: 19/06/2012 N° 67509/12 v. 19/06/2012/in</Empieza> Vence> % <NumeroPagHasta>11</NumeroPagHasta> % </NumeroTramite>#4344178#</NumeroTramite> </Aviso>	central s r por reunión socios fecha resolvió unanimidad aceptar renuncia sr marcelo eduardo jose costa cargo gerente titular designar sr marcos perez gerente titular sr gonzalo jorge mateos gerente suplente gerencia queda integrada siguiente manera gerente titular marcos perez gerente suplente gonzalo jorge mateos gerentes constituyen domicilio especial supacha piso capital federal luciana zucatzosta autorizada acta reunion socios fecha abogada luciana v zucatzosta

Figura 1: Registro base

Se procedió a la (4) agrupación por categorías de tipo de aviso figura 2, detectando algunos tipos de avisos con menos de 20 ocurrencias que fueron descartados por no ser suficientes para el entrenamiento. Fijando la cantidad de avisos representativos mínima en 450 para poder extraer una muestra aleatoria y balanceada de cada categoría. Pues *“Los conjuntos de datos desequilibrados afectan negativamente el rendimiento de los algoritmos de aprendizaje automático”*, Mooijman[6].

RubroNombre	
AVISOS COMERCIALES	180641
BALANCES	527
CITACIONES Y NOTIFICACIONES, CONCURSOS Y QUIEBRAS, OTROS	30172
CONSTITUCION SA	35408
CONSTITUCION SAS	4710
CONTRATO SRL	47270
CONVOCATORIAS	23162
ESTATUTO OTRAS SOCIEDADES	146
ESTATUTO SCA	6
INFORMACION Y CULTURA	481
MODIFICACIONES SRL	26441
PARTIDOS POLITICOS	4538
REFORMA OTRAS SOCIEDADES	452
REFORMA SA	35664
REFORMA SCA	641
REINATES COMERCIALES	1248
REINATES JUDICIALES	8183
SUCESIONES	103539
TRANSF. FONDO DE COMERCIO	1994

Figura 2. Agrupación por categoría

Se procedió a establecer una (5) base de referencia con modelo de aprendizaje automáticos, procesando el set de entrenamiento normalizado (7.000 avisos) con AutoNLP, que lo agrupo y clasificó por tipos societarios, dividiendo el set en dos grupos (Entrenamiento, 70% y Test, 30%).

(6) Creamos 7 modelos (Naïve Bayes NB, GS, Logistic regression LogR, linear regression LinR, random forest classifier RFC, K-Nearest Neighbor KNC, Support Vector Machines SVC) candidatos chowdhury[7] en su respectiva notebook jupyter. Cada uno de ellos fue entrenado con el mismo 70% del set de 7.000 avisos y luego validado contra el 30% obtenido los resultados expresados en la figura 3.

	Model	Accuracy	f1	precision	recall
0	NB	86.90	86.56	88.04	86.90
1	GS	93.77	93.70	93.88	93.77
2	LogR	91.73	91.67	91.93	91.73
3	LinR	93.68	93.62	93.74	93.68
4	RFC	91.45	91.33	91.65	91.45
5	KNC	84.76	84.61	85.42	84.76
6	SVC	93.68	93.62	93.74	93.68

Figura 3. Resultados de modelos de IA

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Los ejes principales del trabajo son:

- Implementar un modelo de IA que aprenda a identificar avisos de 3ra sección.
- Implementar un modelo que valide y permita depurar la base de datos existente.
- Evidencia la performance y vialidad del empleo de IA en el

proceso de depuración de base de datos clasificadas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Una vez probado en distintos modelos, se define el de mejor eficacia y se procede a aplicar a los 550.000 avisos disponibles en la 2da. Sección. Obteniendo una eficacia de entre el 84,76% y 93,77%. Se obtuvo un total de 30.000 avisos sobre los 550.000 catalogados para revisión, de los cuales encontramos que sobre una muestra de 300 avisos 200 estaban mal clasificados y 100 no tenían el formato correspondiente o faltaban datos sobre el contenido del aviso.

Concluimos que ante las sucesivas migraciones de la base de datos a lo largo 20 años, se ha cambiado el formato y estructura de los avisos, pese a esto al ser un formato legal los principales indicadores están presentes permitiendo al modelo de aprendizaje supervisado aprender y clasificar con alto grado de asertividad (superiores al 84%).

El proceso total de trabajo desde la adquisición de los datos, depuración y entrenamiento demanda 3 días de trabajo, contra un proceso manual que se estima en más de un año si realizara de forma manual por personal de publicaciones con una dotación de 10 empleados.

El resultado es alentador y ya con el set de datos a verificar el personal de publicaciones estima un mes de trabajo para verificar y corregir las inconsistencias.

Como proyecto a futuro se contempla

incorporar los modelos de IA a las aplicaciones existentes, tanto para validar el ingreso de los avisos, como para contar con una herramienta de validación histórica en tiempo real. Y aplicar los modelos de aprendizaje a 1ra sección de la base de datos Boletín Oficial que supera los 2.000.000 de avisos.

También una nueva línea de investigación que se desprende del presente trabajo, sobre el análisis mediante inteligencia artificial de la base de datos legales. Como lo expresan varios autores *“la clasificación automatizada de textos legales es un tema de investigación destacado en el campo legal. Sienta las bases para construir un sistema legal inteligente.”* Haihua [8]. Enmarcados dentro de los proyecto de investigación y avances de los sistemas con inteligencia artificial en el sistema legal Argentino y su aplicabilidad Dobratinich[9] y Dobratinich [10].

4. FORMACION DE RECURSOS HUMANOS

El equipo está formado por 1 docente Magister en tecnología informática graduado de la Universidad Abierta Interamericana (UAI), 1 alumnos de 5to año de ingeniería en sistemas (UAI) y un investigador del laboratorio de robótica física (LRF-UAI).

Este proyecto también cuenta con la participación de colaboradores externos del área de sistemas y publicaciones de boletín oficial.

5. BIBLIOGRAFIA

[1]Chu, Xu & Ilyas, Ihab & Krishnan, Sanjay & Wang, Jiannan. (2016). Data Cleaning: Overview and Emerging

Challenges. 2201-2206.
10.1145/2882903.2912574.

<https://dl.acm.org/doi/proceedings/10.1145/2882903>

[2] Gudivada, Venkat & Apon, Amy & Ding, Junhua. (2017). Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations. International Journal on Advances in Software. 10. 1-20.
https://www.researchgate.net/publication/318432363_Data_Quality_Considerations_for_Big_Data_and_Machine_Learning_Going_Beyond_Data_Cleaning_and_Transformations

[3] Huxiao Liu, Lianhai Wang, Weinan Zhang, Wei Wang (2019). An Illegal Billboard Advertisement Detection Framework Based on Machine Learning
<https://doi.org/10.1145/3358528.3358549>

[4] Khayyat, Zuhair & Ilyas, Ihab & Jindal, Alekh & Madden, Samuel & Ouzzani, Mourad & Papotti, Paolo & Quiané-Ruiz, Jorge-Arnulfo & Tang, Nan & Yin, Si. (2015). BigDancing: A System for Big Data Cleansing. 10.1145/2723372.2747646.
https://www.researchgate.net/publication/274256636_BigDancing_A_System_for_Big_Data_Cleansing

[5] Badia Antonio (2023). Data Science in the Database: Using SQL for Data Preparation . University of Louisville, USA - DOI: 10.4018/978-1-7998-9220-5.ch069

[6] Paul Mooijman, Cagatay Catal, Bedir Tekinerdogan, Arjen Lommen, Marco Blokland. The effects of data balancing approaches: A case study, Applied Soft Computing,

<https://doi.org/10.1016/j.asoc.2022.109853>.

[7] Shovan Chowdhury, Marco P. Schoen (2020), Research Paper Classification using Supervised Machine Learning Techniques - Intermountain Engineering, Technology and Computing (IETC) – IEEE
<https://ieeexplore.ieee.org/document/9249211>

[8] Haihua Chen, Lei Wu, Jiangping Chen, Wei Lu, Junhua Ding (2022). A comparative study of automated legal text classification using random forests and deep learning, Information Processing & Management, Volume 59, Issue 2, 2022, 102798, ISSN 0306-4573,
<https://doi.org/10.1016/j.ipm.2021.102798>.

[9] Dobratinich Gonzalo (2022). INTELIGENCIA ARTIFICIAL Y JUSTICIA: APLICABILIDAD DE LA TECNOLOGÍA EN LAS DECISIONES JUDICIALES EN ARGENTINA. Revista Direitos Culturais. 17. 203-216. 10.20912/rdc.v17i42.761.

https://www.researchgate.net/publication/363593342_INTELIGENCIA_ARTIFICIAL_Y_JUSTICIA_APLICABILIDAD_DE_LA_TECNOLOGIA_EN_LAS_DECISIONES_JUDICIALES_EN_ARGENTINA

[10] Gonzalo Ana Dobratinich (2021)- Evaluación De La Preparación Del Sistema Judicial Para La Adopción De Inteligencia Artificial – Universidad de San Andrés -
<https://repositorio.udesa.edu.ar/jspui/handle/10908/18634>

Una herramienta de Inteligencia Artificial para el modelado de distintos escenarios con la determinación del Índice de Calidad del Agua (ICA) para aplicar a ríos urbanos de Argentina

Carlos G. López

Universidad Nacional del Oeste
gustavolopez@uno.edu.ar

Eugenia C. Artola

Universidad Nacional de Cuyo
ecartola@hotmail.com

Resumen

Una de las actividades más importantes para la gestión del recurso hídrico es el monitoreo periódico de los cuerpos de agua. El Índice de Calidad del Agua (ICA) es un indicador compuesto que integra información de varios parámetros de calidad del agua y presenta diferentes metodologías para su cálculo.

Con el objetivo de simular la ocurrencia de distintos escenarios y su impacto ambiental se propone diseñar una plataforma que permita anticipar el impacto que producen las descargas contaminantes en el curso de agua principal de las cuencas de los ríos, basándose en criterios lógicos y científicos, a través del desarrollo e implementación de un sistema de inteligencia artificial denominado Red Neuronal Artificial (RNA).

Los distintos escenarios permiten analizar diferentes formas de gestionar los recursos tendientes a lograr una Gestión Integrada del Recurso Hídrico (GIRH), que sea eficiente.

Palabras Claves: Inteligencia Artificial, medio ambiente, recurso hídrico, Red Neuronal Artificial (RNA).

Introducción

La contaminación de algunos recursos hídricos es mayor de lo esperado para un país que se encuentra en etapa de desarrollo como es Argentina [1]. Este hecho es uno de los problemas percibidos en el Plan Nacional Federal de Recursos Hídricos (PNFRH), que tiene como uno de sus objetivos “reducir y prevenir la contaminación de los recursos hídricos, y proteger y preservar los ecosistemas vinculados a los mismos” [2]. Si bien los recursos hídricos se ven degradados a lo largo de todo su curso, puede distinguirse el tipo de contaminación de acuerdo a la altura de su cuenca. Así, por ejemplo, encontramos ríos cuya contaminación varía de acuerdo al siguiente esquema [3]:

- En la cuenca alta, las aguas de encuentran afectadas por la contaminación de origen agrícola.
- En la cuenca media, se evidencia una degradación progresiva de la calidad del agua con aumento del grado de afectación de las aguas, por el mayor aporte de vertidos de sustancias contaminantes de origen domiciliario e industrial.
- En la cuenca baja, las aguas del río se encuentran severamente degradadas debido al aporte acumulado de

contaminantes vertidos en toda la cuenca y por los vertidos de contaminantes de origen domiciliario e industrial.

Se observa que este esquema de contaminación es frecuente en los ríos de Argentina.

La calidad del agua de las cuencas puede ser estudiada mediante bioensayos ecotoxicológicos que muestran la degradación del recurso de una forma integrada, no solamente desde los aspectos fisicoquímicos sino desde la forma en que afecta a la biota y principalmente en su etapa de ciclo de vida más susceptible, la embrionaria [4].

Objetivos

Principal:

- Desarrollar e implementar una Red Neuronal Artificial (RNA) que permita modelar distintos escenarios posibles a fin de anticipar el impacto que producen las descargas contaminantes en un río urbano de Argentina.

Secundarios:

- Seleccionar los parámetros más representativos para la definición del Índice de Calidad de Agua (ICA) de los ríos urbanos de Argentina.
- Seleccionar un río del área de influencia para su estudio.
- Recolectar sistemáticamente muestras que permitan entrenar la RNA implementada.
- Modelar distintos escenarios posibles ante eventuales catástrofes por contaminación.
- Analizar los datos obtenidos, para su discusión en la problemática planteada.

Antecedentes

El estado de contaminación de origen orgánico e inorgánico en que se encuentra en general el agua superficial de los ríos urbanos argentinos, limita sus usos potenciales afectando la calidad de vida de la población que habita en su entorno y representa un riesgo para la salud no solo de los habitantes sino también de los ecosistemas que dependen del recurso hídrico. La potencial contaminación de las aguas subterráneas en la cuenca, especialmente la freática, representa un riesgo para la salud de un número significativo de hogares sin servicio público de provisión de agua potable, en particular para los de bajos recursos, que dependen del agua subterránea para sus necesidades diarias [3].

Los usos residenciales, comerciales y productivos han transformado a los ríos en urbanos sin haberse resuelto adecuadamente las cuestiones ambientales que esta nueva situación plantea. En particular, la degradación de la calidad de sus aguas, ha restringido la mayoría de sus usos, entre los que cabe destacar el de protección de vida acuática; recreación con y sin contacto directo; fuente de agua para bebida de ganado; uso industrial; y fuente de agua para consumo humano con tratamiento convencional [3]. En línea con las recomendaciones internacionales [5, 6, 7] y con el PNFRH [2], las nuevas intervenciones requeridas necesitan de la evaluación de la calidad del agua y de la dinámica de los procesos de contaminación con el objeto de desarrollar herramientas que permitan efectuar procesos de saneamiento más eficientes [8].

Los Índices de Calidad del Agua reducen en su construcción, la información correspondiente a una gran cantidad de parámetros físicos, químicos y microbiológicos a un solo valor numérico resultante de la sumatoria de determinados parámetros seleccionados en base a su representatividad respecto de la calidad del recurso hídrico evaluado, dadas las particularidades conocidas de este último [9]. Existen antecedentes de ICA que son representados a través de un conjunto de índices, produciendo de igual forma la

mencionada reducción. La calidad del agua no es un concepto absoluto, es decir, su análisis se asocia a un uso concreto preestablecido o deseable, como puede ser el de la protección de la vida acuática, riego, consumo humano con tratamiento convencional, recreación con o sin contacto directo, entre los principales. La calidad del agua según su uso, se define en función de un conjunto de variables o parámetros, así como de sus valores de aceptación o de rechazo, que son los indicadores de la calidad de la misma [9]. Las exigencias y/o restricciones son diferentes de acuerdo al uso preestablecido o a establecer y, por lo tanto, la estructura y expresión matemática de los índices, deben tener en cuenta esta condición. Por lo tanto, los ICA varían según la información que refleje el valor total resultante de la agregación de parámetros. Algunos ICA, a través de la escala o rango establecido, aportan información general que no discrimina usos, o en todo caso, brinda una idea general de usos recomendados o restricciones concretas según el caso o categoría en la que se presenta el valor final del ICA [10]. Otros ICA permiten claramente discriminar, para un único valor, la calidad del agua y su aptitud para diferentes usos; es decir, la escala o rango jerárquico cambia según el uso. Existen otros Índices de Calidad del Agua más específicos que no apuntan a usos, sino a causas de deterioro o tipo de contaminación. Son índices independientes entre sí y complementarios, que brindan información individual respecto de diferentes situaciones, como, por ejemplo, aportes de nutrientes responsables de la eutrofización, descargas de materia orgánica, contaminación por patógenos, presencia de sustancias tóxicas, etcétera [8].

Los ICA son instrumentos de síntesis y, en este sentido, por sus propias características estructurales y funcionales son simplificaciones de la realidad, condicionadas por la complejidad del sistema abordado, por lo que deben utilizarse con restricciones y en asociación-combinación con otras herramientas diagnósticas [11, 10].

Metodología

El modelo a desarrollar para el cálculo del ICA en ríos urbanos de Argentina, incluye la revisión de todas las fuentes de contaminantes identificadas para dicho recurso hídrico.

Para cada una de las fuentes de contaminantes, se evaluará el aporte aproximado de efluentes domiciliarios, industriales y agrícolas. Las fuentes pueden volcar sus efluentes al río a través de tributarios de caudal permanente y significativo (arroyos), o volcar sus efluentes a través de tributarios de menor caudal (zanjones, pluviales, etc.). Se relevarán las descargas de afluentes, teniendo en cuenta los aportes recibidos y especialmente su procedencia.

Estudios previos [8, 11] sugieren que el ICA debe estar integrado por los siguientes parámetros: Oxígeno Disuelto (OD), Demanda Bioquímica de Oxígeno a cinco días (DBO5), Demanda Química de Oxígeno (DQO), Fósforo Total (P total), Nitrógeno Amoniacal (N-NH_4^+), Sólidos Suspendidos Totales (SST), Conductividad Eléctrica (CE), Escherichia coli, pH, Cromo Total (Cr total), Plomo Total (Pb total) e Hidrocarburos Totales (HCT). En este sentido, se procederá a la convalidación de dichos parámetros o a definir los que se consideren pertinentes a los fines de la presente investigación.

Se seleccionará un río urbano de la zona de influencia y se procederá a la extracción de muestras y su posterior análisis. Se medirán los parámetros hidrológicos. Los datos obtenidos en campo deberán registrarse mediante la aplicación de un protocolo.

Se deberá proceder al análisis de los resultados obtenidos y se discutirá la problemática relevada.

Se diseñará, mediante la implementación de técnicas de inteligencia artificial, un Sistema Experto basado en conocimiento a fin de obtener un sistema de predicción de la calidad del agua superficial ante posibles vuelcos de contaminantes.

La Red Neuronal Artificial (RNA) implementada basará su aprendizaje en la asociación de patrones existentes. Tal como lo hace el cerebro humano, la capacidad de generalizar se logra mediante un conjunto de muestras escogidas como patrones de entrenamiento. Este tipo de red requiere una búsqueda de un tipo de comportamiento que se acople a los valores muestreados y al comportamiento del sistema, con un valor mínimo de error. Asimismo, esta red necesita de una etapa de aprendizaje.

Se propone para la RNA implementada un tipo de aprendizaje supervisado, es decir basada en deducir una función a partir de datos o información de entrenamiento. Dichos datos están compuestos por vectores u objetos dados en pares de datos. Una de las componentes del par es la entrada mientras que la otra es el resultado deseado. La finalidad de este sistema de aprendizaje es poder predecir un resultado de manera eficaz a partir de una entrada válida que se ingrese al sistema.

Ventajas de la propuesta

Al implementar esta propuesta, se pretende lograr:

- Aprendizaje

La Red Neuronal Artificial (RNA) propuesta aprenderá a realizar tareas o tomar acciones a partir del conjunto de datos de entrada, obtenidos mediante la toma de muestras en el río, siendo así que, en el proceso de aprendizaje a partir de estos datos, serán representados mediante entradas y pesos asignados a las mismas.

- Auto-organización

Las redes neuronales organizan la información que reciben durante el proceso de aprendizaje utilizando el método matemático Perceptron. Pueden crear su propia organización o representación de la información recibida.

- Tolerancia a fallos

La destrucción parcial de la red daña el funcionamiento de la misma, pero no la

destruye completamente. Esto es debido a la redundancia de la información contenida, es decir, esto conlleva que la información no se pierda ya que funciona en forma similar al cuerpo humano.

Desventajas de la propuesta

- Tiempo de aprendizaje considerable. Esto dependerá del número de patrones a reconocer y la flexibilidad para reconocer patrones que sean bastante parecidos.

- No tiene la capacidad de interpretar los resultados que la RNA genere siendo necesaria la interpretación del desarrollador de la aplicación para que le dé significado a tales resultados.

Conclusiones

El presente trabajo permitirá simular la ocurrencia de distintos escenarios, para lo que se diseñará una plataforma que, mediante la toma sistemática de muestras en un río urbano de Argentina, permita anticipar el impacto que producen las descargas contaminantes en el curso de agua principal de la cuenca, basándose en el análisis de los datos recogidos y por la implementación de un sistema de inteligencia artificial.

La motivación principal de la investigación se basa en el hecho de considerar, que es posible diseñar un sistema basado en inteligencia artificial para que mediante un procedimiento de aprendizaje continuo y evolutivo modele distintos escenarios posibles de ocurrencia en función de los muestreos que se tomarán a lo largo de dicho estudio.

Lo descripto anteriormente, permitirá desarrollar herramientas que posibiliten proponer procesos de saneamiento más eficientes en línea con las recomendaciones internacionales y con el PNFRH del Consejo Hídrico Federal, que establecen como prioritarias las intervenciones para la evaluación de la calidad del agua y del estudio de la dinámica de los procesos de contaminación ambiental en Argentina.

Referencias

- [1] Castilla V., Canevaro, S., López, B.. Migración, degradación ambiental y percepciones del riesgo en la cuenca del río (Buenos Aires, Argentina), Revista de Estudios Sociales [En línea], 76 | 01 abril de 2021, Publicado el 04 abril 2021, consultado el 15 julio 2022. URL: <http://journals.openedition.org/revestudsoc/49626>
- [2] Consejo Hídrico Federal. Plan Nacional Federal de los Recursos Hídricos. 2007
- [3] ACUMAR, Autoridad de Cuenca Matanza-Riachuelo, https://www.acumar.gob.ar/wp-content/uploads/2016/12/CDCA_ACUMAR_INFO_AGUA_abril-2022_completo-web-con-anexos_compressed.pdf Accesed 2022.
- [4] Merlinsky, G., Tobías, M.. Conflictos por el agua en las cuencas de los ríos Matanza-Riachuelo y Reconquista. Claves para pensar la justicia hídrica a escala metropolitana. Punto Sur, (5). 2021. <https://doi.org/10.34096/ps.n5.10998>
- [5] Dourojeanni, A., Jouravlev, A., & Chávez, G.. Gestión del agua a nivel de cuencas: teoría y práctica. CEPAL. 2002.
- [6] Global Water Partnership. Manejo integrado de recursos hídricos (p. 80). Global Water Partnership. 2000. Obtenido de www.gwpforum.org
- [7] United Nations. Agenda 21 - United Nations Conference on Environment & Development (p. 351). Rio de Janeiro, Brazil. 1992.
- [8] Nader, G. M.. Evaluación de la calidad del agua en un río urbano. Universidad Nacional de San Martín. 2015.
- [9] Acero Vargas, J.. Desde el fondo: espacio, conflicto y acción colectiva en la periferia bonaerense. Universidad Nacional de Colombia. 2021.
- [10] Cantera, C.G., Tufo, A.E., Scasso, R.A. . Geochemical characterization and the assessment of trace element retention in sediments of the Reconquista River, Argentina. Environ Geochem Health **44**, 729–747 (2022). 2022. <https://doi.org/10.1007/s10653-021-00970-7>
- [11] Anastopoulos, I. y Kyzas, G.. Progress in batchbiosorption of heavy metals ontoalgae. Journal of Molecular Liquids. 2015.

MOVILIDAD INTELIGENTE EN UNA CIUDAD SOSTENIBLE

Villagra A¹., Errecalde M^{1,2}., Pandolfi D¹., Molina D¹., Mercado V¹., Rasjido J.¹, Ramos L¹., Del Do M¹., Orozco S¹., Villagra A.M¹.

¹Laboratorio de Tecnologías Emergentes (LabTEM)
Instituto de Tecnología Aplicada (ITA) - Unidad Académica Caleta Olivia
Universidad Nacional de la Patagonia Austral

²Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)
Departamento de Informática - Universidad Nacional de San Luis

avillagra@uaco.unpa.edu.ar, merreca@unsl.edu.ar, {dpandolfi, dmolina, vmercado, jrasjido, lramos, mdeldo, srozc, amvillagra}@uaco.unpa.edu.ar

RESUMEN

El tráfico en las grandes ciudades se ha vuelto una problemática de primera necesidad hoy en día. No solo por los problemas logísticos que supone, sino por las emisiones de gases asociadas. Tener en cuenta el impacto que tiene la movilidad en las ciudades en el medio ambiente es de vital importancia para un futuro sostenible. Precisamente, en los Objetivos de Desarrollo Sostenible de la Agenda 2030 el objetivo número 11 “Ciudades y Comunidades Sostenibles” potencian las ciudades inteligentes desde la perspectiva de la sostenibilidad.

En esta línea de trabajo se presenta una propuesta de investigación enfocada en los desafíos relacionados con la movilidad. Uno de nuestros objetivos es mejorar y optimizar las ciudades para que sean más adaptables a los ciudadanos y a las nuevas preocupaciones ambientales. Principalmente nos estamos enfocando en optimizar la planificación de los ciclos semafóricos para hacer el tráfico más fluido reduciendo atascos y tiempos de viaje. Además, se reducen las emisiones de gases a la atmósfera que impactan en el ambiente y en la salud de las personas.

Palabras clave: Ciudades Inteligentes, Optimización, Planificación semafórica, Movilidad Inteligente.

CONTEXTO

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/B273 “Ciudades inteligentes y sostenibles: iniciativas y desafíos”. Este proyecto se desarrolla en cooperación con el LIDIC de la UNSL, y el Grupo NEO de la UMA (España).

1. INTRODUCCIÓN

Desde la gran migración del campo a la ciudad, los centros urbanos han sido el escenario del crecimiento de nuestra sociedad. No sólo porque permiten un número más importante de puestos de trabajo, sino también por los muchos beneficios que ofrecen salud, seguridad, ocio, etc. Múltiples estudios afirman que en sólo 30 años más del 66% de la población mundial vivirá en ciudades [18].

Clásicamente, los diferentes servicios ofrecidos eran concebidos y gestionados por personas, generalmente pertenecientes a los dis-

tintos organismos públicos de la ciudad. Sin embargo, los enormes avances tecnológicos de los últimos tiempos han permitido la comunicación y la automatización de diferentes tareas, algo nunca visto antes. Estas ventajas técnicas incluyen avances en las telecomunicaciones inalámbricas sensores de bajo costo, actuadores y recolección automática y masiva de datos de la ciudad, por citar algunos.

Estos avances tecnológicos han llevado a pasar de una sociedad tradicional a una en la que la vida es inconcebible sin la tecnología. Por supuesto, este cambio no sólo se ha producido a pequeña escala, a nivel individual. También se ha producido a nivel social y, por supuesto, a nivel urbano.

En este contexto, las ciudades se están modernizando y convirtiendo en las llamadas Ciudades Inteligentes (CI). Mejorar todos los procesos de la ciudad desde una perspectiva holística es el aspecto clave que caracteriza a la CI. Toda esta mejora es gracias a la tecnología, los avances en las comunicaciones y la inteligencia artificial.

Las ciudades son grandes ecosistemas en los que interactúan diversos actores.

Podemos clasificar las partes interesadas en ciudadanos, empresas y gestores de la ciudad. Una de las visiones más utilizadas es la definida por Cohen [19] que propuso seis aspectos que conformaban las CI: Gobernanza Inteligente, Movilidad Inteligente, Ambiente Inteligente, Vida Inteligente, Economía Inteligente y Gente Inteligente.

En este trabajo tratando de aportar al logro de los ODS (en particular el nro. 11) nos centramos en Movilidad Inteligente. La fluidez del tráfico en las grandes ciudades inteligentes se ha convertido en uno de los problemas más graves a los que se enfrentan las grandes ciudades. En algunos casos, este problema se agrava aún más debido a la gran cantidad de atascos, accidentes de tráfico, o incluso heridos o muertes. Por ello es necesario regular el tráfico con algunos elementos como semáforos. Cuanto mayor sea el área metropolitana, mayor será el número de semáforos necesarios para regular el

flujo de tráfico. Una gestión óptima del tráfico puede ser beneficiosa para minimizar la duración de los trayectos, reducir el consumo de combustible y las emisiones nocivas.

Los semáforos son elementos simples pero esenciales utilizados en entornos urbanos para organizar el tránsito, principalmente de vehículos [1]. A diferencia de técnicas como los semáforos inteligentes, que requieren nuevas infraestructuras, la ubicación de sensores y modificaciones en obras civiles ya estructuradas, el uso de técnicas de inteligencia artificial en la optimización de los ciclos de los semáforos se presenta como una herramienta viable, rápida, eficiente y de bajo costo [2].

Las metaheurísticas se han utilizado ampliamente para abordar el problema de planificación del tránsito, especialmente el algoritmo genético (GA), la optimización por colonia de hormigas (ACO) y la optimización por enjambre de partículas (PSO) y Evolución Diferencial (DE) [3-12], entre otras.

Para abordar la movilidad inteligente será necesario la aplicación de técnicas de vanguardia de diferentes dominios que permitan flexibilidad, autoadaptabilidad, robustez, alta dimensionalidad (escalabilidad) y eficiencia. Además, de la construcción nuevos prototipos basados en sistemas inteligentes, mejorados con metodología y tecnologías diferentes con el fin de exhibir "inteligencia holística". En particular se pretende contribuir a la reducción del consumo de combustibles, la emisión de gases, y el tiempo de traslado, así como al mejoramiento de la calidad del aire.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se describe la línea de investigación que se lleva a cabo en el proyecto.

Como en muchas ciudades del mundo real, el control en tiempo real de los semáforos [20] no es factible por diversas razones (legales, técnicas, etc.), y, en su lugar, debemos encontrar una programación global de semáforos altamente fiable que funcione bien en el sistema

de tráfico dinámico e incierto [2,5,10,16,17]. A la hora de optimizar los programas de ciclos de luces de los semáforos dentro de una ciudad para mejorar el flujo de tráfico y reducir la contaminación, la idoneidad de un programa semafórico candidato se evalúa mediante la simulación de rutas de vehículos y velocidades sobre una red de tráfico dada [10, 11, 17, 21]. La simulación de flujos de tráfico en una ciudad concreta requiere recopilar datos de la red (topología de la zona e información sobre los semáforos), que suele ser precisa y estática, y los datos de tráfico (el número de vehículos, sus desplazamientos y velocidades) se estiman a partir de datos reales muy dinámicos. Dada la incertidumbre inherente de esta estimación, es posible generar distintos escenarios de tráfico que son coherentes con el sistema del mundo real [15]. Una forma de tener en cuenta esta incertidumbre es calcular un valor de aptitud agregado simulando la misma solución candidata varias veces utilizando diferentes escenarios de tráfico [22]. Otra forma alternativa (o adicional) es hacer que cada simulación sea estocástica introduciendo cambios aleatorios en el escenario de tráfico durante la simulación [5]. La aptitud de un programa semafórico fiable no debería presentar una alta varianza cuando se evalúa en el tráfico dinámico de la ciudad real.

A pesar de la incertidumbre inherente a los escenarios de tráfico simulados, la literatura sobre optimización de semáforos suele basarse en la simulación determinista de un único escenario de tráfico [5, 10, 17, 21]. Cuando se consideran múltiples escenarios, se utilizan para evaluar la flexibilidad del algoritmo de optimización optimizando cada escenario por separado. En [22] los autores validaron la fiabilidad de las soluciones candidatas en múltiples escenarios tras la fase de optimización. Sin embargo, cada solución se sigue optimizando en relación con un único escenario.

Para el desarrollo de esta línea se han utilizado escenarios reales de la ciudad de Málaga y de París y se ha seguido el método científico, a fin de analizar cuestiones de investiga-

ción abiertas en estos campos, definir los desafíos y validar nuestros resultados con prototipos reales. Dar soluciones significa resolver muchos problemas tecnológicos, necesidad de plataformas de alto rendimiento, y visualización intuitiva (mapas digitales y gemelo). Las tecnologías necesarias son muchas y modernas, pero serán inútiles sin una inteligencia real capaz de extraer información de extensos conjuntos de datos junto con una gran cantidad de complejos problemas de optimización por resolver.

Nuestra hipótesis es que creando nuevos algoritmos bioinspirados podremos cumplir con los requisitos, especialmente cuando se combinan con técnicas de aprendizaje automático para ofrecer una neuro-evolución rápida, algoritmos paralelos multiobjetivo, selección de características y metaheurísticas dinámicas, por nombrar algunos. Además, la incorporación de funciones subrogadas aportará mayor eficiencia computacional al proceso de optimización.

3. RESULTADOS OBTENIDOS/ESPERADOS

Hemos abordado el problema de programación de semáforos con dos versiones de un Algoritmo Genético Celular, cGA (síncrono y asíncrono) para resolver instancias grandes y reales. Nuestros algoritmos superan las técnicas de vanguardia y las configuraciones de expertos. Se llevaron a cabo varios análisis en profundidad de los resultados, estudio genotípico y fenotípico (para mayor detalle ver [17]).

La investigación actual y futura se centra en lograr disminuir el número de evaluaciones con el objetivo de reducir el esfuerzo computacional en términos de tiempo de procesamiento. Planeamos definir una función de aptitud más precisa que explore el espacio de búsqueda de manera más eficiente. También, aplicar diferentes modelos de funciones subrogadas utilizando redes neuronales artificiales (ANN) y otros modelos de SMT (*Surrogate Modeling Toolbox*). Integrar los modelos con el cGA y analizar resultados.

Tenemos previsto modelar diferentes problemas con datos de varias ciudades para nuestros algoritmos de optimización., basándonos en trabajos como [15 y 16] donde se realiza la importación de mapas de *OpenStreetMap* en el simulador de tráfico SUMO. Los mapas realistas incluirán rotondas reales, semáforos, etc. Además, utilizaremos datos abiertos publicados por las autoridades locales, para mejorar la precisión. También se pretende desarrollar varios prototipos y servicios web para poner en práctica nuestras ideas de movilidad inteligente, ofreciendo servicios web para optimizar las rutas.

Recientemente para reducir el esfuerzo computacional de cGA, se utiliza una ANN como función subrogada, es decir, en lugar de evaluar la solución con la simulación con SUMO, la red neuronal artificial entrenada predice el valor de aptitud [23]. Esta ANN recibe una solución obtenida por cGA y devuelve la función de fitness correspondiente (sin evaluar la solución con SUMO). Para analizar el rendimiento de cGA con la función subrogada, definimos diferentes formas de aplicar el modelo ANN. La primera, denominada cGA-s1, aplica la ANN en el 100% de las evaluaciones, la segunda, denominada cGA-s2, aplica la ANN en el primer 50% de las evaluaciones, y la tercera, denominada cGA-s3 utiliza SUMO en el primer 50% de las evaluaciones y para el resto el modelo ANN. Con los experimentos realizados, hemos comprobado que el uso de una función subrogada reduce el esfuerzo computacional (en tiempo) para obtener el mejor valor de fitness para las instancias analizadas. Hemos reducido el tiempo de ejecución entre 60% en cGA-s2 y cGA-s3 (utilizando sólo parcialmente el modelo sustituto) y 99% en el enfoque cGA-s1 (que no utiliza el simulador).

En futuros trabajos, hay varias cuestiones que pretendemos abordar, como el uso de otras funciones subrogadas o estrategias de optimización como la bayesiana y, de combinar las evaluaciones SUMO con las subrogadas. Además, en lo que respecta a la reducción de los parámetros iniciales, sería interesante aplicar procesos de selección de características para

reducir el espacio de búsqueda y analizar el comportamiento del algoritmo. Además, utilizar un modelo ANN de ajuste dinámico preentrenado que nos permita probar nuevas técnicas y realizar más experimentos (otras rutas). También en cuanto a la generación de conjuntos de datos, estudiar diferentes tamaños de conjuntos de datos y diferentes formas de generar soluciones (muestreo aleatorio, hipercubos latinos, entre otros). Por último, definir la función objetivo como una función multiobjetivo.

Finalmente, se procurará colaborar con los gobiernos locales para la implementación de políticas y acciones inteligentes y sostenibles que impacten en la calidad de vida de los ciudadanos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado dos doctores y dos Magister en Ciencias de la Computación, seis Ingenieros y Licenciados en Sistemas cursando estudios de Posgrado.

Este proyecto de investigación proporcionará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes.

5. BIBLIOGRAFIA

- [1] Papageorgiou, M., Diakaki, C., Dinopoulou, D., Kotsialos, A., Wang, Y. Review of road traffic control strategies, *Proc. IEEE* 91 (2003) 2043–2067.
- [2] Garcia-Nieto, J., Ferrer, J., Alba, E. Optimising traffic lights with metaheuristics: reduction of car emissions and consumption, in: *International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2014, pp. 48–54.
- [3] Celtek, S. A., Durdu, A., & Ali, M. E. M. (2020). Real-time traffic signal control with swarm optimization methods. *Measurement*, 166, 108206.
- [4] Shaikh, P. W., El-Abd, M., Khanafer, M., & Gao, K. (2020). A review on swarm intelligence and evolutionary algorithms for solving

the traffic signal control problem. *IEEE transactions on intelligent transportation systems*, 23(1), 48-63.

[5] Sánchez, J., Galán, M., Rubio, E., Applying a traffic lights evolutionary optimization technique to a real case: “Las Ramblas” area in Santa Cruz de Tenerife, *IEEE Trans. Evol. Comput.* 12 (2008) 25–40.

[6] Stolfi, D. H., & Alba, E. (2021). Yellow Swarm: LED panels to advise optimal alternative tours to drivers in the city of Malaga. *Applied Soft Computing*, 109, 107566.

[7] Baskan, O., Haldenbilen, S., Ant colony optimization approach for optimizing traffic signal timings, in: *Ant Colony Optimization-Methods and Applications*, InTech, 2011.

[8] Nguyen, T. H., & Jung, J. J. (2021). Ant colony optimization-based traffic routing with intersection negotiation for connected vehicles. *Applied Soft Computing*, 112, 107828.

[9] Cui, C.-Y., Lee, H.-H., Distributed traffic signal control using PSO based on probability model for traffic jam Intelligent Autonomous Systems, 12, Springer, 2013, pp. 629–639.

[10] García-Nieto, J., Alba, E., Olivera, A., Swarm intelligence for traffic lightscheduling: application to real urban areas, *Eng. Appl. Artif. Intell.* 25 (2012) 274–283.

[11] Garcia-Nieto, J., Olivera, A., Alba, E., Optimal cycle program of traffic lights withparticle swarm optimization, *IEEE Trans. Evol. Comput.* 17 (2013) 823–839.

[12] Souravlias, D., Luque, G., Alba, E., Parsopoulos, K., Smart traffic lights: a first parallel computing approach, in: *Intelligent Networking and Collaborative Systems (INCoS)*, International Conference on Intelligent Networking and Collaborative Systems (INCoS), IEEE, 2016, pp. 229–236.

[13] Shaheen, S., Rodier, C. J., & Finson, R. S. (2003). University of California, Davis Long-Range Development Plan: A Davis Smart Mobility Model. California Partners for Advanced Transit and Highways (PATH).

[14] Arnott, R., Rave, T., & Schöb, R. (2005). Alleviating urban traffic congestion. MIT Press Books, 1.

[15] Stolfi, D. H., & Alba, E. (2014). Red Swarm: Reducing travel times in smart cities by using bio-inspired algorithms. *Applied Soft Computing*, 24, 181-195.

[16] Stolfi, D. H., & Alba, E. (2014, July). Eco-friendly reduction of travel times in european smart cities. In *Proc. of the 2014 Annual Conference on Genetic and Evolutionary Computation* (pp. 1207-1214).

[17] Villagra, A., Alba, E., & Luque, G. (2020). A better understanding on traffic light scheduling: New cellular GAs and new in-depth analysis of solutions. *Journal of Computational Science*, 41, 101085.

[18] Economic, United Nations Department of and Social Affairs (2019). *World Urbanization Prospects: The 2018 Revision*.

[19] Cohen, Boyd (2012). “What exactly is a smart city”. In: *Co. Exist* 19.

[20] Cao, Z., S. Jiang, J. Zhang, and H. Guo (2017). “A unified framework for vehicle rerouting and traffic light control to reduce traffic congestion”. In: *IEEE Trans. on Intelligent Transportation Systems* 18.7, pp. 1958–1973.

[21] Péres, M., Ruiz G., Nesmachnow S., and Olivera C. (2018). “Multiobjective evolutionary optimization of traffic flow and pollution in Montevideo, Uruguay”. In: *Applied Soft Computing* 70, pp. 472–485.

[22] Ferrer, J., García-Nieto J., Alba E., and Chicano F. (2016). “Intelligent Testing of Traffic Light Programs: Validation in Smart Mobility Scenarios”. In: *Mathematical Problems in Engineering* 2016, pp. 1–19. DOI:10.1155/2016/3871046.

[23] Villagra A., Luque G. (2023). A surrogate function in cellular GA for the traffic light scheduling problema. *Applications of Evolutionary Computation – 26th International Conference, EvoApplications*. A Springer Nature Computer Science Series (en proceso de publicación).

Asistente virtual con comportamiento definible de consultas referidas a la legislación argentina para personas con discapacidad.

Caffetti, Yanina Andrea¹, Acosta Nelson², Horacio Kuna³.

¹ Facultad de Ciencias Forestales, Universidad Nacional de Misiones.

² Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires.

³ Facultad de Exactas Químicas y Naturales, Universidad Nacional de Misiones.

yanina.caffetti@fcf.unam.edu.ar

el.nelson.acosta@gmail.com

hdkuna@gmail.com

Palabras claves: Interacción persona-computadora, agente inteligente, procesamiento del lenguaje natural, inteligencia artificial conversacional.

RESUMEN

El presente artículo forma parte de un proyecto de investigación que se enmarca dentro de una tesis de Doctorado en Informática UNNE-UNaM-UTN-FRRe. Tiene como objetivo la realización de un asistente virtual de consultas referidas a la legislación argentina para personas con discapacidad ofreciendo respuestas con connotaciones emocionales, en este punto es de destacar que el desarrollo presenta la posibilidad de elección del tipo de connotación emocional del asistente por parte del usuario.

CONTEXTO

En la actualidad el procesamiento del lenguaje natural, el aprendizaje profundo y el aprendizaje automático, son las bases de la interacción entre un usuario y una computadora, siendo los asistentes virtuales el ejemplo destacado en la literatura[1]. Por otro lado, las consultas por medio de la escritura no fluida debido a espasmos musculares o de construcciones gramaticales puntuales, por ejemplo de personas que presentan un desarrollo cognitivo particular, son un impedimento en dicha interacción.

En este contexto, la línea de investigación presentada por medio de este artículo, desglosa las tareas planificadas en el desarrollo del asistente virtual con comportamiento definible basado en reglas, quién aportará entonces conocimiento en el área de inteligencia artificial, modificando al aprendizaje automático del agente a través de las necesidades y elecciones diarias del usuario.

En lo relativo al procesamiento del lenguaje natural y poniendo énfasis en la consulta, se plantea que el asistente virtual tenga un comportamiento definible por el usuario utilizando la retroalimentación humana como parte del entrenamiento y será desarrollado utilizando reglas del asistente robot para fines sociales (RASA, Robot Assistant for Social Aims).

1. INTRODUCCIÓN

A los grandes modelos lingüísticos (LM) se les puede “pedir” que realicen una serie de tareas de procesamiento del lenguaje natural (NLP), dándoles como entrada algunos ejemplos de la tarea. Sin embargo, estos modelos a menudo expresan comportamientos no intencionados, como inventar hechos, generar textos sesgados o tóxicos, o simplemente no seguir las instrucciones del usuario[2]-[7]. Esto se debe a que el objetivo de modelado del lenguaje utilizado para muchos grandes LM recientes -predecir el siguiente token en una página web de Internet por ejemplo- es diferente del objetivo “seguir las instrucciones del usuario de forma útil y segura”[8]-[12]. Por lo tanto, decimos que el objetivo de modelado del lenguaje está desfasado. Evitar estos

comportamientos no deseados es especialmente importante para los modelos lingüísticos que se despliegan y utilizan en cientos de aplicaciones[13]. En este contexto se presenta la posibilidad de creación de un agente inteligente basado en reglas en donde el modelado está nivelado con la intención del usuario y ajustado a la retroalimentación del mismo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Corresponde señalar que si el asistente virtual con comportamiento definible basado en reglas, presenta la función de responder a una consulta, con una connotación emocional predeterminada por el usuario, como ser: empatía, antipatía e indiferencia, entonces, el algoritmo de aprendizaje del asistente será modificado, predefiniendo el usuario al consultar, la emoción de la respuesta. En base a esta hipótesis se presenta como línea de investigación el modelo InstructGPT[13] que básicamente ajusta los parámetros de entrada a través de la intervención de datos humanos. Los modelos InstructGPT generan salidas más apropiadas y siguen con mayor fiabilidad las restricciones explícitas de la instrucción.

Los pasos a seguir son los siguientes, con esta línea de investigación como base:

- Etapa estado del arte: desarrollo de una revisión sistemática de la literatura (RSL, Review Systematic of Literature).
- Etapa construcción base de datos: el sistema de gestión de base de datos relacional orientado a objetos y de código abierto que se utilizará para la creación de la base de datos del marco regulatorio argentino será posiblemente PostgreSQL. En GitHub se realizará la documentación de los sprint necesarios correspondientes a la metodología scrum para el desarrollo del mismo.

- Etapa modelo algorítmico y evaluación del agente conversacional: la plataforma de código abierto RASA complementa la creación del agente inteligente conversacional junto con el lenguaje de programación Prolog. GitHub nuevamente se utilizará para la documentación de los sprints. Los requisitos del software en formato sprint son necesarios para adaptar las características específicas de las entradas del usuario y construir la metodología del aprendizaje supervisado del agente inteligente conversacional utilizando el ajuste de parámetros por medio de etiquetadores humanos. Esta etapa se realizará en paralelo junto con la etapa de construcción de la base de datos, ya que es necesario definir entidades de entrenamiento y emociones con un primer etiquetado.

3. RESULTADOS OBTENIDOS/ESPERADOS.

De verificarse la hipótesis del presente proyecto, generará un aporte innovador a la inteligencia artificial conversacional, incorporando al usuario como partícipe necesario en la construcción del aprendizaje del asistente. Específicamente modificará las reglas de entrenamiento en consideración a los datos de entrada y utilizará la verificación del etiquetado planteado por InstructGPT.

El asistente virtual será entrenado acerca de la legislación argentina, ofreciendo un entorno accesible e inclusivo a personas con discapacidad, lo que representa una posibilidad de inclusión y equidad en el uso de una herramienta informática.

4. FORMACIÓN DE RECURSOS HUMANOS.

El proyecto representa una de las líneas de investigación abordada por la doctorando en Informática Mgter. Yanina A. Caffetti.

5. BIBLIOGRAFÍA

- [1] Ellen Adamopoulou, Lefteris Moussiades, “Chatbots: History, technology, and applications.” *Machine Learning with Applications*, Volume 2, ISSN 2666-8270, 2020.
- [2] Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. “On the dangers of stochastic parrots: Can language models be too big?” *In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623.
- [3] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg J., Bosselut, A., Brunskill, E., et al. “On the opportunities and risks of foundation models.” <https://arxiv.org/abs/2108.07258>, 2022.
- [4] Kenton, Z., Everitt, T., Weidinger, L., Gabriel, I., Mikulik, V., and Irving, G. “Alignment of language agents”. <https://arxiv.org/abs/2103.14659>, 2021.
- [5] Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., et al. “Ethical and social risks of harm from language models”. <https://arxiv.org/abs/2112.04359>, 2021.
- [6] Tamkin, A., Brundage, M., Clark, J., and Ganguli, D. “Understanding the capabilities, limitations, and societal impact of large language models”. <https://arxiv.org/abs/2102.02503>, 2021.
- [7] Gehman, S., Gururangan, S., Sap, M., Choi, Y., and Smith, N. A. “Real Toxicity Prompts: Evaluating neural toxic degeneration in language models”. <https://arxiv.org/abs/2009.11462>, 2020.
- [8] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. “Language models are unsupervised multi task learners”. *OpenAI Blog*, 1(8):9, 2020.
- [9] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. “Language models are few-shot learners”. <https://arxiv.org/abs/2005.14165>, 2020.
- [10] Fedus, W., Zoph, B., and Shazeer, N. “Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity”. <https://arxiv.org/abs/2101.03961>, 2022.
- [11] Rae, J. W., Borgeaud, S., Cai, T., Millican, K., Hoffmann, J., Song, F., Aslanides, J., Henderson, S., Ring, R., Young, S., et al. “Scaling language models: Methods, analysis & insights from training gopher”. <https://arxiv.org/abs/2112.11446>, 2022.
- [12] Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., Du, Y., et al. “LaMDA: Language models for dialog applications”. <https://arxiv.org/abs/2201.08239>, 2022.
- [13] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... & Lowe, R. “Training language models to follow instructions with human feedback.” <https://arxiv.org/abs/2203.02155>, 2022.

AUTOMATIZACIÓN ROBÓTICA DE PROCESOS EN LAS PYMES

Mirabete M., Villagra A., Pandolfi D.

Laboratorio de Tecnologías Emergentes (LabTEM)

Instituto de Tecnología Aplicada (ITA) - Unidad Académica Caleta Olivia

Universidad Nacional de la Patagonia Austral

{mmirabete, avillagra, dpandolfi}@uaco.unpa.edu.ar

RESUMEN

La Automatización Robótica de Procesos, o RPA (Robotic Process Automation), es una tecnología que tiene como objetivo reducir la intervención manual en el uso de programas y aplicaciones informáticas. Automatizando funciones y procesos administrativos, mecánicos y bien definidos. Tareas manuales de poco valor añadido, ya que no requieren inteligencia humana, para dejar en manos de los humanos aquellas cuestiones que requieran interpretación o análisis.

Básicamente es una forma de automatizar mediante fuerza de trabajo virtual. Tiene la particularidad de que son agentes de software, no son robots físicos como podríamos encontrar en una fábrica, no altera la infraestructura, usa las aplicaciones que ya existen y es una tecnología muy rápida de implementar.

Esta línea de investigación se propone como objetivo identificar un modelo que haga posible de una manera ágil y práctica, la implementación de RPA. El modelo será aplicado metodológicamente a un Caso de Estudio en el dominio de las PyMES (pequeñas y medianas empresas), con el propósito de evaluar su eficiencia, evidenciar sus beneficios y hallar oportunidades de mejora.

Palabras clave: Automatización Robótica de Procesos; RPA; Transformación Digital; PyMES; Tecnologías.

Esta línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia de la Universidad Nacional de la Patagonia Austral, en el marco del trabajo de tesis en la Maestría de Informática y Sistemas (MIS) del Lic. Martín Mirabete.

1. INTRODUCCIÓN

El dramaturgo Karel Capek acuñó el término "robot" en su drama satírico Rossum's Universal Robots [1]. La palabra robot deriva del término checo robota (trabajo esclavo). Desde entonces, varias obras famosas de ciencia ficción han popularizado la noción de Capek de los robots como máquinas totalmente autónomas. Aunque las versiones ficticias han caricaturizado nuestra concepción popular de un robot, una definición menos glamorosa, pero más científica, sería: Un robot es un dispositivo mecánico reprogramable, controlado por computadora, equipado con sensores y actuadores [2].

El término de automatización (del griego autos que significa "por sí mismo" y maiomai que significa "lanzar") se ha venido empleando en muchos campos; un ejemplo de esto son los procesos industriales donde se dio forma a esta palabra debido a la necesidad de minimizar la intervención humana en los procesos de gobierno directo en la producción,

CONTEXTO

optimizando los costos y tiempos de aplicación [3].

La automatización de procesos por robótica es una forma de procesar automáticamente actividades que típicamente son repetitivas y basadas en reglas de operación. La ejecución de estas actividades se realiza normalmente en procesos estructurados, centralizados o dentro del back office de las empresas. De acuerdo con lo descrito por Deloitte en su artículo "La era de la Automatización - Implementación de Robotics en las organizaciones" [4].

En estos últimos años debido a la pandemia de COVID-19 hemos vivido un contexto muy especial, que ha representado una amenaza existencial para las organizaciones, la Automatización Robótica de Procesos surgió como una palanca vital para garantizar la resiliencia y la agilidad empresarial. Todo esto, sumado a la necesidad de reducción de costos y la priorización de la eficiencia operativa han sido impulsores clave de la adopción de RPA [5].

Cuando se habla de robots tendemos a pensar en máquinas pesadas que llevan a cabo tareas industriales, principalmente en cadenas de montaje. RPA extrapola ese concepto de automatización a los modelos puramente informáticos, por lo que hablamos de robots de software, y es la tendencia que promueve añadir nuevos niveles de calidad y productividad en la estrategia de Transformación Digital de toda organización. RPA es una tecnología no intrusiva de acuerdo a lo que afirma el IRPA-AI (Insitute for Robotic Process Automation and Artificial Intelligence) [10, 11].

La automatización de procesos mediante robots es una tecnología que permite crear y utilizar una fuerza de trabajo virtual, flexible y escalable que interactúa con los

sistemas y el software de aplicaciones existentes (ERP, CRM, etc.) imitando la acción de un ser humano, lo que permite realizar tareas y procesos de una forma más eficiente, rápida, sin errores y con disponibilidad 24x7 [6, 7, 8].

Esta fuerza de trabajo virtual (robots), es capaz de asumir, de manera coordinada con las personas, aquellas tareas rutinarias, repetitivas y de poco valor que son necesarias dentro de una organización, de tal manera que se pueda proteger y potenciar el talento digital de los empleados con tareas de más valor dentro de la organización, aquellas que requieren: aportar creatividad, emitir juicios de valor, resolver problemas, gestionar excepciones, conocimiento humano, interpretación de datos, entre otras [9].

En relación con el costo, Capgemini [12] sugiere que una licencia de software RPA puede costar entre 1/3 y 1/5 del precio de un empleado de tiempo completo. Además, Lacity y Willcocks [13] argumentan que un robot puede realizar tareas estructuradas equivalentes a dos o cinco humanos. De todos modos, el uso de RPA por parte de las empresas proporciona las siguientes ventajas [8]:

- RPA es fácil de configurar, por lo que los desarrolladores no necesitan conocimientos de programación.
- El software RPA no es invasivo, se basa en sistemas existentes, sin necesidad de crear, reemplazar o desarrollar costosas plataformas.
- RPA es seguro para la empresa, RPA es una plataforma robusta que está diseñada para cumplir con los requisitos de IT de la empresa en términos de seguridad, escalabilidad, auditabilidad y gestión de cambios.

En este contexto, y como parte de un proyecto de tesis de postgrado, nos planteamos como objetivo proponer un modelo que haga posible de una manera ágil y práctica, la implementación de Automatización Robótica de Procesos en el dominio de las PyMEs, centrando el esfuerzo en tres líneas de investigación y desarrollo que observamos como primordiales: *identificar patrones en los procesos candidatos, especificar una metodología para seleccionar las herramientas RPA y definir un modelo de ciclo de vida para implementar RPA.*

El resto de este artículo describe estas líneas de trabajo en la Sección 2 y los resultados esperados/obtenidos en la Sección 3.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

En esta sección se describen las 3 líneas de investigación y desarrollo que se llevan a cabo como parte de esta tesis de postgrado.

2.1. Patrones en los procesos candidatos

Considerando la gran variedad de investigaciones presentes en la literatura tales como: [14-26], se advierte que existe una clara tendencia a que empresas de diferentes entornos comiencen a incluir software RPA en sus procesos tratando de: (1) aprovechar las ventajas que RPA proporciona con el objetivo de reducir costos y (2), mejorar la producción.

Actualmente no hay un estándar de referencia común obligatorio para identificar un proceso que pueda ser automatizado. De acuerdo a esto se planteó la pregunta de investigación *¿qué características nos permiten identificar los procesos candidatos a ser automatizados en el dominio de las pequeñas y medianas empresas (PyMEs)?*

2.2. Metodología para seleccionar las herramientas RPA

La compatibilidad del software RPA con los estándares y la arquitectura de IT, es un tema importante para la mayoría de las empresas. Además, el conocimiento de los sistemas de IT involucrados es esencial, de hecho, saber con qué frecuencia se actualizan los sistemas y se producen cambios, podría ser en factor de estabilidad. Si el sistema alcanzado por el proyecto RPA está experimentando cambios recurrentes en la interfaz de usuario, es esencial que el software RPA pueda manejar esto.

Costos de la inversión: licencias de software RPA, costos de implementación y costos de mantenimiento, también deben ser analizados.

La selección del proveedor adecuado e implementación a menudo puede ser tan difícil como seleccionar el software, ya que la mayoría de los proveedores de software RPA deben tener operaciones en el país y brindar la consultoría.

Todo esto sugiere la siguiente pregunta de investigación: *¿Cómo seleccionar la herramienta más adecuada para implementación de RPA en el dominio de las PyMEs?*

2.3. Modelo de ciclo de vida para implementar RPA.

Comenzar con RPA es más rápido que implementar sistemas de IT tradicionales.

RPA se caracteriza por implementarse con un enfoque ágil paso a paso para garantizar resultados rápidos en iniciativas de amplio alcance con alto potencial de beneficio. Las actividades de ciclo de vida para la implementación de RPA se adoptan del ciclo de vida de desarrollo de software. Esto nos lleva a realizar la tercera pregunta de investigación en este trabajo: *¿Qué acciones habría que ejecutar, en el ciclo de*

vida, para implementar RPA en el dominio de las PyMEs?

3. RESULTADOS OBTENIDOS/ ESPERADOS

Como resultado de esta línea se espera lograr probar que existen características recurrentes que permiten identificar los procesos candidatos a ser automatizados con un robot. Identificar el conjunto de herramientas adecuado para implementar RPA en el dominio de las PyMEs. Y verificar que exista un conjunto de acciones similares al ciclo de vida de desarrollo de software para la implementación de RPA en el dominio de las PyMEs.

Actualmente nos encontramos realizando la Revisión Sistemática de la Literatura (RSL) y avanzando en el estudio de cada línea de la investigación.

4. FORMACION DE RECURSOS HUMANOS

En esta línea de trabajo actualmente el primer autor está desarrollando su tesis correspondiente a la Maestría en Informática y Sistemas de la UNPA.

5. BIBLIOGRAFÍA

[1] Capek K. (1923). Rossum's Universal Robots. Playfair N, Selver P, trans. Landes WA, ed. New York: Doubleday.

[2] Robles, C. T. (2020, 3 junio). ¿Qué es un robot? Análisis jurídico comparado de las propuestas japonesas y europeas. Mirai. Estudios Japoneses. Recuperado 8 de septiembre de 2022, de <https://revistas.ucm.es/index.php/MIRA/article/view/67530>

[3] Córdoba Nieto, Ernesto (2006). Manufactura y automatización. Ingeniería e Investigación, 26(3),120-128. ISSN: 0120-5609. Recuperado el 8 de septiembre de 2022, de

<https://www.redalyc.org/articulo.oa?id=64326315>

[4] Deloitte. (2018, Junio). La era de la Automatización - Implementación de Robotics en las organizaciones. Deloitte. Recuperado el 8 de septiembre de 2022, de <https://www2.deloitte.com/content/dam/Deloitte/gt/Documents/technology/180605-Robotics.pdf>

[5] Siderska, J. (2021, Julio 18). "The Adoption of Robotic Process Automation Technology to Ensure Business Processes during the COVID-19 Pandemic". Ewa Ziembka. Recuperado el 8 de septiembre de 2022, de <https://doi.org/10.3390/su13148020>

[6] H. P. Fung (2014), "Criteria, use cases and effects of information technology process automation (ITPA)", Adv. Robot. Autom., vol. 3, no. 3, pp. 1–11.

[7] J. R. Slaby (2012), "Robotic automation emerges as a threat to traditional low-cost outsourcing", HfS Res., vol. 1, no. 1, p. 3.

[8] L. Willcocks and M. Lacity (2016), "A new approach to automating services", MIT Sloan Manage. Rev., vol. 58, no. 1, pp. 40–49.

[9] V. Kommera (2019), "Robotic process automation", Amer. J. Intell. Syst., vol. 9, no. 2, pp. 49–53.

[10] Modi, A., Makan, H., Mittal, S., & Khanna, R. (2021). "Defining Attended Robotic Process Automation (RPA) - What to Look for in an Enterprise-grade Solution". Everest Group. Recuperado el 8 de septiembre de 2022, de https://irpaai.com/wp-content/uploads/dlm_uploads/2022/06/Everest-Group_NICE-Defining-Attended-Robotic-Process-Automation.pdf

[11] C. Frank (2015), "Introduction to robotic process automation", in Proc. Inst. Robot. Process Automat., 2015, p. 35.

- [12] C. Capgemini (2017), "Robotic process automation - Robots conquer business processes in back offices", Capgemini Consulting Capgemini Bus. Services, Paris, France, Tech. Rep. 2016.
- [13] M. Lacity and L. Willcocks (2015), "What Knowledge Workers Stand to Gain from Automation", *Harvard Bus. Rev.*, vol. 19, no. 6, pp. 1–7. Recuperado el 8 de septiembre de 2022, de <https://hbr.org/2015/06/what-knowledge-workers-stand-to-gain-from-automation>
- [14] S. Aguirre and A. Rodriguez (2017), "Automation of a business process using Robotic Process Automation (RPA): A case study", in *Proc. Appl. Comput. Sci. Eng.*, 4th Workshop Eng. Appl. WEA, pp. 65–71.
- [15] A. Asatiani and E. Penttinen (Nov. 2016), "Turning robotic process automation into commercial success – case OpusCapita", *J. Inf. Technol. Teach. Cases*, vol. 6, no. 2, pp. 67–74.
- [16] M. Buckingham and A. Goodall (2015), "How deloitte killed forced rankings: Interaction", *Harvard Bus. Rev.*, vol. 93, no. 6, pp. 18–19.
- [17] B. Bygstad (Jun. 2017), "Generative innovation: A comparison of lightweight and heavyweight IT", *J. Inf. Technol.*, vol. 32, no. 2, pp. 180–193.
- [18] J. Hultin, C. Trudell, A. Vashistha, and T. Glover, (2017), "Implications of technology on the future workforce", *Defense Bus. Board Washington United States, Washington, DC, USA*, Tech. Rep. AD1040713.
- [19] L. P. Lacity & M. Willcocks (2016), "Robotic process automation at telefónica O2", *MIS Quart. Executive*, vol. 15, no. 1, pp. 21–35.
- [20] C. Lamberton, D. Brigo, & D. Hoy (2017), "Impact of Robotics, RPA and AI on the insurance industry: Challenges and opportunities", *J. Financial Perspective, Insurance*, vol. 4, no. 1, pp. 8–20.
- [21] C. Le Clair, A. Cullen, and M. King (2017), "The forrester wave robotic process automation, Q1 2017", *Forrester Res.*, Cambridge, MA, USA, Tech. Rep. Q1 2017.
- [22] P. Mijovic, E. Giagloglou, P. Todorovic, I. Macužic, B. Jeremic, & I. Gligorijevic (2014), "A tool for neuroergonomic study of repetitive operational tasks", in *Proc. Eur. Conf. Cognit. Ergonom. (ECCE)*, pp. 1–2.
- [23] V. Naik, P. Garbacki, and A. Mohindra (2006), "Architecture for service request driven solution delivery using grid systems", in *Proc. IEEE Int. Conf. Services Comput. (SCC)*, pp. 414–422.
- [24] F. C. Varghese (2017), "The impact of automation in IT industry: Evidences from India", *Int. J. Eng. Sci.*, vol. 7, no. 3, pp. 5000–5004.
- [25] L. Willcocks & A. Craig (2015), "Robotic process automation at Xchanging", in *Proc. Outsourcing Unit Work. Res. Paper Ser.*, pp. 1–26.
- [26] L. Willcocks, M. Lacity, & A. Craig (2017), "Robotic process automation: Strategic transformation lever for global business services?" *J. Inf. Technol. Teaching Cases*, vol. 7, no. 1, pp. 17–28.

CLASIFICACIÓN DE PARTÍCULAS DE ARENA A TRAVÉS DE REDES NEURONALES CONVOLUCIONALES

Mg. Ing. Carlos Gustavo Rodriguez Medina ¹, Dr. Ing. Oscar Daniel Chuk ²,
Lic. Adriana Luna, Ing. Regina Bertero.

Instituto de Investigaciones Mineras / Facultad de Ingeniería /
Universidad Nacional de San Juan

Av. Libertador Gral. San Martin 1109 oeste. San Juan
0264-4211700 (int. 285 ¹, int. 389 ^{1,2})

grodriguez@unsj.edu.ar ¹, dchuk@unsj.edu.ar ²

RESUMEN

Se presenta una alternativa para clasificar partículas de arenas, en cuanto a la característica de redondez, utilizando redes neuronales convolucionales. Se ha seleccionado este tipo de redes neuronales, por la atractiva posibilidad de poder operar directamente sobre las imágenes digitales de partículas de arena, sin la necesidad de tener que extraer previamente características o medidas alternativas de las imágenes. Esta propuesta resulta ser una opción diferente a la utilización de redes neuronales que no aplican convolución, y también a técnicas de visión artificial (que emplean cálculos, métodos y medidas alternativas para estimar valores de redondez).

La necesidad de poder clasificar las partículas en cuanto a su redondez, radica en que estas arenas se utilizan en procesos de fracturación hidráulica para la extracción de hidrocarburos de manera no convencional. Estos yacimientos no convencionales requieren de la inyección de arenas que actúan como apuntalantes de la fractura. La calidad de las arenas para cumplir dicha función se evalúa por la norma API19C, que establece medidas geométricas –esfericidad y redondez– más un porcentaje de fractura cuando la muestra es sometida a compresión. Todas estas mediciones deben encontrarse dentro de valores límites para asegurar la calidad de la arena.

El método indicado por la norma (tradicional no automatizado) para la medición de redondez tiene una alta dependencia del criterio del observador o laboratorista, que analiza una muestra mediante un microscopio eligiendo un número muy reducido de partículas al azar, y comparándolas visualmente con formas gráficas establecidas en una cartilla.

Palabras clave: Redes neuronales convolucionales, Partículas de arena, Redondez, Fracturación hidráulica.

CONTEXTO

Este trabajo presenta de manera resumida los resultados del Proyecto de Investigación Científica (PIC – UNSJ) “*Clasificación de partículas de arena de fracturación mediante redes neuronales convolucionales*”, ejecutado durante el periodo 2020 – 2022, financiado por la Universidad Nacional de San Juan.

Dicho proyecto se inscribe dentro de la línea de investigación que lleva adelante un grupo de trabajo en la temática de Inteligencia y Visión Artificial orientado al control automático de procesos mineros, en el Instituto de Investigaciones Mineras de la Facultad de Ingeniería de la UNSJ.

1. INTRODUCCIÓN

En la industria de extracción de petróleo de manera no convencional, y con la finalidad de garantizar la capacidad de las arenas con destino a los procesos de fracturación hidráulica (*Fracking*), para actuar como apuntalantes o propantes, se reconocen como base las normas de referencia del *American Petroleum Institute* (API) 19C [1, 2]. Esta norma recomienda una serie de ensayos a realizar sobre las arenas con el fin de conocer la calidad del material a utilizarse en el proceso. Entre los ensayos a aplicarse se encuentran los de redondez y esfericidad, entre otros.

Tales ensayos padecen de una fuerte dependencia del criterio del observador, dado que el procedimiento se aplica mediante la utilización de un microscopio y las mediciones de redondez y esfericidad se realizan por comparación respecto a tablas que contienen una serie de formas provistas por la norma, mediante unas cartas gráficas conocidas como tablas o cartillas de *Krumbein y Sloss* [3]. Esto deriva en resultados sustancialmente disímiles que invalidan el procedimiento, aunque es el que se sigue usando por norma.

El uso de técnicas de Visión Artificial se presenta como la alternativa tecnológica que puede permitir obtener medidas objetivas no dependientes del criterio de un operador [1]. En este sentido el grupo de investigación ha elaborado varios algoritmos con diversos objetivos aplicados al procesamiento digital de partículas que van desde la segmentación hasta el cálculo de medidas geométricas, tales como el de esfericidad, redondez, tamaño, distribución granulométrica, etc. [1, 4]

Las partículas pueden presentar entre ellas distintas variaciones, pudiendo ser algunas más redondeadas, o con más puntas, más o menos achatadas, más o menos alargadas, de diferentes tamaños, etc.

Dado que esta situación que se presenta, es un problema relacionado con la forma de las partículas, se sabe que existen herramientas de la inteligencia artificial orientadas a resolver problemas de reconocimiento de

formas, patrones y su clasificación; éstas son las redes neuronales artificiales [4, 9, 10].

Además, hay un tipo particular de redes neuronales que permiten operar directamente sobre un conjunto de imágenes, sin necesidad de tener que realizar ningún pre-procesado para la extracción de las características de interés, estas son las redes neuronales convolucionales [5, 8].

Las redes neuronales convolucionales son especialmente útiles para localizar patrones en imágenes con el objetivo de reconocer objetos, formas, etc. Se puede decir que aprenden directamente a partir de los datos de las imágenes [4].

Al igual que otras redes neuronales, una red neuronal convolucional se compone de una capa de entrada, una capa de salida y muchas capas ocultas intermedias. A su vez, las capas intermedias ocultas se organizan de forma alternada con capas convolucionales y de pooling, finalizando con una serie de capas de neuronas completamente conectadas (*Full connected*) [5].

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el Laboratorio de Control Automático del Instituto de Investigaciones Mineras (Facultad de Ingeniería de la Universidad Nacional de San Juan) se realiza investigación y desarrollo respecto a la automatización de procesos de molienda de minerales aproximadamente desde el año 2000. Desde hace algunos años se viene desarrollando el Área de Procesamiento de Imágenes.

Actualmente se está trabajando en la caracterización de arenas especiales (calidad, forma, color, tamaños, clasificación del tipo de material, etc.) mediante el procesamiento de imágenes para ser utilizadas en la extracción de petróleo mediante el método de *Fraking*.

3. RESULTADOS OBTENIDOS

Dado a que las redes neuronales requieren de un proceso de entrenamiento previo a su utilización, se debe contar con un conjunto de

datos lo más representativo posible de las imágenes con las cuales se necesita trabajar; para este caso se trata de imágenes de partículas de arena. A la vez, este conjunto de imágenes se divide en un conjunto de entrenamiento y otro de validación para poder comprobar el desempeño de la red entrenada. En la figura 1 se muestra un ejemplo de imágenes en blanco y negro (del tipo binarias) de partículas de arena. Para este caso, cuando se utilizan imágenes capturadas mediante un microscopio, las mismas deben convertirse de RGB a binario, para luego ser ingresadas a la red neuronal.



Figura 1. Imágenes en blanco y negro de partículas de arena.

Para poder entrenar una red neuronal y que luego la misma logre determinado nivel de acierto en la predicción de la variable de interés, y con mayor razón si se trata de una red del tipo convolucional, se necesitan grandes cantidades de imágenes para llevar a cabo dicho proceso adecuadamente. Por tal motivo, y debido a que no se disponía de un banco de imágenes de partículas de arena con tal característica, es que se trabajó en la elaboración de un banco de imágenes de partículas realizadas mediante la utilización de funciones matemáticas para graficar elementos geométricos básicos y la aplicación de procesamiento digital computacional [6]. Ha tales imágenes se las ha denominado como partículas “artificiales” de arena [7], la figura 1 corresponde a éstas.

Además, se debe tener en cuenta que para este tipo de redes se debe contar con un grupo de

etiquetas o valores de redondeces que se correspondan con cada imagen del conjunto de datos. Con estos pares de datos (imágenes y etiquetas) es con lo que se nutre la red (entradas) para llevar a cabo el aprendizaje.

El conjunto inicial de imágenes (artificiales) para entrenamiento y validación, obtenido en un primer momento, era aproximadamente de 150 imágenes de partículas de arena por cada etiqueta (unas 600 en total), el cual resultaba ser un número importante comparado con el que se contaba si solo se disponía la tabla de *Krumbein y Sloss* [3], obsérvese en la figura 2. Pero ya se mencionó que para el entrenamiento de una red neuronal convolucional se requieran grandes cantidades de imágenes, por lo cual seguía siendo insuficiente dicha cantidad.

Por lo tanto, aplicando procesamiento digital de imágenes y mediante funciones de rotación y también de espejado se obtuvieron más de 50.000 imágenes. De esta manera, el conjunto que se obtuvo fue el adecuado para poder comenzar a obtener algunos resultados de interés una vez llevado a cabo el entrenamiento.

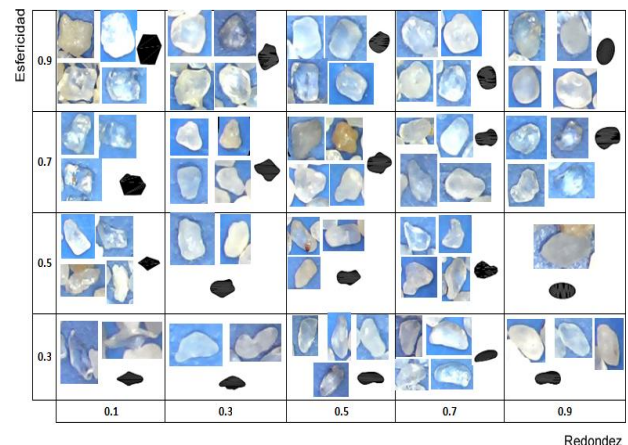


Figura 2. Cartilla de Krumbein y Sloss.

De acuerdo a las tablas de *Krumbein y Sloss* [3], la redondez puede tomar valores entre 0.1 y 0.9, figura 2. Debido a los importantes recursos de hardware que requiere una red neuronal convolucional para llevar a cabo el proceso de entrenamiento, y más cuando se necesita disponer de varias salidas, es que en este trabajo se acota a entrenar la red para clasificar las partículas con redondeces entre

0.1 y 0.4. El tiempo insumido en el entrenamiento lógicamente también es mayor cuando no se cuenta con el procesamiento adicional de una placa grafica con GPU dedicada.

Para el entrenamiento y validación de la red, se ha utilizado un equipo informático con procesador Intel i5 de 4 núcleos, 8GB de RAM y no se dispone de GPU dedicada.

Se ha utilizado el software matemático Matlab para la implementación de los algoritmos de las redes neuronales convolucionales, para el proceso de entrenamiento y pruebas, como así también para la generación de las partículas artificiales de arena, procesamiento de imágenes y armado de los distintos conjuntos de entrenamiento y validación.

Red neuronal

La red neuronal convolucional desarrollada se conforma de 5 capas de Convolución y Pooling alternadas, para luego unirse a una capa Completamente conectada (*Full-connected*). La salida se compone de los posibles 4 valores de redondeces antes indicados (0.1, 0.2, 0.3 y 0.4).

En la tabla I se resume el conjunto de imágenes de partículas de arena elaborado para estas pruebas, en la que se presentan las etiquetas que referencian a los valores de redondez y a la cantidad de imágenes que contiene cada una.

Etiqueta (redondez)	Cant. Imágenes
1 (0.1)	13.320
2 (0.2)	14.400
3 (0.3)	14.580
4 (0.4)	14.760

Tabla I. Conjunto de datos.

El conjunto de entrenamiento utilizado para la prueba que se presenta es de 8.000 imágenes para cada valor de redondez, por lo que por ej. para el caso de la primera etiqueta de redondez se tiene aproximadamente un 60% del conjunto de datos subdividido para el proceso de entrenamiento, y el resto para validar dicho entrenamiento.

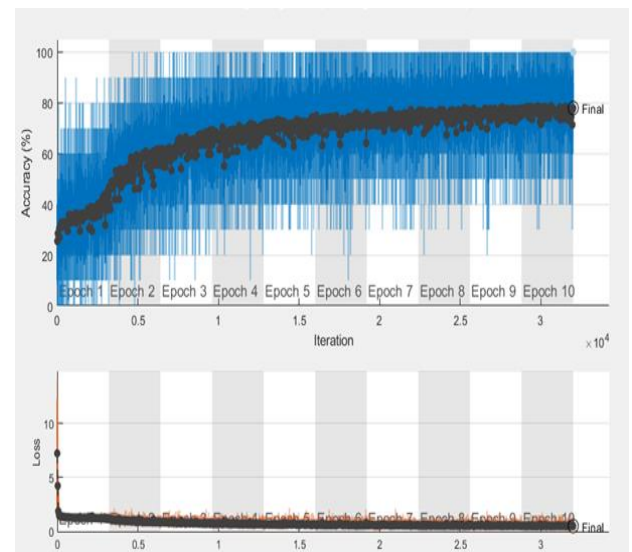


Figura 3. Resultado de entrenamiento de la red neuronal convolucional.

El entrenamiento realizado para este ejemplo ha insumido un tiempo de 908 min (poco más de 15hs). El mismo logró una precisión del 77.98%. En la figura 3 se presenta la pantalla de resultado arrojada por Matlab.

También se aprecia que la red neuronal a medida que transcurre el tiempo de entrenamiento va aumentando su precisión, hasta casi alcanzar prácticamente el 80%.

True Class	1	4749	479	65	27
	2	870	4636	758	136
	3	80	503	4356	1641
	4	7	32	298	6423
		1	2	3	4
		Predicted Class			

Figura 4. Matriz de Confusión de la prueba realizada a la red neuronal convolucional.

En la figura 4 se presenta el resultado de testear el desempeño de la red neuronal, a través de la Matriz de Confusión.

De acuerdo a la Tabla I, y teniéndose en cuenta que para el entrenamiento se utilizaron 8.000 imágenes de partículas de arena de cada

etiqueta, entonces en la Tabla II se indican las cantidades de imágenes utilizadas para la prueba realizada a la red neuronal y así poder comparar y analizar junto a la Matriz de Confusión.

Etiqueta (redondez)	Cant. imágenes
1 (0.1)	5.320
2 (0.2)	6.400
3 (0.3)	6.580
4 (0.4)	6.760

Tabla II. Datos utilizados para la prueba.

De la figura 4, se aprecia que cuando se ingresan como entradas a la red neuronal imágenes de partículas de arena con una redondez de 0.1 (etiqueta 1), la red clasificó adecuadamente (eje indicado como *predicted class*) con redondez 0.1 (etiqueta 1) a 4749 imágenes (89.2%). También se puede observar que 479 imágenes (9%) fueron identificadas con una redondez de 0.2 (etiqueta 2) erróneamente. Así también, alejándose un poco más del resultado correcto, se observa que a 65 imágenes (1.2%) la red las identificó con la etiqueta 3 erróneamente (redondez 0.3). Finalmente, solo 27 imágenes (0.5%) identificó erróneamente con una redondez de 0.4 (etiqueta 4). Obsérvese que los elementos de la diagonal principal de dicha matriz se encuentran coloreados en azul, indicando las cantidades de aciertos obtenidos por la red neuronal. Por lo tanto, los elementos fuera de la diagonal corresponden a valores erróneos (predicciones incorrectas de la red). De esta manera se puede recorrer toda la matriz para analizar el desempeño de la red neuronal, y apreciar la cantidad de imágenes con determinado valor de redondez ingresadas como entradas, y observar los valores predichos por la red y obtenidos a la salida.

Si se calcula el porcentaje promedio de errores cometidos por la red (analizando la matriz de confusión completamente), para este caso se tiene que es del 19,27%, el cual resulta ser razonable teniendo en cuenta el porcentaje de aciertos (*Accuracy*) obtenido durante el entrenamiento, indicado en párrafos anteriores.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por Ingenieros Electrónicos, Electromecánicos, en Minas, Metalurgistas y Geólogos, con formación diversa en posgrados, tal como Doctorado en sistemas de control, Maestría en informática, Especialización en gestión y vinculación tecnológica, etc.

Actualmente, uno de los integrantes del equipo de trabajo se encuentra desarrollando su Tesis de Doctorado en Ingeniería de Procesamiento de Minerales, siendo su director de Tesis otro de los integrantes del grupo de investigación.

5. BIBLIOGRAFIA

- [1] Rodriguez Medina, C.G.; Chuk, O.D.; Bertero R.; Luna, A.; Núñez, E.; Trigo, P. Visión artificial aplicada a la determinación de propiedades geométricas de arenas de fracturación para el análisis de calidad. Informe de Final de Proyecto PIC CICITCA. Universidad Nacional de San Juan. 2019.
- [2] American Petroleum Institute API 19C, Second Edition: Measurement of Properties of Proppants Used in Hydraulic Fracturing and Gravel-Packing Operations; API Publishing Services: Washington, D.C., 2008
- [3] Krumbein, W.C.; Sloss, L.L. Stratigraphy and Sedimentation; 2nd ed.; W.H. Freeman: San Francisco, 1963;
- [4] Rodriguez Medina, C. G.; Chuk, O.D.; Bertero, R. Luna, A.; Núñez, E.. Redes neuronales para la clasificación de partículas de arena. WICC 2021. XXIII Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional de Chilecito. Pág. 149 – 154. 2021. ISBN 978-987-24611-3-3.
- [5] Mathworks. Redes neuronales convolucionales. 1994-2022. <https://la.mathworks.com/discovery/convolutional-neural-network-matlab.html>
- [6] Charoenpong, T.; Chamnongthai, K.; Kamhom, P.; Krairiksh, M. Volume Measurement of Mango by Using 2D Ellipse Model. In Proceedings of the 2004 IEEE

- International Conference on Industrial Technology, 2004. IEEE ICIT '04.; December 2004; Vol. 3, pp. 1438-1441.
- [7] Rodriguez Medina, C.G.; Chuk, O.D.; Bertero, R. Luna, A.; Núñez, E. Generación gráfica de partículas artificiales de arena para su utilización en redes neuronales. WICC 2020. XXII Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional de la Patagonia Austral. Pág. 307 – 312. 2020. ISBN 978-987-3714-82-5.
- [8] Liang, Z.; Nie, Z.; An, A.; Gong, J.; Wang, X. A Particle Shape Extraction and Evaluation Method Using a Deep Convolutional Neural Network and Digital Image Processing. *Powder Technology* Pag. 156-170. 2019. doi:10.1016/j.powtec.2019.05.025.
- [9] Jain, A. K. Statistical Pattern Recognition: a review. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol 22, Nro. 1. 2000.
- [10] Haykin S. Neural Networks - A comprehensive foundation, IEEE Press - Macmillan College Publishing Company, Inc. 1994.
- [11] Hryciw, R.D.; Zheng, J.; Shetler, K. Particle Roundness and Sphericity from Images of Assemblies by Chart Estimates and Computer Methods. *Journal of Geotechnical and Geoenvironmental Engineering* 2016, 142, 04016038, doi:10.1061/(ASCE)GT.1943-5606.0001485.
- [12] Hernandez, E.M.; Chavez, G.M.; Hernandez, J.V. Roundness Estimation of Sedimentary Rocks Using Elliptic Fourier and Deep Neural Networks. In Proceedings of the 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC); IEEE: Ixtapa, Mexico, November 4 2020; pp. 1–5.
- [13] Wadell, H. Volume, Shape, and Roundness of Rock Particles. *Jour. Geol.* 1932, XL, N°5, 443–451.

Resolución de problemas regionales NP-duros usando técnicas de optimización

Carolina Salto^{2,3}, Gabriela Minetti², Hugo Alfonso², Carlos Bermúdez², M.Juliana Dielschneider Del Bono², Luciano Bessoni², Javier Vargas², José Luis Hernandez¹, Mercedes Carnero¹

¹ Facultad de Ingeniería - Universidad Nacional de Río Cuarto

² Laboratorio de Investigación en Sistemas Inteligentes (LISI)

Facultad de Ingeniería - Universidad Nacional de La Pampa – ³ CONICET

e-mail: {saltoc, minetti, alfonsoh, bermudezc, juliana, jvargas}@ing.unlpam.edu.ar
{mcarnero,jlh}@ing.unrc.edu.ar

Resumen

En este artículo se introducen dos de las líneas de investigación que se desarrollan en el Laboratorio de Sistemas Inteligentes (LISI), relacionadas con problemas de optimización NP-duros de carácter regional: diseño de redes distribución de agua y logística de distribución de carga y paquetería.

En la actualidad, las empresas públicas de suministro de agua proporcionan más del 90% del abastecimiento mundial, por lo que un sistema seguro de distribución de agua es fundamental para cualquier ciudad. La importancia, el enorme costo de capital del sistema y el creciente tamaño de las ciudades conducen a la optimización de la red de distribución de agua. En esta línea de investigación, se proponen y analizan dos algoritmos metaheurísticos para optimizar el diseño de la red de agua de un nuevo barrio de la ciudad de General Pico, con el objetivo de minimizar el costo total de la inversión.

La logística remite a flujos de materiales y de información; a lugares de manipulación, depósito y transformación de las mercancías; a redes y nodos de circulación; y a tiempos de movimiento y de espera que responden a aspectos materiales (las infraestructuras, los transportes y las cargas) y también a aspectos funcionales (los servicios, las normativas y regulaciones). En suma, la logística implica un uso del territorio en el tiempo, una convergencia espacio-temporal, una organización y sincronización de flujos a través de estrategias sobre los nodos y las redes. La Provincia de La Pampa cuenta con gran potencial para constituirse

en un centro logístico de distribución a nivel nacional, por ello se considera importante que esté consolidada en la búsqueda de propuestas innovadoras para introducir al sector. En consecuencia, la segunda línea de investigación se enfoca en modelar las características de la logística de distribución de cargas y paquetería en la región con el objetivo de minimizar los costos operativos y maximizar la calidad del servicio.

Palabras claves: Optimización, Logística, Inteligencia Artificial, Diseño de redes

CONTEXTO

Desde su creación en 1998, el Laboratorio de Investigación de Sistemas Inteligentes (LISI), perteneciente a la Facultad de Ingeniería (FI) de la Universidad Nacional de La Pampa (UNLPam), se ha abocado al estudio de algoritmos cada vez más eficientes para la solución de problemas complejos, tanto de optimización como de diseño. Actualmente, dos de sus líneas de investigación están relacionadas con la resolución de problemas de optimización NP-duros de la región de influencia de dicho laboratorio: diseño óptimo de redes de distribución de agua y optimización de la logística de distribución de cargas y paquetería. Estas líneas de investigación se desarrollan en el marco de dos proyectos de investigación: "Big data optimization con algoritmos metaheurísticos utilizando frameworks de computación distribuida", acreditado por la FI, y "Optimización de la logística de distribución utilizando técnicas de la Inteligencia Artificial" (POIRE), acreditado por la UNLPam.

En ambas líneas, el objetivo principal consiste en obtener algoritmos inteligentes que, además de brindar soluciones de calidad a las problemáticas planteadas con poco esfuerzo computacional, requieran bajo costo de implementación. Cabe destacar que desde hace varios años, los integrantes de este laboratorio mantienen una importante vinculación con investigadores de universidades argentinas (Universidad Nacional de San Luis y Universidad Nacional de Río Cuarto-UNRC) y de la Universidad de Málaga (España), con quienes se realizan trabajos conjuntos. En particular este trabajo se desarrolló con investigadores de la UNRC.

La primera línea de investigación, diseño de redes de agua, se enmarca en la ciudad de General Pico, donde el agua potable proviene de unas 100 perforaciones para extracción del acuífero subterráneo que se extiende desde General Pico a la localidad de Dorila, uno de los pocos recursos hídricos accesibles y aptos para el consumo. Cada pozo extrae un promedio de 9000 litros/hora. El agua extraída se transporta hasta una cisterna principal subterránea, que junto a otras cinco cisternas abastecen a toda la ciudad. Todo esto contribuye a que la optimización del diseño de redes de distribución de agua para nuevos barrios sea un problema de alta complejidad.

La segunda línea de investigación, logística de distribución de carga, se encuadra en la Provincia de La Pampa situada en el centro geográfico de la República Argentina y considerada la puerta de ingreso a la Región Patagónica. Por ende, cuenta con gran potencial para constituirse en un centro logístico de distribución para el territorio argentino. Dado que los costos de transporte se hallan entre un tercio y dos tercios de los costos logísticos totales, mejorar la eficiencia mediante la máxima utilización del equipo de transporte y de su personal es una preocupación importante. Un problema frecuente en la toma de decisiones es reducir los costos de transporte y mejorar el servicio al cliente encontrando las mejores rutas.

I. INTRODUCCIÓN

El diseño óptimo de redes de distribución de agua potable y la optimización de la logística de distribución de cargas y paquetería son problemas de decisión complejos del mundo real que, mode-

lados como problemas de optimización, pertenecen a la clase NP-Duros.

El objetivo del problema de optimización del diseño de redes de distribución de agua (Water Distribution Network Design Optimization - WDND) es minimizar el costo total de inversión (Total Investment Cost - TIC) de una red de distribución de agua. Un grafo conexo conformado por un conjunto de nodos $N = \{n_1, n_2, \dots\}$, un conjunto de tuberías $P = \{p_1, p_2, \dots\}$, un conjunto de subredes o loops internos $L = \{l_1, l_2, \dots\}$, y un conjunto de tipos de tuberías disponibles en el mercado $T = \{t_1, t_2, \dots\}$ representan la red de agua. A partir de este modelo, se minimiza el costo total de la inversión, siguiendo la fórmula mostrada en la Ecuación 1.

$$\min TIC = \sum_{p \in P} \sum_{t \in T} L_p IC_t x_{p,t} \quad (1)$$

donde IC_t es el costo de un tubo p del tipo t , L_p es la longitud del tubo y $x_{p,t}$ en una variable binaria que indica si el tubo p es del tipo t o no. Función restringida por las leyes físicas de conservación de masa y energía, la demanda de presión mínima en cada nodo y la máxima velocidad en la tubería, en cada momento $\tau \in \mathcal{T}$.

En el campo de la investigación operativa [1], la optimización de la logística de distribución de cargas y paquetería se agrupa en la clase de problemas del enrutamiento vehicular (VRP, sus siglas en Inglés) [2]. El VRP clásico consiste en determinar las rutas que debe tomar una determinada flota de vehículos para recolectar y/o distribuir artículos en ubicaciones conocidas de clientes. Cada artículo suele tener asociado un determinado tamaño o peso. La cantidad total (en términos de tamaño o peso) de los artículos recolectados por un vehículo no puede exceder su capacidad. En esta versión de VRP, se supone que todos los datos (demandas del cliente, tiempos de viaje, ventanas de tiempo, etc.) se conocen de antemano. El responsable de la toma de decisiones, tanto en grandes compañías como en pequeñas empresas, debe planificar con anticipación las rutas de los vehículos para satisfacer las demandas de los clientes a un costo de viaje mínimo. El costo del viaje puede incluir, entre otros: tiempo de viaje, horas hombre, viáticos, combustible y peajes [3], [4]. Desafortunadamente, el VRP es fuertemente NP-duro incluso para un

solo objetivo, ya que el problema del viajante de comercio (TSP, por sus siglas en inglés) [5] puede reducirse polinomialmente a él [6].

En el VRP clásico se establecen rutas de menor costo desde un depósito central a un conjunto de puntos dispersos geográficamente (clientes, tiendas, escuelas, ciudades, almacenes, etc.) con diversas demandas. Cada cliente debe ser atendido exactamente una vez por un solo vehículo, y cada vehículo tiene una capacidad limitada. Una extensión más realista es el Problema de generación de rutas para vehículos con ventanas de tiempos (VRPTW), que cuenta con muchas aplicaciones en el mundo real; aquí se asocia una ventana de tiempo a cada cliente. Es decir, además de la restricción de capacidad del vehículo, cada cliente proporciona un marco de tiempo dentro del cual se debe completar un servicio o tarea en particular, como cargar o descargar un vehículo. Un vehículo puede llegar temprano, pero debe esperar hasta que sea posible la hora de inicio del servicio. El objetivo del VRPTW es minimizar la cantidad de vehículos y la distancia total recorrida para atender a los clientes sin violar las limitaciones de capacidad y ventana de tiempo.

La resolución de estos problemas de alta complejidad requiere de métodos eficaces, confiables y fáciles de usar que, además de tener en cuenta los costos de capital y operativos, consideren el rendimiento, la confiabilidad hidráulica y la gestión competente de la energía en el caso de WDND y la satisfacción del cliente y la fiabilidad del sistema para el VRPTW. En consecuencia, las técnicas de solución capaces de producir soluciones de alta calidad en un tiempo limitado, como las heurísticas y las metaheurísticas [7], son de primordial importancia. Estas técnicas, que pertenecen a la clase de algoritmos de búsqueda estocásticos, reducen de forma significativa la complejidad temporal del proceso de búsqueda. Por otro lado, la paralelización de la búsqueda, además de disminuir los tiempos, permite aumentar la calidad de las soluciones halladas. Esto implica el estudio de diferentes plataformas actuales de hardware (procesadores multicore, unidades de procesamiento gráficas, ambientes cloud) y de software (Hadoop, Spark, MPI, High Processing Computing en general).

II. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

En esta sección se presenta el tratamiento dado a los problemas, WDND y VRP, abordados en cada una de las líneas de investigación enunciadas anteriormente.

A. Diseño de redes de distribución de agua

Actualmente, las formulaciones de WDND incluyen la extensión a múltiples períodos, es decir, patrones de demanda que varían en el tiempo, que son formulaciones de problemas realistas, pero también más complejas que las tradicionales. Algunos trabajos expresaron el problema de diseño como un problema de optimización multiobjetivo y aplicaron un algoritmo evolutivo multiobjetivo [8]. También se desarrolló un algoritmo genético para resolver seis pequeñas redes [9], considerando la restricción de velocidad del agua que fluye a través de las tuberías. En [10] también consideraron esta restricción, pero los autores usaron programación matemática en redes más grandes y más cercanas a la realidad. Los autores en [11], [12] utilizaron otras metaheurísticas para abordar formulaciones de WDND más complejas. En la búsqueda local iterativa (ILS) presentada en [13] se consideró que cada nodo de demanda tiene un patrón de demanda de agua de 24 horas, incluyendo una nueva restricción relacionada con el límite de la velocidad máxima del agua a través de las tuberías.

El crecimiento de las ciudades también requiere de este complejo proceso de diseño, como es el caso de General Pico (La Pampa, Argentina). En particular, se necesita optimizar un WDND independiente de un nuevo barrio de cinco km^2 , sujeto a demandas multiperiodo, restricciones hidráulicas, entre otras restricciones. Una cooperativa pública está a cargo de distribuir este elemento esencial en esta ciudad, lo que requiere un sistema de optimización para determinar la mejor solución de ingeniería que cumpla con los criterios de diseño establecidos y minimice los costos de capital. Nuestro objetivo de investigación es desarrollar una solución óptima basada en algoritmos de optimización de última generación, que pretenda apoyar la toma de decisiones para diseñar, planificar y gestionar sistemas de agua complejos.

Proponemos y comparamos metaheurísticas basadas en *Tabu Search* (TS) y *Simulated Annealing* (SA) para resolver el problema WDND, op-

timizando los diámetros de las tuberías y minimizando el costo total de inversión. El método basado en TS, llamado SOTS (*Strategic Oscillations Tabu Search*), adapta la técnica constructivo-destructiva llamada Oscilación Estratégica [14] para mejorar las capacidades de intensificación y diversificación de la búsqueda. Se utiliza un algoritmo de optimización híbrido basado en *Simulated Annealing* (HSA) [15] mejorado con un procedimiento de búsqueda local basado en GRASP [13] para minimizar el costo de la red de agua. Probamos el rendimiento de SOTS y HSA con una red de distribución de tamaño medio del mundo real. Las principales contribuciones de este trabajo son las siguientes: *i*) La adaptación de la técnica SO ad hoc al problema WDND (fases constructiva y destructiva de SOTS), y *ii*) Resolución de una red de distribución del mundo real.

B. Logística de la distribución de cargas y paquetería

Para abordar este problema de optimización relacionado con la logística de distribución de cargas y paquetería, se propone el desarrollo de software logístico, que incorpore herramientas basadas en inteligencia artificial, como soporte para la toma de decisiones a nivel gerencial y asistido por herramientas que permitan evaluar la incidencia de la matriz de costos y buscar el equilibrio del sistema. Debe ser capaz de diseñar rutas de distribución eficientes para incrementar la eficiencia operacional, reducir costos y aumentar la satisfacción de los clientes. Además, se pretende determinar los factores de costos relacionados a cada ruta planificada, su clasificación en cuanto a variabilidad y la asignación de un valor, para estudiar la incidencia de cada uno de ellos en el costo total de dicha ruta y contar con información importante para la toma de decisiones. El VRPTW ha recibido mucha atención debido a la aplicabilidad de las restricciones de ventana de tiempo en situaciones del mundo real. Se trata de un problema NP-Completo y las instancias con 100 clientes o más son muy difíciles de resolver de manera óptima. Representamos el VRPTW como un problema multiobjetivo, en el que las dos dimensiones objetivas son la cantidad de vehículos y el costo total (distancia). Una ventaja de este enfoque es que no es necesario derivar pesos para

una fórmula de puntuación de suma ponderada. Esto evita la introducción de un sesgo de solución hacia cualquiera de las dos dimensiones.

Muchos investigadores han propuesto técnicas exactas para resolver este problema, pero ha sido Kohl [16] quien ha introducido uno de los métodos exactos más eficientes para el VRPTW; al resolver varias instancias de hasta 100 clientes. Sin embargo, el costo computacional de estos métodos es muy alto, siendo los métodos aproximados, como las metaheurísticas, los que han resuelto este inconveniente (algoritmos genéticos, estrategias evolutivas, enfriamiento simulado, búsqueda tabú, entre otras) [17], [18], [19]). Debido a que resolver eficientemente las instancias de VRPTW con más de 100 clientes sigue siendo un desafío, en nuestra línea de investigación se propone combinar las ventajas de las metaheurísticas multiobjetivo con heurísticas exactas para obtener soluciones de calidad con bajo costo computacional.

III. RESULTADOS ENCONTRADOS Y ESPERADOS

En esta sección se muestran y analizan, en primer lugar, los resultados hallados a partir de la concreción de los objetivos propuestos en la primera línea de investigación. En segundo lugar, se describe qué resultados se esperan obtener en la línea de trabajo orientada a VRP.

A. Diseño de redes de distribución de agua

A continuación se analiza el desempeño de los algoritmos propuestos para optimizar el WDND de un nuevo barrio en la ciudad de General Pico: SOTS y HSA. El algoritmo SOTS puede iniciar la búsqueda de soluciones con costos de inversión totales muy altos y alcanza buenas soluciones en las primeras etapas de la búsqueda, con un esfuerzo computacional relativamente bajo en comparación con la heurística HSA. Esto se debe principalmente a la fase de destrucción inteligente implementada. Sin embargo, a medida que avanza la búsqueda, la disminución de los costos totales de inversión encontrados se hace menor, la capacidad exploratoria disminuye y aumenta la cantidad de veces que se ejecuta la función de evaluación. Un comportamiento similar se puede observar en HSA. Las herramientas disponibles para evitar mínimos locales son diferentes para los dos métodos: memoria a corto plazo y memoria a largo plazo en el caso de

SOTS, y aceptación de peores soluciones cuando la temperatura es alta en el caso de HSA. Considerando que el espacio de posibles soluciones para el problema dado es del orden de 10^{256} , la posibilidad de explorar diferentes regiones del problema utilizando una sola técnica de búsqueda es limitada, y los esfuerzos para diseñar algoritmos para el WDND podrían considerar estrategias cooperativas entre los dos métodos. El trabajo futuro tiene como objetivo desarrollar algoritmos híbridos, combinando las ventajas de SOTS y HSA y así minimizar sus debilidades.

B. Logística de la distribución de cargas y paquetería

A partir de las características del VRPTW relevadas en la región pampeana y del problema modelado en consecuencia, se está en proceso del diseño e implementar un software logístico como soporte para la toma de decisiones a nivel gerencial de empresas que incluyan en sus actividades tareas de distribución de cargas y paquetería. Dicha tecnología permitirá, a partir de la consideración de las distintas variables y restricciones, conseguir el mayor rendimiento y eficacia en los procesos logísticos. Para determinar la eficacia del software logístico desarrollado para resolver el VRPTW, se utilizará un conjunto de datos de referencia bien conocidos. Se esperan hallar soluciones competitivas con las más conocidas en la literatura, así como también posibles nuevas soluciones que no estén sesgadas hacia algunos de los objetivos (cantidad de vehículos y el costo total).

IV. FORMACIÓN DE RECURSOS HUMANOS

Cada año se incorporan al LISI alumnos avanzados en la carrera Ingeniería en Sistemas como becarios de investigación y también para realizar trabajos finales de carrera. El trabajo con problemas actuales y regionales favorece la formación de los recursos humanos en la actividad de transferencia de los resultados de la investigación desarrollada. También es importante resaltar la interdisciplinariedad en el tratamiento de las distintas propuestas ya que permite vincular áreas ingenieriles. Por otra parte, los docentes-investigadores que integran el proyecto realizan diversos cursos de posgrado relacionados con la temática del proyecto, con el objetivo de sumar los créditos necesarios para cursar carreras de posgrado.

REFERENCES

- [1] F. S. Hillier and G. J. Lieberman, *Introducción a la investigación de operaciones*, 9th ed. McGRAW-HILL/INTERAMERICANA EDITORES, 2012.
- [2] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Management Science*, vol. 6, no. 1, pp. 80–91, 1959.
- [3] S. Chopra and P. Meindl, *Administración de la Cadena de Suministro: Estrategia Planeación y Operación*, 5th ed. Pearson, 2013.
- [4] R. Ballou, *LOGÍSTICA. Administración de la cadena de suministros*, 5th ed. Pearson, 2004.
- [5] G. Dantzig, R. Fulkerson, and S. Johnson, "Solution of a large-scale traveling-salesman problem," *Journal of the Operations Research Society of America*, vol. 2, no. 4, pp. 393–410, 1954. [Online]. Available: <http://www.jstor.org/stable/166695>
- [6] J. K. Lenstra and A. H. G. R. Kan, "Complexity of vehicle routing and scheduling problems," *Networks*, vol. 11, no. 2, pp. 221–227, 1981. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/net.3230110211>
- [7] E. Talbi, *Metaheuristics: from Design to Implementation*. Wiley, 2009.
- [8] R. Farmani, G. A. Walters, and D. A. Savic, "Trade-off between total cost and reliability for anytown water distribution network," *Journal of Water Resources Planning and Management*, vol. 131, no. 3, pp. 161–171, 2005.
- [9] I. Gupta, A. Gupta, and P. Khanna, "Genetic algorithm for optimization of water distribution systems," *Environmental Modelling & Software*, vol. 14, no. 5, pp. 437–446, 1999.
- [10] C. Bragalli, C. D'Ambrosio, J. Lee, A. Lodi, and P. Toth, "On the optimal design of water distribution networks: a practical MINLP approach," *Optimization and Engineering*, vol. 13, no. 2, pp. 219–246, 2012.
- [11] R. Uma, "Optimal design of water distribution network using differential evolution," *International Journal of Science and Research (IJSR)*, vol. 5, no. 11, pp. 1515–1520, 2016.
- [12] R. Mansouri and M. Mohamadizadeh, "Optimal design of water distribution system using central force optimization and differential evolution," vol. 7, no. 3.
- [13] A. De Corte and K. Sörensen, "Hydrogen," available online:
- [14] M. Carnero, J. Hernández, and M. Sánchez, "A new meta-heuristic based approach for the design of sensor networks," *Computers Chemical Engineering*, vol. 55, pp. 83–96, 2013.
- [15] C. Bermúdez, C. Salto, and G. Minetti, "Designing a multi-period water distribution network with a hybrid simulated annealing," in *XLVIII JAIIO: XX Simposio Argentino de Inteligencia Artificial (ASAI 2019)*. Universidad Nacional de Salta, 2019, pp. 39–52.
- [16] N. Kohl, "Exact methods for time constrained routing and related scheduling problems," 1995.
- [17] B. Ombuki, B. Ross, and F. Hanshar, "Multi-objective genetic algorithms for vehicle routing problem with time windows," *Applied Intelligence*, vol. 24, pp. 17–30, 02 2006.
- [18] V. F. Yu, H. Susanto, P. Jodiawan, T.-W. Ho, S.-W. Lin, and Y.-T. Huang, "A simulated annealing algorithm for the vehicle routing problem with parcel lockers," *IEEE Access*, vol. 10, pp. 20 764–20 782, 2022.
- [19] J.-F. Cordeau and G. Laporte, *Tabu Search Heuristics for the Vehicle Routing Problem*. Boston, MA: Springer US, 2005, pp. 145–163. [Online]. Available: https://doi.org/10.1007/0-387-23667-8_6

Confección de dataset de patentes de automóviles argentinos para entrenamiento de redes neuronales.

Ferreira Biron, Martín, Maidana, Carlos Eduardo

Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza,
Buenos Aires, Argentina.

{mferreira,cmaidana} @unlam.edu.ar

RESUMEN

En el marco de una investigación más extensa, se conformó un vasto dataset compuesto por imágenes de escenas donde se encuentran automóviles, imágenes de automóviles, patentes, caracteres y dígitos de las mismas en ambas versiones de la patente argentina. Dicho dataset fue utilizado para entrenar distintas redes neuronales y depurado de diversas formas en varias ocasiones. El mismo posee más de 297.000 imágenes y puede ser utilizado en una investigación de posgrado o doctorado (con la previa anuencia de la UNLaM) para entrenar una red neuronal o para cualquier otro tipo de investigación en donde se necesiten este tipo de datos. En este trabajo se muestran las características de dicho dataset y los resultados al utilizarlo en ciertas redes neuronales.

Palabras Clave: Dataset, Redes neuronales, Inteligencia artificial, Chapa patente argentina.

CONTEXTO

El presente documento es resultado de la investigación “*Obtención de dataset a partir de imágenes capturadas para su utilización en sistema experto de reconocimiento de imágenes*” la cual es una continuación de la investigación anterior “*Obtención de dataset a partir de imágenes*

capturadas del estacionamiento de UNLaM para su utilización en sistema experto de reconocimiento de imágenes” inconclusa por distintas desavenencias y la imposibilidad de llevarla a cabo debido a las diversas situaciones ocasionadas por la pandemia de COVID-19.

1. INTRODUCCIÓN

En la situación en la que se encuentran desarrolladas las redes neuronales hoy en día, es de absoluta y vital importancia tener un nutrido dataset, bien clasificado y depurado para poder entrenar correctamente una red neuronal [1]. Y el hincapié de hoy en día está dado porque una red neuronal va a ser más precisa si la cantidad de datos utilizados en esta etapa es verdaderamente abundante. Es cierto que existen técnicas como *data augmentation* [2] que permiten mitigar una deficiencia de esta materia prima, pero sin embargo poseer un importante dataset, cuantioso y variado hace a un mejor desempeño a la hora de utilizar una red neuronal. También es una verdad insoslayable que existen diversos datasets que cubren, en general, una gran variedad de objetos que pueden ir desde números, como la del famoso dataset MNIST [3], pasando por datasets de rostros humanos como DigiFace-1M [4] hasta llegar a datasets de mariposas como Leeds Butterfly Dataset [5] e incluso estos

datasets pueden contener varias clases de objetos como Imagenet [6] o COCO [7]. La dificultad comienza cuando los datasets que se necesitan no son de un objeto general, como un carácter escrito o como un rostro humano, sino algo más local o específico de una región, como por ejemplo un dataset de imágenes de patentes argentinas. En estos casos, desde nuestra perspectiva, no es una opción utilizar datos sintéticos [9].

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

Al anhelar poseer una forma de confirmar qué automóviles ingresaban en el estacionamiento universitario y también, por otras razones se propuso entrenar un sistema experto con el fin de poder detectar cuales eran las patentes de los automóviles que ingresaba al mismo. Pero nos encontramos con el inconveniente de no tener a nuestra disposición una base de datos de imágenes con patentes argentinas. Debido a esto llevamos a cabo la primera investigación mencionada en la sección Contexto, la cual no pudo finalizarse por los motivos ya explicados. En esta investigación, en líneas generales, se planteaba obtener un dataset de imágenes de patentes argentinas, para su posterior utilización en el entrenamiento de un sistema experto. Al no poder cumplirse completamente con los objetivos de la misma, se continuó la investigación con la siguiente: *“Obtención de dataset a partir de imágenes capturadas para su utilización en sistema experto de reconocimiento de imágenes”*. En esta se propone completar el dataset que se deseaba confeccionar en la investigación anterior y utilizarlo en el entrenamiento de diversas redes neuronales. Este documento presenta las características del dataset obtenido, y a modo de ejemplo cómo se desempeñó en el

entrenamiento de una red neuronal YoloV3-Tiny[8].

Para elaborar este dataset se propuso obtener al menos 100 horas de grabación de tráfico, intentado que estas fueran lo más heterogéneas posibles. Luego de haber obtenido estas, se revisaron en su totalidad en búsqueda de vehículos, patentes del tipo 1994 y 2015 y sus respectivos caracteres. Estas búsquedas persiguieron el fin de obtener y resguardar la información de las posiciones de los elementos encontrados en la imagen, su alto, su ancho y tipo de objeto, entre otros sumados a la imagen del objeto. Además se conservaron los fotogramas donde se encuentran cada una de las imágenes recolectadas. Para este fin fue desarrollado un software interno que ayudó en gran proporción a realizar esta ardua tarea, haciéndola más tolerable, además de automatizar el resguardo de estos datos. El acopio de la información se realizó en archivos de texto plano que contienen estos valores separados por comas (intentando hacer más fácil a futuro su manejo). Los datos de los objetos obtenidos se encuentran divididos en tres archivos que poseen estructuras distintas dependiendo el caso.

Luego de una extensa comprobación de los datos, valiéndonos de algunas automatizaciones programadas por nosotros y de varias inspecciones oculares pudimos concluir la creación del dataset y realizar los cálculos estadísticos.

Con el dataset concluido, nos dispusimos a desarrollar un sistema que fuese capaz, a partir del uso de redes neuronales y del dataset construido, de detectar vehículos en una escena, luego su patente y finalmente los caracteres que la conforman. De esta forma, si realmente el dataset confeccionado es idóneo, las redes

neuronales entrenadas deberán ser capaces de realizar las detecciones pertinentes.

Para cumplir dichos objetivos consideramos la red neuronal Yolo V3 , en su versión tiny, para la detección de vehículos y patentes. Respecto a los caracteres, luego de numerosos intentos desistimos de utilizar Yolo para el reconocimiento de los mismos decantándonos por la red neuronal Xception. En cambio utilizamos Yolo para la localización de los caracteres en la patente.

En resumen el pipeline propuesto es el siguiente:

1. Captura de la escena
2. Detección del vehículo (Yolo)
3. Localización y categorización de la patente. (Yolo)
4. Localización de los caracteres en la patente (Yolo) (Se utilizan dos redes neuronales para esto , una para las patentes 1994 y otra para las patentes 2015)
5. Reconocimiento de los caracteres y dígitos (Xception)

3. RESULTADOS OBTENIDOS/ESPERADOS

Los primeros resultados a destacar son las estadísticas del dataset confeccionado. Como aseveramos en el apartado anterior, la cantidad de horas de vídeo conseguidas asciende a la cifra de 100 horas en distintos lugares como las localidades de Morón, San Justo, Villa Luzuriaga, Ramos Mejía, Haedo, Ituzaingó y CABA en las cuales pudimos obtener **29.278 imágenes de vehículos**, **33.829 imágenes de patentes** , donde **11.301 corresponden a patentes 2015** y **22.528 corresponden a patentes 1994**. Cabe

mencionar que no todas las patentes tienen una imagen de vehículo asociado y no toda imagen de vehículo obtenido posee una imagen de una patente asociada.

Finalmente de las patentes recolectadas pudimos obtener **62.810 caracteres y 62.958 dígitos de patentes 1994** y **42.704 caracteres y 32.028 dígitos de patentes 2015**. Todas las imágenes se encuentran en las más variadas condiciones meteorológicas, tamaños y perspectivas. La distribución de los datos obtenidos se puede consultar en [10]

Respecto al entrenamiento de las redes neuronales, tuvimos en general una respuesta positiva al entrenar cada una de ellas, más allá de los inconvenientes con los que nos pudimos encontrar a lo largo de esta tarea. En todos los casos utilizamos redes neuronales pre entrenadas intentando hacer uso del concepto de *transferencia de aprendizaje* [2]. Las razones de esto son diversas , aunque la principal es la gran cantidad de tiempo que se requiere para entrenar correctamente estas grandes redes y la segunda es la posibilidad de caer rápidamente en *overfitting* debido a las características de estas.

La primera red neuronal que entrenamos fue la correspondiente a la localización de una patente en la imagen de un vehículo y su categorización en una patente del tipo 1994 o 2015. Para realizar dicha tarea se entrenó una red neuronal YoloV3-Tiny con 24.223 imágenes de vehículos en los cuales se podían observar las patentes. De este conjunto de datos se utilizó un 70% para entrenamiento y un 30% para testeo. Estas imágenes fueron convertidas a escala de grises. Luego de aproximadamente medio millón de epochs, comprobamos el desempeño en el conjunto de entrenamiento y en el conjunto de testeo,

obteniéndose un **porcentaje de acierto del 95%**.

Después de haber concluido con el entrenamiento anterior nos enfocamos en otras dos redes neuronales YoloV3-Tiny con el fin de hallar la ubicación de los caracteres y dígitos. Para el caso de los caracteres de las patentes 2015 tomamos aproximadamente todas las patentes 2015 mencionadas en este trabajo, convertimos estas a escala de grises y dividimos el conjunto de datos en tres, un conjunto de validación con el 10% de las imágenes totales, un conjunto de imágenes de entrenamiento conformado por el 70% del 90% del total de imágenes, y un conjunto de testeo conformado por el 30% del 90% del total de imágenes. Entrenamos esta red neuronal con 1 millón de epochs, modificando aleatoriamente los conjuntos de entrenamiento y testeo cada 50000 epochs. Al finalizar comprobamos el desempeño de la red neuronal con el conjunto de validación (que nunca se utilizó en el entrenamiento) obteniéndose un **porcentaje de acierto del 79%**.

Del mismo modo y aproximadamente con la misma cantidad de epochs entrenamos otra red neuronal Yolo para poder detectar la localización de los caracteres y dígitos en las patentes 1994. **Obtuvimos un desempeño en el set de validación del 73,73%**

Continuando con la validación del dataset nos dedicamos a entrenar dos redes neuronales Xception, una de ellas dedicada a reconocer los caracteres de la patente y la otra dedicada a reconocer dígitos de las mismas en ambos casos sin importar la variante de la patente.

Para el caso de la red neuronal destinada a reconocer caracteres, eliminamos las capas finales de esta red neuronal, y colocamos 3 capas de MLP

totalmente conectados de 800 perceptrones junto con una capa de salida de 26 perceptrones. Estas capas, junto con las dos últimas convolucionales fueron las que entrenamos durante 40 epochs, de una forma similar a como lo hicimos con la red neuronal Yolo para la localización de caracteres y dígitos. Al finalizar el entrenamiento, pusimos a prueba esta red con el conjunto de datos de validación (que no conformó parte del entrenamiento) y obtuvimos un **95,11% de acierto en el reconocimiento de los caracteres**.

Para el caso de la red neuronal destinada a reconocer dígitos, eliminamos las capas finales y colocamos 3 MLP de 800 perceptrones junto con una capa de salida de 10 perceptrones. Estas capas, unidas con la última convolucional fueron las entrenadas durante 120 epochs. El entrenamiento fue realizado de la misma forma que la red neuronal anterior. Al finalizar el entrenamiento y comprobar el rendimiento de esta red conseguimos un **97,79% de acierto en el reconocimiento de dígitos** sobre el set de datos correspondientes a la validación.

Por último y no menos importante entrenamos una red YoloV3-Tiny para detectar los vehículos que aparecen en una escena. Para esto aplicamos el mismo método de entrenamiento explicado anteriormente para la localización de las patentes. Utilizamos el 70% de las imágenes de los vehículos para entrenamiento y el 30% restante como imágenes de testeo. Luego de aproximadamente medio millón de epochs obtuvimos un **92% de efectividad** al poner a prueba esta red con el total de las imágenes de vehículos del dataset.

Si bien los datos mencionados anteriormente son alentadores, en las pruebas realizadas con imágenes que no

corresponden a ninguno de los sets de datos utilizados , no obtuvimos un buen desempeño en algunos casos, ocurriendo esto principalmente en el reconocimiento de caracteres y de dígitos y de vehículos.

Respecto a la detección de vehículos, cometimos un error al capturarlos. No todos los vehículos en las escenas fueron recolectados, solamente aquellos en los cuales se podía observar la patente. Esto produjo un inconveniente al entrenar la red neuronal, ya que en el entrenamiento se ofrecían datos confusos: en una escena existían vehículos , que no fueron marcados como vehículos, y algunos vehículos que sí lo fueron.

Por otra parte, en el caso de los caracteres y dígitos , el reconocimiento no era aceptable. Pocas patentes eran reconocidas correctamente. Entendemos que esto se debía a una incompleta segmentación de los caracteres y dígitos, solucionando dicho inconveniente redimensionando la imagen de las patentes cinco veces, antes de la etapa de segmentación. De esta forma la misma era más precisa y los caracteres podían ser reconocidos con más certeza.

4.FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de este desarrollo forman parte del plantel docente e investigadores del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de la Matanza. En conjunto, los integrantes de este grupo

poseen conocimientos de electronica e informatica.

6. Bibliografía

- [1]S. Raschka y V. Mirjalili, *Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*, Third edition. Birmingham Mumbai: Packt, 2019.
- [2] U. Michelucci, *Advanced applied deep learning: convolutional neural networks and object detection*. New York: Apress, 2019.
- [3]Yann Lecun, “The MNIST database of handwritten digits”, <https://www.clarin.com>
- [4]G. Bae *et al.*, “DigiFace-1M: 1 Million Digital Face Images for Face Recognition”. arXiv, doi: [10.48550/arXiv.2210.02579](https://arxiv.org/abs/10.48550/arXiv.2210.02579).
- [5]Wang, Josiah, “Leeds Butterfly Dataset”. Zenodo, doi: [10.5281/ZENODO.7559420](https://zenodo.org/record/10.5281/ZENODO.7559420).
- [6]“ImageNet”. <https://www.image-net.org/>
- [7]“COCO - Common Objects in Context”. <https://cocodataset.org/#home>
- [8] J. Redmon y A. Farhadi, “YOLOv3: An Incremental Improvement”. arXiv, 8 de abril de 2018. doi: [10.48550/arXiv.1804.02767](https://arxiv.org/abs/10.48550/arXiv.1804.02767).
- [9] “Synthetic data”, *Wikipedia.*: https://en.wikipedia.org/w/index.php?title=Synthetic_data&oldid=112992158
- [10]<https://github.com/mfcattg/DatasetWicc2023>

Metaheurísticas aplicadas al alineamiento múltiple de secuencias de ADN y/o de aminoácidos

Adrián Díaz¹, Gabriela Minetti²

¹ Universidad Nacional de Quilmes

² Laboratorio de Investigación en Sistemas Inteligentes (LISI)

Facultad de Ingeniería - Universidad Nacional de La Pampa

e-mail: diazadriang@gmail.com, minettig@ing.unlpam.edu.ar

Resumen Actualmente en la biología, muchos análisis requieren como punto de partida los resultados de la comparación de la composición de macromoléculas que forman parte de la vida, dado que cuanto más cercanas sean dos macromoléculas en términos evolutivos, más similares y conservadas serán sus secuencias. Esto significa alinear tres o más secuencias de nucleótidos en el caso de ADN/ARN o de aminoácidos en el caso de proteínas. Por ende el problema de alineamiento múltiple de secuencias es una de las tareas primordiales en la Bioinformática y, también, de alta complejidad ya que el problema de alineamiento es NP-completo. Por ende, el objetivo de esta línea de investigación es desarrollar un algoritmo metaheurístico, basado en enfriamiento simulado e híbrido con heurísticas del problema, que permita resolver el alineamiento múltiple eficientemente y, además, sea competitivo con los propuestos en la literatura.

Palabras claves: MSA, ADN, proteína, metaheurística, enfriamiento simulado

CONTEXTO

Desde su creación en 1998, el Laboratorio de Investigación de Sistemas Inteligentes (LISI) de la Facultad de Ingeniería (FI) de la Universidad Nacional de La Pampa (UNLPam), mantiene una importante vinculación con investigadores de las universidades argentinas, Universidad Nacional de San Luis (UNSL), de Río Cuarto (UNRC) y de Quilmes (UNQ), y de la Universidad de Málaga (España), con quienes se realizan trabajos de investigación conjuntos. Esta línea de investigación surge de la interrelación con la UNQ y el desarrollo

de una tesis de maestría en el área de la Bioinformática.

I. INTRODUCCIÓN

En las últimas décadas, los avances en el campo de la Biología Molecular y en las tecnologías de genómica han provocado un crecimiento muy fuerte en la información biológica generada por la comunidad científica. La secuenciación de genomas y de proteomas, la identificación de genes, la generación de perfiles de expresión genética y otras áreas genéticas han demostrado la necesidad de la participación de expertos matemáticos, ingenieros y físicos para obtener resultados en corto tiempo y de mayor calidad en estas áreas. La Bioinformática es, entonces, un campo interdisciplinar que analiza la secuencia genómica, identifica y predecir las estructuras moleculares, determina el perfil de expresión genética, etc. Esta disciplina incluye técnicas computacionales y aplicaciones que llevan a cabo esas actividades, siendo enorme el número de estas técnicas y aplicaciones y muy significativas las diferencias entre ellas.

El *genoma* o secuencia completa de ADN de un organismo constituye la información genética heredable del núcleo celular, es el material genético almacenado en cada una de las células de un organismo. El *ácido desoxirribonucleico* (ADN, DNA son sus siglas en Inglés) contiene la información genética usada en el desarrollo y el funcionamiento de los organismos vivos conocidos y de algunos virus, además de ser responsable de su transmisión hereditaria. EL ADN es un polímero¹

¹Los polímeros son macromoléculas (generalmente orgánicas) formadas por la unión de moléculas más pequeñas.

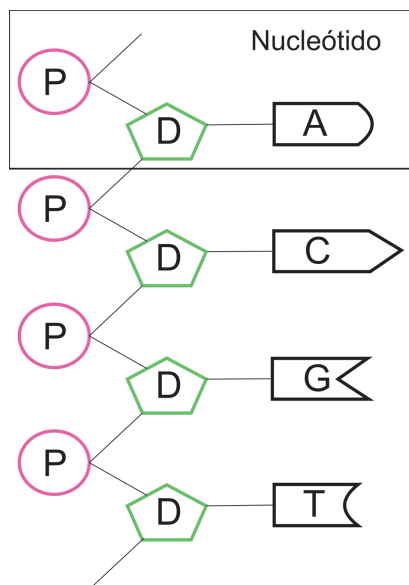


Fig. 1. Representación esquemática de cuatro eslabones de una cadena de nucleótidos. Ácido fosfórico (P), desoxirribosa (D) y bases nitrogenadas (A, C, G, T).

de unidades menores denominados nucleótidos, a saber adenina (A), citosina (C), timina (T) y guanina (G) (véase figura 1). Estructuralmente la molécula de ADN se presenta en forma de dos cadenas helicoidales alrededor de un mismo eje (imaginario); las cadenas están unidas entre sí por las bases de los pares que la forman (véase figura 2). Los emparejamientos son siempre adenina-timina y citosina-guanina.

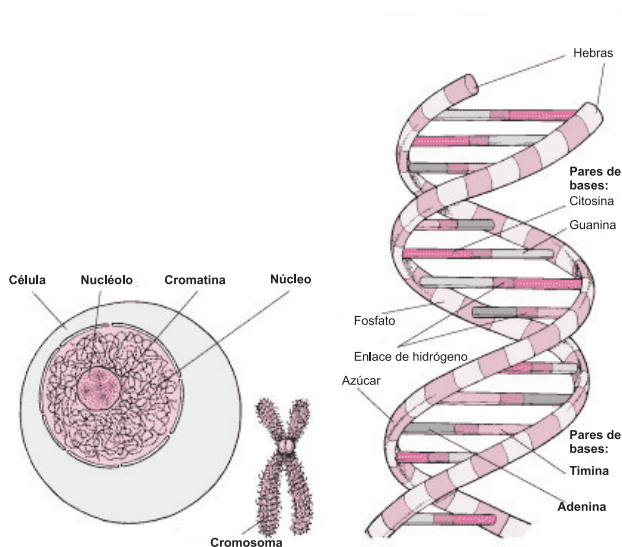


Fig. 2. Molécula de ADN.

La información del ADN es mantenida y ordenada de forma lineal y es la que guía la conformación de la estructura tridimensional de la proteína. Inicialmente, la información lineal de los nucleótidos se transcribe a la información lineal de las secuencias de nucleótidos pertenecientes al ácido ribonucleico (ARN); después de eso, las secuencias de nucleótidos del ARN son traducidas a la información lineal de las secuencias de aminoácidos². En el polipéptido³, la secuencia lineal (estructura primaria) dispone su estructura secundaria y esto determina su organización tridimensional (estructura terciaria). La determinación de una conformación óptima constituye el plegamiento de proteínas. Es un proceso muy complejo, que suministra grandes cantidades de información sobre la presencia de sitios activos⁴ y la posible interacción con medicamentos. Las proteínas en diferentes organismos que están relacionados entre sí por la evolución de un ancestro común, se denominan homólogas. Esta relación puede ser reconocida por la comparación de secuencias múltiples.

El resultado de alinear tres o más secuencias de nucleótidos, en el caso de ADN/ARN, o de aminoácidos, en el caso de proteínas, es utilizado para analizar diversos propósitos: secuenciación genómica, modelado estructural, determinación de familias de proteínas homólogas, reconstrucción de relaciones filogenéticas entre las secuencias, o incluso registración de regiones conservadas que han tenido baja variabilidad a lo largo de la evolución en el tiempo [1].

Cuando se hace un alineamiento de un par de secuencias, se busca maximizar una puntuación o "score" que refleje las características compartidas de las secuencias, es decir que haya una mayor coincidencia de bases entre ambas secuencias: si ambas secuencias son cercanas evolutivamente hablando, una no-coincidencia puede ser interpretada como una mutación, mientras que un espacio o "gap" se considera como una inserción/delección,

²Aminoácido: sustancia química orgánica que constituye el componente básico de las proteínas, compuesta por un grupo amino y un grupo carboxílico. Todas las proteínas de los seres vivos están compuestas por la combinación de 20 aminoácidos.

³Péptido: unión de varios aminoácidos lo suficientemente grande, que cuando tiene una estructura tridimensional única y estable, es una proteína.

⁴Sitio activo: sitio específico en la superficie de la proteína reconocido por los sustratos.

pues ambas secuencias divergieron a través del tiempo. Las regiones compartidas dan un indicio de que esos nucleótidos o aminoácidos tienen una importancia vital para el funcionamiento correcto de las secuencias y una mutación en esa región puede desencadenar que el individuo no logre trascender a lo largo del tiempo.

El objetivo del Alineamiento de múltiples secuencias (*Multiple Sequence Alignment*, MSA) es reproducir las relaciones homólogas entre los residuos en un conjunto de secuencias genéticas o proteínicas. Las mutaciones (sustituciones, eliminaciones/deleciones o inserciones) que se han producido durante el proceso evolutivo hacen que las relaciones homólogas entre residuos sean difíciles de detectar. Por lo tanto, es vital obtener un MSA confiable a partir de un conjunto de secuencias remotamente relacionadas. Por otro lado, los datos de secuencia que se van a analizar se acumulan a un ritmo creciente y, por lo tanto, son requeridos métodos MSA rápidos y fiables. En [2], Bonizzoni et al. demuestran que MSA es un problema NP-completo, en consecuencia es necesario evitar algoritmos que resuelvan este problema en tiempos no polinomiales. Esto conlleva a la utilización de algoritmos de búsqueda aproximados, como son las metaheurísticas estudiadas en esta línea de investigación e introducidas en la siguiente sección.

II. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

Las herramientas informáticas de análisis de datos genómicos han madurado mucho desde su aparición. Los métodos para alinear las secuencias de ADN y proteínas se han desarrollado hasta convertirse en los núcleos necesarios de análisis de la secuencia automática. El objetivo de la alineación de secuencias de ADN es alinear las dos bases en dos o más secuencias reduciendo al mínimo el número de discrepancias. El método de Needleman-Wunsch ha sido uno de los métodos originales que hizo su aparición en 1970 [3]. Esta técnica utiliza la programación dinámica para encontrar una alineación óptima para un par de secuencias. Desde Needleman-Wunsch se han incorporado una serie de nuevos métodos de alineación refinados basados en: programación dinámica [4], [5], BLAST [6], [7], la visualización de secuencias [8], el procesamiento de señales [9], [10], [11], [12], el enfriamiento simulado [13], los algoritmos

genéticos [14], [15], los sistemas de colonias de hormigas [16], entre otros.

En función al análisis de las técnicas existentes surgen las siguientes preguntas de investigación:

- 1) ¿Es posible aplicar un algoritmo metaheurístico basado en enfriamiento simulado (*Simulated Annealing*, SA) hibridado con heurísticas y/o otras metaheurísticas, para construir un alineamiento múltiple de secuencias (MSA)?
- 2) ¿Los resultados obtenidos por el SA propuestos son competitivos con respecto a los obtenidos por las herramientas más usadas en el ámbito de la Bioinformática en la actualidad?

Con el objetivo de responder estas preguntas, se plantea el diseño y el desarrollo de un SA que combinado con heurísticas relacionadas con el MSA sea capaz de obtener resultados comparables o mejores a los de la literatura, tanto para cadenas de longitudes cortas como largas, y para una cantidad amplia de secuencias.

III. RESULTADOS ENCONTRADOS Y ESPERADOS

A partir del desarrollo del SA hibridado para resolver MSA y de su aplicación a un conjunto de casos de estudio de diversa complejidad, se pretende realizar un análisis comparativo entre herramientas de MSA representativas del estado del arte de la Bioinformática y el SA propuesto. A partir de esta comparación se espera detectar las ventajas y desventajas del SA propuesto a la hora de realizar alineamientos múltiples de casos con diferentes características.

IV. FORMACIÓN DE RECURSOS HUMANOS

Con esta línea de investigación se propende a la formación, del primer autor, en investigación siguiendo el método científico. Además, este trabajo le permitirá llevar a cabo gran parte de la investigación requerida en su tesis de maestría.

REFERENCES

- [1] H. J. Notredame C, Higgins DG, "T-coffee: A novel method for fast and accurate multiple sequence alignment," *J. Mol. Biol.*, vol. 302, pp. 205–217, 2000.
- [2] P. Bonizzoni, G. D. Vedova, and G. Mauri, "Experimenting an approximation algorithm for the lcs," *Discrete Applied Mathematics*, vol. 110, no. 1, pp. 13–24, 2001, proceedings of the First Conference on Algorithms and Experiments.

- [3] S. Needleman and C. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *Molecular Biology*, vol. 48, no. 3, pp. 443–453, 1970.
- [4] T. Smith and M. Waterman, "Comparison of biosequences," *Adv. Appl. Math.*, no. 2, pp. 482–489, 1981.
- [5] —, "Identification of common molecular subsequences," *Journal of Molecular Biology*, no. 147, pp. 195–197, 1981.
- [6] S. Altschul, W. Gish, W. Miller, E. Myers, and D. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, no. 190, pp. 403–410, 1990.
- [7] S. Altschul, T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman, "Gapped blast and psi-blast: a new generation of protein database search programs," *Nucleic Acids Research*, no. 25, pp. 3398–3402, 1997.
- [8] Y. Wu, A. Liew, H. Yan, and M. Yang, "Db-curve: a novel 2d method of dna sequence visualization and representation," *Journal of Chemical Physics Letters*, no. 367, pp. 170–176, 2003.
- [9] D. Anastassiou, "Genomic signal processing," *IEEE Signal Processing Magazine*, pp. 8–20, July 2001.
- [10] H. Chang, N. Lo, W. Lu, and C. Kuo, "Visualization and comparison of dna sequences by use of three-dimensional trajectories," *Proceedings of the First Asia-Pacific Bioinformatics Conference APBC2003*, February 2003.
- [11] E. Cheever, G. Overton, and D. Searls, "Fast fourier transform-based correlation of dna sequences using complex plane encoding," *Comput. Appl. Biosci.*, vol. 7, no. 2, pp. 143–154, 1991.
- [12] J. Ning, C. Moore, and J. Nelson, "Preliminary wavelet analysis of genomic sequences," *Proceedings of the Second IEEE Computer Society Bioinformatics Conference CSB2003*, pp. 509–510, August 2003.
- [13] H. Mamitsuka, "Finding the biologically optimal alignment of multiple sequences," *Artificial Intelligence in Medicine*, vol. 35, no. 1-2, pp. 9–18, 2005.
- [14] C. Notredame and D. Higgins, "SAGA: sequence alignment by genetic algorithm," *Nucleic Acids Research*, vol. 24, no. 8, pp. 1515–1524, 1996.
- [15] C. Notredame, L. Holm, and D. Higgins, "COFFEE: an objective function for multiple sequence alignments," *Bioinformatics*, vol. 14, no. 5, pp. 407–422, 1998.
- [16] O. Karpenko, J. Shi, and Y. Dai, "Prediction of mhc class ii binders using the ant colony search strategy," *Artificial Intelligence in Medicine*, vol. 35, no. 1, pp. 147–156, 2005.

APLICACIÓN DE INTELIGENCIA COMPUTACIONAL Y COMPUTACIÓN DE ALTO DESEMPEÑO EN EL DESARROLLO DE UN MODELO DE PREDICCIÓN DE LAS CONDICIONES PREDISPONENTES AL QUEMADO DEL ARROZ (PYRICULARIA ORYZAE)

Miriam Asselborn⁽⁴⁾, Julian Escalante⁽¹⁾, Mariela Lopresti⁽²⁾, Natalia Miranda⁽²⁾, Esteban Schab^(1,3), Karina Cedaro⁽¹⁾, Pablo Fontanini⁽⁶⁾, Malvina Martinez⁽⁵⁾, Carlos Casanova^(1,3), Virginia Pedraza⁽⁴⁾, Fabiana Piccoli^(1,2)

⁽¹⁾ Universidad Autónoma de Entre Ríos, Facultad de Ciencia y Tecnología, Concepción del Uruguay.

⁽²⁾ LIDIC- Univ. Nacional de San Luis, San Luis.

⁽³⁾ Univ. Tecnológica Nacional, Facultad Regional Concepción del Uruguay, Concepción del Uruguay.

⁽⁴⁾ INTA, Estación Experimental Agropecuaria Concepción del Uruguay, Concepción del Uruguay.

⁽⁵⁾ Instituto de Clima y Agua. CIRN. INTA Castelar, Buenos Aires.

⁽⁶⁾ Bolsa de Cereales de Entre Ríos.

Argentina

{omlopres,ncmiran,mpiccoli}@unsl.edu.ar

{schabe,casanovac}@frcu.utn.edu.ar

RESUMEN

Un sistema complejo o crítico con toma de decisiones se caracteriza por la imposibilidad de reproducir para su estudio un escenario sin consecuencias reales, o cuando su resolución implica gran cantidad de recursos para obtener resultados en un tiempo prudencial. La complejidad puede darse por las características del problema o por la cantidad de datos con los que se trabaja. Tomar decisiones en estos contextos debe conjugar dos atributos usualmente contrapuestos: calidad y velocidad. En este trabajo proponemos una línea de investigación enfocada en analíticas, principalmente las prescriptivas, capaces de determinar acciones a ser ejecutadas en el momento (decisiones operativas) o en el futuro (decisiones tácticas: corto y mediano plazo, decisiones estratégicas: largo plazo) para lograr un objetivo deseado. A esta línea se suman investigaciones en Inteligencia Computacional y Computación de Alto Desempeño con el fin de obtener, de forma colaborativa, calidad y velocidad en las decisiones.

Palabras clave: Inteligencia Computacional. Analíticas. Big Data. Computación de Alto Desempeño.

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro de los proyectos de investigación: “Inteligencia Computacional aplicada al desarrollo de un modelo Eficiente de predicción de las condiciones predisponentes al Quemado del Arroz (Pyricularia oryzae) en Argentina” (PICTO-UADER-UNER-00020, UADER e INTA) y “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” (LIDIC, UNSL).

1. INTRODUCCIÓN

El modelado computacional es una gran herramienta para poder generar nuevos conocimientos en diferentes áreas de la ciencia como la biología, física y agronomía [11, 18,24] entre otras. El uso de herramientas y técnicas computacionales permite llevar a cabo el estudio de un sistema y el análisis de los posibles resultados para tomar buenas decisiones. Entre ellas se encuentran la inteligencia computacional, la simulación y la computación de alto desempeño [4,25].

Desarrollar sistemas complejos para la toma de decisiones basados en técnicas de simulación e inteligencia artificial es una de las prácticas más difundidas a la hora de resolver computacionalmente problemas muy grandes, ya sea por su complejidad o por la cantidad de datos involucrados, generalmente obtenidos de la realidad en forma off-line o en tiempo real. Los sistemas reales complejos, generalmente, no son deterministas ni dicotómicos y requieren, frecuentemente de gran cantidad de datos, imposibles de reconocer, procesar y comprender al mismo tiempo por un humano [9,27,28]. Un modelo computacional de estos sistemas debe maximizar su utilidad, y para ello se debe lograr un balance apropiado entre: complejidad, credibilidad e incertidumbre [19]. Uno de los enfoques para lograrlo es el de la Inteligencia Computacional (IC), también referida como Soft Computing, la cual se inspira en la naturaleza para desarrollar sistemas inteligentes basados en computadoras. La estrategia de la IC es explotar la tolerancia a la incertidumbre, aproximación y falta de información para lograr eficiencia, robustez y soluciones de bajo costo. Ejemplos de ella son las Redes Neuronales Artificiales, los Sistemas Difusos y la Computación Evolutiva [7]. Todos estos métodos pueden ser utilizados juntos y complementariamente para resolver un problema [23]. En un contexto de toma de decisiones, la IC tiende a soportar el proceso aprovechando la imprecisión e incertidumbre del mismo para brindar “soluciones más satisfactorias”. Las decisiones se derivan de analíticas, las cuales generan el conocimiento necesario según los datos disponibles. Existen distintos tipos y niveles de analíticas, el nivel depende del grado de automatización en el proceso. Su diseño no es una tarea sencilla, tampoco su implementación computacional, ya que se deben considerar, por ejemplo, aspectos como la toma de decisiones multicriterio y multipersona, la fusión de información de

diferentes fuentes y naturalezas, el análisis de datos e información, etc.

Por todo lo expuesto, implementar una solución computacional para un sistema complejo mediante un modelo que soporte la toma de decisiones demanda gran cantidad de recursos a fin de obtener resultados en un tiempo adecuado. La inclusión de técnicas y estrategias de computación de alto desempeño es mandatoria, principalmente la Computación Paralela [8,14,17], más teniendo en cuenta la naturaleza de la IC y las características de los recursos informáticos actuales [20,21].

Un sistema real complejo es el relacionado con la producción de arroz. El arroz es el alimento básico más importante a nivel mundial, después del trigo, en Argentina tiene importancia regional, se cultiva bajo riego por inundación, en las provincias de Entre Ríos, Corrientes, Santa Fe, Chaco y Formosa [22]. El Quemado del Arroz (QA), causado por el hongo *Pyricularia oryzae*, es la enfermedad más importante del cultivo a escala mundial [3,10]. En Argentina, es la enfermedad más temida, su aparición es esporádica, pero puede causar pérdidas totales a nivel de lote [16].

En cuanto al manejo del QA, si bien la resistencia genética es la principal herramienta, las variedades más difundidas en Argentina son susceptibles al patógeno. El control químico, mediante aplicaciones de fungicidas registrados para el QA, durante la floración o principio de llenado de granos, ayuda a disminuir los daños. No obstante, el patógeno posee una gran facilidad para desarrollar resistencias a los fungicidas utilizados para su tratamiento [5].

Los síntomas del QA pueden aparecer en cualquier etapa de desarrollo. La intensidad varía con el ambiente y con la susceptibilidad del cultivar. En Argentina, el síntoma más temido es la necrosis del raquis de panoja, afectando directamente al rendimiento. Según las condiciones naturales (suelo, clima, entre otras) y la intensidad de los monitoreos de los lotes, la

necrosis del raquis de panoja puede ser el primer síntoma de la enfermedad detectado por el productor/asesor en el campo y, en general, suele ser demasiado tarde como para una aplicación exitosa de fungicidas.

Las condiciones meteorológicas predisponentes al QA han sido ampliamente estudiadas a nivel mundial. Los factores más importantes en el desarrollo de la enfermedad son la temperatura, la humedad relativa, la nubosidad, y las horas de mojado de hoja [6]. En condiciones naturales de campo, las prácticas de manejo del cultivo, como la elección del cultivar, el riego y la fertilización nitrogenada, pueden alterar la precisión de las predicciones de aparición de la enfermedad basadas en datos meteorológicos. La relación entre factores zonales y locales, y la intensidad de enfermedad, están siendo estudiadas para el desarrollo de modelos de predicción en Entre Ríos [16].

Por todo lo expuesto, y según la bibliografía existente, el desarrollo del QA constituye un sistema candidato a ser estudiado mediante métodos computacionales de distinta índole. Se han reportado al menos 52 modelos predictivos, de los cuales sólo ocho se encuentran en uso actualmente en todo el mundo, y sólo 4 han sufrido posteriores revisiones [6]. No obstante, nuevos modelos de predicción surgen debido a dificultades que los existentes exhiben, sobre todo relativos a la calidad e integridad de los datos, y al alcance geográfico de los mismos. Recientemente, se ha estudiado la efectividad de algunos métodos provenientes del *machine learning* (por ejemplo, generación de reglas M5 y Redes Neuronales Artificiales LSTM) para la predicción de la aparición del QA, con resultados competitivos respecto a modelos clásicos como el de Yoshino y WARN [2,12,15,26]. Sin embargo, existen aún muchas técnicas avanzadas de la inteligencia computacional que no se han probado en este problema y que pueden resultar superadoras respecto de las ya analizadas. Entre ellas se

encuentran los autómatas celulares como modelos, la lógica difusa para la generación de reglas y otros tipos de arquitecturas de redes neuronales artificiales, así como las posibles hibridaciones entre tales técnicas.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Accionar ante la real o posible presencia de QA en los cultivos, puede determinarse que es un sistema complejo y apto para el desarrollo de herramientas que permitirán analizar la interacción de distintos parámetros en un corto tiempo, y generar herramientas para el productor agropecuario que le permita tomar decisiones y actuar en forma predictiva adelantándose a los hechos o ante los primeros indicios de la enfermedad. Asimismo, se evitarían aplicaciones preventivas innecesarias de fungicidas, con beneficios al medioambiente y a la economía de la producción.

Desarrollar un sistema computacional para la detección temprana del QA implica diferentes líneas de investigación, podemos destacar a:

- Modelos de inteligencia computacional.
- Procesamiento de imágenes y realidad aumentada.
- Tecnologías del Lenguaje Humano y Visualización de la información.
- Computación de alto desempeño, a aplicar en cada una de las otras líneas y en la mayor parte del proceso.

Cada una de estas líneas aportarán al desarrollo de una herramienta de software con emisión de alertas para un accionar preventivo, de manera inteligente, ante la presencia de condiciones predisponentes para el desarrollo de QA.

Con toda la investigación que implica y el desarrollo final, no solamente se introduciría valor agregado al sistema productivo arrocerero, permitiendo a los productores un accionar predictivo y proactivo, resultando en una reducción de los costos y, consecuentemente,

una mejora sensible en la seguridad de la cosecha, sino también generar conocimiento relacionado con el comportamiento del QA y las formas de tratamiento, como así también respecto a la aplicación del modelo en problemáticas similares, ya sea en el mismo dominio u otro distinto.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se está trabajando en la modelización del sistema mediante Autómatas Celulares y Agentes de Aprendizaje. Para ellos se necesita el estudio previo del comportamiento de los cultivos y del QA, como así también de los parámetros que influyen en su desarrollo. Inicialmente, se está trabajando de manera incremental respecto a los parámetros. Además, se están considerando otras herramientas de IC para incluir. En todos los casos se plantean las soluciones mediante la aplicación de HPC en GPU.

Como uno de los objetivos es lograr soluciones paralelas portables, de costo predecible, capaz de explotar las ventajas de modernos ambientes HPC a través de herramientas y “frameworks de computación” de alto nivel [14], los primeros desarrollos se están haciendo en GPU Nvidia con CUDA [13], y se prevé el uso de otras tecnologías como OpenCL[1].

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados esperados respecto a la formación de recursos humanos son el desarrollo de 2 tesis de doctorado y de varias tesinas de grado en las universidades intervinientes.

5. BIBLIOGRAFÍA

- [1] M. Breyer, G. Daiss and D. Pflüger. “Performance - Portable Distributed k-Nearest Neighbors Using Locality - Sensitive Hashing and SYCL”. Int. Workshop on OpenCL. ISBN 9781450390330. Association for Computing Machinery. 2021
- [2] Bregaglio, S., Titone, P., Cappellia, G, Tamborini, L., Mongiano, G., Confalonieri, R.. Coupling a generic disease model to the WARM rice simulator to assess leaf and panicle blast impacts in a temperate climate. *Eur J Agron.* 76:107–17. 2016.
- [3] Carmona, M. Tizón del arroz. Herbario. UBA:http://herbariofitopatologia.agro.uba.ar/?page_id=5571. 2021.
- [4] Huerta, E.A., Khan, A., Davis, E., Bushell, C., Gropp, W., Katz, D., Kindratenko, V., Koric, S., Kramer, W., McGinty, B., McHenry, K., Saxton, A.. Convergence of artificial intelligence and high performance computing on NSF-supported cyberinfrastructure. *J Big Data* 7, 88. Springer. 2020.
- [5] Ishii, H., Hollomon, D.. *Fungicide Resistance in Plant Pathogens : Principles and a Guide to Practical Management*. Springer. ISBN 978-4431556411. 2015.
- [6] Katsantonis, D., Kadoglidou. K., Dramalis, C., Puigdollers, P. Rice blast forecasting models and their practical value: a review. *Phytopathologia Mediterranea*, 56, 2. ISSN 0031-9465. Pp 187–216. 2017.
- [7] Keller, J., Liu, D. Fogel, D.. *Fundamentals of computational intelligence: Neural networks, fuzzy systems, and evolutionary computation*. John Wiley & Sons. ISBN 9781119214342. 2016.
- [8] Kirk, D., Hwu, W.. *Programming Massively Parallel Processors: A Hands-on Approach*. Morgan Kaufmann. 3r ed. ISBN 978012811987. 2016
- [9] Klir, G.. "Uncertainty and information: foundations of generalized information theory." *Kybernetes* . ISSN: 0368-492X, 35 (7-8). Pp. 1297-1299. 2006.
- [10] Li, S., Zhang, X., Wang, W., Guo, X., Wu, Z., Du, W., ... & Xia, L.. Expanding the scope of CRISPR/Cpf1-mediated genome editing in rice. *Molecular Plant*, 11(7). Pp 995-998. 2018.
- [11] Liu, J., Zheng, Q., Mu, X., Zuo, Y., Xu, B., Jin, Y., Li, Y.. Automated tumor proportion score analysis for PD-L1 (22C3) expression

- in lung squamous cell carcinoma. Scientific reports, 11(1). Pp 1-9. 2021.
- [12] Nettleton, D. F., Katsantonis, D., Kalaitzidis, A., Sarafijanovic-Djukic, N., Puigdollers, P., Confalonieri, R. Predicting rice blast disease: machine learning versus process-based models. BMC bioinformatics. vol. 20, no 1, p. 1-16. 2019.
- [13] Nvidia. "CUDA C++ Programming Guide". https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf. 2023.
- [14] Pacheco, P., Malensek, M. An Introduction to Parallel Programming. Morgan Kaufmann. 2 Ed. ISBN 0128046058. 2021.
- [15] Pagani, V., Guarneri, T., Busetto, L., Ranghetti, L., Boschetti, M., Movedi, E., Campos-Taberner, M., Garcia-Haro, F., Katsantonis, D., Stavrakoudis, D., Ricciardelli, E., Romano, F., Holecz, F., Collivignarelli, F., Granell, C., Casteleyn, S., Confalonieri, R.. A high-resolution, integrated system for rice yield forecasting at district level. Agricultural Systems, 168. Pp 181-190. ISSN 0308-521X. 2019.
- [16] Pedraza, M.V.; Liberman, C.A.; Nuñez Bordoy, E.; Sanso, C.; Panozzo, J.; Asselborn, M.N. Control químico del Quemado del Arroz (*Pyricularia oryzae*). Jorn. Técn. Nac. del Cultivo de Arroz. Concordia, Entre Ríos. Pp 119-127. Resultados Experimentales 2012-2013. Vol XXII. Ed. Fundación Proarroz. Pp 67-74. ISSN 1853-8754. 2013.
- [17] Piccoli, F.. Computación de alto desempeño en GPU. La Plata: EDULP, ISBN:978-950-34-0759-2. 2011.
- [18] Raja,R.; Nagwanshi,K.; Kumar, S.; Ramya Laxmi; K.. Data Mining and Machine Learning Applications. John Wiley & Sons. ISBN 9781119791782. 2022
- [19] Ross, T.. Fuzzy logic with engineering applications. John Wiley & Sons, ISBN 9781119235866. 2016.
- [20] Schab, E., Casanova, C., Piccoli, F.. Solving an Instance of a Routing Problem Through Reinforcement Learning and High Performance Computing. In: Rucci,E., Naiouf,M., Chichizola,F., De Giusti,L., De Giusti,A.(eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2022. Comm. in Computer and Information Science, vol 1634. Springer, Cham. https://doi.org/10.1007/978-3-031-14599-5_8. 2022.
- [21] Schab, Esteban, Lopresti, Mariela, Miranda, Natalia Carolina, Casanova, Carlos, Piccoli, María Fabiana. "Toma de decisiones en sistemas de eventos mediante inteligencia computacional y computación de alto desempeño". XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza). Red de Universidades con Carreras en Informática. ISBN: 978-987-48222-3-9. Pp 535-539. Universidad Champagnat, Mendoza. Abril 2022.
- [22] SIBER. Inf. Sup. Sembrada y Prod. del Arroz. Camp. 2020/21. <https://www.bolsacer.org.ar/Fuentes/siberd.php?Id=1259>. 2021.
- [23] Siddique, N.; Adeli, H.. Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing. John Wiley and Sons. ISBN 9781118534816. 2013.
- [24] Skoda, P.; Fathallahman, A. Knowledge discovery in big data from astronomy and Earth observation : astrogeoinformatics. Elsevier. ISBN 9780128191545. 2020.
- [25] Wu, K., Simon, H.D.. High-Performance Computational Intelligence and Forecasting Technologies. Lawrence Berkeley National Laboratory. Ch: 22. Wiley. ISBN: 9781119482086. 2021.
- [26] Yoshino R. Ecological studies on the penetration of rice blast fungus, *Pyricularia oryzae* into leaf epidermal cells. Bulletin of the Hokuriku National Agricultural Experiment Station. ;22: 163–221. 1979
- [27] Zadeh, L. A.. "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," in IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, no. 1, pp. 28-44, 1973.
- [28] Zimmermann, H. . Testability and meaning of mathematical models in social sciences. Mathematical Modelling 1(1). ISSN 0270-0255. Pp 123-139. 1980.

Lenguaje natural aplicado a los sistemas conversacionales para el aprendizaje de lenguas extranjeras

A.Vega (1), M.Bilbao(1), M.Falappa(2)

(1) Grupo de Investigación en Inteligencia Artificial Facultad de Ingeniería UNPSJB
Ruta 1 Km 4 Comodoro Rivadavia - Chubut

alejandrovega@unpata.edu.ar, martinbilbao@ing.unp.edu.ar,

(2) Departamento de Ciencias e Ingeniería en Computación UNS
San Andrés 800 – Campus de Palihue - Bahía Blanca
mfalappa@cs.uns.edu.ar

RESUMEN

La computación afectiva es un área de investigación emergente que reúne a investigadores y profesionales de varios campos, que van desde el estudio de la IA, el procesamiento del lenguaje natural hasta las ciencias cognitivas y sociales (Poria et al., 2017, 98-127). Las mismas se centran en el estudio y desarrollo de sistemas y dispositivos que pueden reconocer, interpretar, procesar y simular los afectos humanos.

El objetivo de la computación afectiva es lograr que una máquina o dispositivo sea capaz de simular la empatía i.e: “La capacidad de comprender el estado emocional de un individuo y, a menudo, expresar una respuesta que resuene con él” (Rodríguez et al., 2021), generando una inteligencia emocional que sea capaz de reconocer y responder de una manera más adecuada ante situaciones cotidianas. Para la detección de las emociones en el último tiempo se ha empezado a trabajar con sistemas que indican la utilización de más de una modalidad, también llamados multimodales, como por ejemplo la modalidad visual, de audio, de texto o

entradas fisiológicas. En este estudio se enfocará en el uso de información audiovisual, es decir bimodal.

Actualmente se utiliza el término deep learning (aprendizaje profundo), a partir de Big Data emocional. En (Shamim Hossain & Muhammad, 2019), se propone el procesamiento de una señal de voz para obtener un Espectrograma de Mel, y transformar las frecuencias de sonido según las escuchamos los humanos. Esto sirve, por ejemplo, para clasificaciones de canciones o clasificaciones del dictado (speech recognition). Luego, se extraen unos cuadros representativos de un segmento de video que alimentan una red neuronal convolucional (CNN, Convolutional Neural Network), y las salidas de esta red se fusionan utilizando dos máquinas de aprendizaje extremo (ELMs, extreme learning machines) consecutivas. La salida de esta fusión las califica una máquina de vectores de soporte (SVM, support vector machine) indicando la emoción encontrada, utilizando bases de datos emocionales audiovisuales. Por último, el estudio indica

si los sistemas que involucran CNN y ELMs son eficaces.

Por otro lado, existe el NLP (Procesamiento del Lenguaje Natural) como rama de la IA que ayuda a una computadora a entender, interpretar y manipular el lenguaje, y es una herramienta importante en el campo de la comunicación humano-máquina. Los avances de la última década han dado como resultados que estas se utilicen cada vez más en una gran variedad de campos como la atención médica, finanzas, marketing, etc. Se podría decir que el objetivo principal de las NLP es traducir el lenguaje humano (o natural) en comandos que puedan ser ejecutados por una computadora.

Palabras claves: Computación Afectiva, Empatía, Sistemas Bimodales

CONTEXTO

El proyecto de investigación se desarrolla bajo convenio de actividad interinstitucional entre el Departamento de Ciencias e Ingeniería en Computación de la UNS y el Grupo de Investigación y el Grupo de Investigación en Inteligencia Artificial de la UNPSJB.

1. INTRODUCCION

Los avances tecnológicos actuales en el área de TICs, permiten disponer de cámaras y micrófonos en entornos familiares, sean dedicados o mediante dispositivos como smartphones o notebooks. Junto al creciente uso de redes

sociales, estos dispositivos permiten capturar grandes cantidades de información para la creación de una base de datos audiovisual para el análisis de sentimientos en tiempo real. Así, se puede dotar a los sistemas de la capacidad de interactuar con los usuarios de una manera más empática. Esto quiere decir, obtener una respuesta lo más cercana posible a la humana en las interacciones humano-máquina. El estudio de la computación afectiva y del análisis de los sentimientos son las claves para obtener Inteligencias artificiales (IAs), más complejas y cercanas al comportamiento humano, como así también para todos los campos de estudios que derivan de ellas.

El objetivo de la computación afectiva es lograr que una máquina o dispositivo sea capaz de simular la empatía i.e: “La capacidad de comprender el estado emocional de un individuo y, a menudo, expresar una respuesta que resuene con él” (Rodriguez et al., 2021), generando una inteligencia emocional que sea capaz de reconocer y responder de una manera más adecuada ante situaciones cotidianas. Para la detección de las emociones en el último tiempo se ha empezado a trabajar con sistemas que indican la utilización de más de una modalidad, también llamados multimodales, como por ejemplo la modalidad visual, de audio, de texto o entradas fisiológicas. En este estudio se enfocará en el uso de información audiovisual, es decir bimodal.

2. LINEAS DE INVESTIGACION, DESARROLLO E INNOVACION

El principal objetivo de este trabajo es la obtención de información relacionada a la computación afectiva y el análisis de sentimientos como campos interdisciplinarios para el procesamiento y comprensión del lenguaje natural (NLP) aplicando enfoques híbridos. De esta manera, se podrá brindar una herramienta de aprendizaje de idiomas extranjeros con la cual sea posible tener una conversación fluida con una máquina y sea capaz de evaluar a la misma en tiempo real y de una manera empática.

Objetivos Específicos:

- Evaluar y seleccionar diferentes métricas para bases de datos y repositorios con información sobre los sistemas conversacionales basados en lenguajes extranjeros.
- Desarrollar un enfoque híbrido (análisis de videos y audio) utilizado para la obtención de datos para los procesos de reconocimiento de emociones.
- Mejorar la respuesta en la interacción humano-computadora, a través del empleo de computación afectiva para el aprendizaje de lenguas extranjeras en los sistemas conversacionales actuales.
- Implementar algoritmos de computación afectiva que puedan ejecutarse en tiempo real.
- Diseñar una app de asistencia conversacional basada en el idioma

español y con aplicación de computación afectiva.

3. RESULTADOS Y OBJETIVOS ESPERADOS

Realizar una buena obtención de datos tanto visuales como auditivos, para la conformación de una buena base de datos, para poder cumplir con ciertos objetivos específicos que podemos enumerar:

- Evaluar y seleccionar diferentes métricas para bases de datos y repositorios con información sobre los sistemas conversacionales basados en lenguajes extranjeros.
- Desarrollar un enfoque híbrido (análisis de videos y audio) utilizado para la obtención de datos para los procesos de reconocimiento de emociones.
- Mejorar la respuesta en la interacción humano-computadora, a través del empleo de computación afectiva para el aprendizaje de lenguas extranjeras en los sistemas conversacionales actuales.
- Implementar algoritmos de computación afectiva que puedan ejecutarse en tiempo real.
- Diseñar una app de asistencia conversacional basada en el idioma español y con aplicación de computación afectiva

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por tres investigadores, dos doctores y un alumno del Doctorado en Ciencias de la Ingeniería de la UNPSJB

Esta línea de investigación proporcionará un marco propicio para la iniciación de estudios de posgrado del alumno.

Dos integrantes de esta línea de investigación dirigen tesis de grado y posgrado en temáticas afines.

5. REFERENCIAS

Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). Affective Computing and Sentiment Analysis. A Practical Guide to Sentiment Analysis. Socio-Affective Computing, 5. https://doi.org/10.1007/978-3-319-55394-8_1

Choi, J. I., Ahmadvand, A., & Agichtein, E. (2019, November). Offline and Online Satisfaction Prediction in Open-Domain Conversational Systems. Conversational Systems. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19), 1281–1290. <https://doi.org/10.1145/3357384.3358047>

Hamed Jelodar, H., & Wang, Y. (2020, 10 24). Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach. IEEE J Biomed Health Inform,

10, 2733-2742. [10.1109/JBHI.2020.3001216](https://doi.org/10.1109/JBHI.2020.3001216)

Lou, C., Liang, B., Gui, L., He, Y., Dang, Y., & Xu, R. (2021, 07). Affective Dependency Graph for Sarcasm Detection. SIGIR '21: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 44, 1844-1849. <https://doi.org/10.1145/3404835.3463061>

Majumder, B., Gupta, A., Surana, H., & Vajjala, S. (2020). Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems. O'Reilly Media.

Moschona, D. S. (2020). An Affective Service based on Multi-Modal Emotion Recognition, using EEG enabled Emotion Tracking and Speech Emotion Recognition. 2020 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia), 1-3. [10.1109/ICCE-Asia49877.2020.9277291](https://doi.org/10.1109/ICCE-Asia49877.2020.9277291).

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. Information Fusion, 37, 98-125. <https://doi.org/10.1016/j.inffus.2017.02.003>

Rodriguez, G., Jofré, N., Alvarado, Y., Fernandez, J., & Guerrero, R. (2021, Abril 15). INTERACCIÓN HUMANO-COMPUTADORA: ESTRATEGIAS HACIA LA EMPATÍA COMPUTACIONAL. XXIII Workshop de Investigadores en Ciencias de la

Computación, 270 - 274. ISBN: 978-987-24611-3-3

Shamim Hossain, M., & Muhammad, G. (2019, September). Emotion recognition using deep learning approach from audio–visual emotional big data. *Information Fusion*, 49, 69-78. 10.1016

Wang, Y., Wu, J., & Hoashi, K. (2019). Multi-Attention Fusion Network for Video-based Emotion Recognition. *International Conference on Multimodal Interaction (ICMI '19)*, 595–601. <https://doi.org/10.1145/3340555.3355720>

Tecnologías Semánticas para el desarrollo de Agentes Inteligentes: Generación de comentarios a partir de código fuente.

Cristian Vincenzini and Sandra Roger

email: `cristian.vincenzini@est.fi.uncoma.edu.ar`
`roger@fi.uncoma.edu.ar`

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

La generación de comentarios de código es la tarea de generar una descripción de lenguaje natural de alto nivel para un método o función de código determinado.

Este trabajo, explora el uso de tecnologías de generación automática de texto explicativos a partir de código fuente. A partir de esta exploración, se pretende implementar un módulo que provea asistencia al desarrollador en la generación de un texto descriptivo, sobre una sección de código a su elección.

Para lograr este objetivo, se hará uso de técnicas de aprendizaje automático que permiten entrenar un modelo a partir de un corpus adecuado. Se empleará la técnica de transferencia de conocimiento en un modelo basado en atención, que permite aprovechar el aprendizaje previo de otros modelos con pequeños conjuntos de datos.

Para la evaluación de la calidad y utilidad de los comentarios generados se utilizarán diferentes lenguajes de programación correspondientes a distintos paradigmas.

Palabras Clave: Generación de Lenguaje Natural, Aprendizaje Automático, Código, Comentarios.

Contexto

Este trabajo está parcialmente financiado por la UNCo, en el marco del nuevo proyecto de investigación *Tecnologías Semánticas para el desarrollo de Agentes Inteligentes*. El proyecto de investigación tiene una duración de cuatro años y ha comenzado en 2022 y se desarrolla en forma colaborativa con docentes-investigadores de la UNS.

1. Introducción

El objetivo de la generación de lenguaje natural (GLN) es producir textos que sean indistinguibles del lenguaje natural hablado o escrito por una persona, de manera tal que el resultado final sea comprensible para cualquiera que reconozca el mismo lenguaje. La GLN a partir de código atañe a proveer de manera automatizada una descripción resumida de un fragmento de código [1]. Estas descripciones, también conocidas como comentarios, suelen ser utilizados para explicar el propósito y el funcionamiento de las diferentes partes del código fuente de un programa, y suelen ser escritos en un lenguaje natural legible para los humanos. Este proceso se realiza mediante el uso de algoritmos de procesamiento de lenguaje natural, algoritmos que permiten identificar la estructura gramatical y el significado de las palabras en un texto.

Desafortunadamente, gran cantidad del código fuente existente no se encuentra comentado, o bien existe pero no describe correctamente la funcionalidad que pretende explicar.

Es importante tener en cuenta que la generación de comentarios a partir de código fuente es una tarea compleja: problemas como la ambigüedad semántica, la falta de contexto o una sintaxis compleja del código analizado invariablemente impactan y reducen la calidad del resultado obtenido.

Existen diferentes enfoques para la generación de comentarios a partir de código fuente [2, 3, 4]. Un enfoque para esta tarea es aplicando técnicas de minería de datos para analizar el código y extraer información relevante del mismo. En la actualidad sin embargo es más común utilizar grandes modelos de lenguaje, que no son más que modelos estadísticos entrenados con millones de datos para realizar una tarea.

Una de las características de las arquitecturas de estos sistemas es que permiten refinar su aprendizaje para realizar una tarea específica, mecanismo que se conoce como *transferencia de aprendizaje*. Mediante esta técnica es posible utilizar estos grandes modelos ya entrenados y adaptarlos, obteniendo resultados aceptables en un tiempo considerablemente inferior al de entrenar un modelo desde cero.

Para este trabajo utilizaremos un *Transformer* conocido como T5 (2020) [5]. Este modelo permite procesar múltiples tareas relacionadas a la GLN. La tarea a desarrollar en esta línea de investigación es la generación de comentarios a partir de código fuente. Para ello tomaremos un modelo ya entrenado que será refinado sobre un conjunto de datos propio y posteriormente analizaremos los resultados obtenidos.

2. Línea de Investigación y Desarrollo

El proyecto de investigación *Tecnologías Semánticas para el desarrollo de Agentes Inteligentes* tiene como objetivo general, generar conocimiento especializado en el área de agentes

inteligentes que accedan, procesen y recuperen información mediante tecnologías semánticas.

En este sentido, se desarrolla una línea de investigación enfocada en la generación de lenguaje natural. Específicamente, se busca generar comentarios descriptivos de funciones utilizando técnicas novedales en el campo del aprendizaje profundo.

El modelo de lenguaje utilizado en esta investigación inicial se basa en la arquitectura de transformadores de codificación-decodificación de CodeTrans [6]. En dicho trabajo se utilizaron tres tamaños del modelo T5 para realizar los entrenamientos: *small*, *base* y *large* (con 60, 220 y 770 millones de parámetros respectivamente). Estos modelos fueron entrenados para diferentes tareas en varios lenguajes de programación (Java, Python, Ruby, GO, entre otros) utilizando transferencia de aprendizaje.

La transferencia de aprendizaje consiste en dos etapas. Una etapa inicial de entrenamiento auto-supervisado donde se utilizan datos sin etiquetar y una segunda etapa, conocida como *fine-tuning* donde el modelo se entrena para una tarea específica utilizando datos etiquetados. Otra técnica utilizada en CodeTrans es el entrenamiento multi-tarea, la cual consiste en entrenar un modelo en múltiples tareas, utilizando datos etiquetados y sin etiquetar. Esta metodología permite, además, realizar un *fine-tuning* posterior a través de la transferencia de aprendizaje.

En una primera instancia en este trabajo tomamos la arquitectura T5 ya entrenada con los datos de CodeTrans, tomando los modelos realizados por transferencia de aprendizaje (TF) y multi-tarea (MT) para los tres tamaños considerados: *small*, *base* y *large*, y realizamos *fine-tuning* sobre cada uno de ellos, utilizando sets de datos propios para realizar diversos experimentos.

Estos conjuntos de datos propios están formado por dos corpus contruidos: uno relacionado al lenguaje de programación GO y otro al lenguaje de programación PROLOG. Se están desarrollando pruebas y validaciones del trabajo de estos modelos sobre estos lenguajes a fin de formular hipótesis de investigación futura.

3. Resultados Obtenidos y Trabajos Futuros

Se estudiaron, analizaron y evaluaron diferentes enfoques y arquitecturas disponibles en investigaciones actuales. En base a este estudio se diseñó e implementó una arquitectura a utilizar [7]. Este diseño está realizado sobre el trabajo de CodeTrans [6].

En base a esto y para comparar resultados se construyeron dos corpus de datos sobre dos lenguajes de programación. Se trabajó sobre un lenguaje que haya sido utilizado en dicho trabajo y que sirva como base, y sobre un lenguaje nuevo que no haya sido utilizado en dicha investigación. Ambos corpus contienen 120 líneas. La estructura de los datos consiste en una lista de tuplas, donde cada tupla está compuesta por una función y su correspondiente comentario en inglés que describe su funcionalidad.

Para nuestro primer conjunto de datos, elegimos el lenguaje de programación imperativo GO¹, el cual fue desarrollado por Google en el año 2009. Utilizando principalmente GitHub Copilot², se extrajeron las funciones y sus respectivos comentarios.

Para el siguiente conjunto de datos, seleccionamos PROLOG, un lenguaje lógico que utiliza predicados como elementos de ejecución. Los datos fueron obtenidos de diversas fuentes, pero durante la selección se evitaron predicados provenientes de módulos o paquetes de terceros, así como la utilización de sintaxis de gramáticas de cláusulas definidas (DCG).

En una primera instancia se llevaron a cabo experimentos para evaluar las tareas de *Transfer Learning* (TF) y de *Multi-Task Learning* con *Fine-Tuning* (MT-TF), debido a que en ambas tareas se emplea la utilización de un *Fine-Tune*. En cada tarea, se realizaron experimentos para los modelos de tres tamaños distintos: *small*, *base* y *large*.

Un primer set de experimentos fueron realizados sobre el lenguaje GO para comparar cómo se comportaba nuestro *finetuning*, el cual fue cons-

truido utilizando el corpus descripto anteriormente con respecto a los modelos de CodeTrans de [6] utilizando nuestro *dataset-test*.

Por otro lado, se hicieron evaluaciones de estas mismas tareas en los modelos pequeños y bases, como así también los modelos grandes en el lenguaje PROLOG. Como mencionamos anteriormente, CodeTrans no cuenta con este lenguaje. Al igual que con el lenguaje GO, se construyó un corpus tanto para realizar el *fine-tuning* como el *dataset* utilizado en el testeo. El objetivo fue probar cómo se comportaban los modelos para un lenguaje no contemplado previamente. Además, elegimos PROLOG por ser un lenguaje de programación declarativo, a diferencia de los otros lenguajes utilizados. Para realizar la comprobación se utilizaron dos métricas: BLEU [8] y ROUGE [9]. Dentro de las variantes de la medida ROUGE nos concentramos en la ROUGE-L, con la ventaja de que puede capturar la estructura del nivel de oración de una manera natural.

Según los datos obtenidos y observaciones preliminares se pueden pensar que los comentarios de los lenguaje declarativo tienen cierta semi-estructuración donde se describen, muchas veces, las características de los parámetros que intervienen haciendo uso de los nombres de las variables y de los predicados. Además, en general, la longitud de los comentarios son relativamente más cortos. Esto hace pensar que las medidas de evaluación automáticas que generalmente se utilizan en estas tareas, no sean del todo adecuadas.

4. Formación de Recursos Humanos

Durante la realización de esta investigación se espera lograr, la culminación de 2 tesis de grado dirigidas y/o codirigidas por los integrantes del proyecto.

Así también, se espera que durante el desarrollo del proyecto, los integrantes del mismo estén en proceso de elaboración de su tesis de posgrado puedan consolidar su formación en investigación, y que el trabajo realizado contribuya

¹<https://go.dev/>

²<https://github.com/features/copilot>

a su graduación.

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

Referencias

- [1] Alexander LeClair, Sakib Haque, Lingfei Wu, and Collin McMillan. Improved code summarization via a graph neural network. In *Proceedings of the 28th International Conference on Program Comprehension*, pages 184–195, 2020.
- [2] Srinivasan Iyer, Ioannis Konostas, Alvin Cheung, and Luke Zettlemoyer. Summarizing source code using a neural attention model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2073–2083, 2016.
- [3] Yanlin Wang, Ensheng Shi, Lun Du, Xiaodi Yang, Yuxuan Hu, Shi Han, Hongyu Zhang, and Dongmei Zhang. Cocosum: Contextual code summarization with multi-relational graph neural network. *arXiv preprint arXiv:2107.01933*, 2021.
- [4] Sonia Haiduc, Jairo Aponte, Laura Moreno, and Andrian Marcus. On the use of automated text summarization techniques for summarizing source code. In *2010 17th Working Conference on Reverse Engineering*, pages 35–44. IEEE, 2010.
- [5] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, Peter J Liu, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, 21(140):1–67, 2020.
- [6] Ahmed Elnaggar, Wei Ding, Llion Jones, Tom Gibbs, Tamas Feher, Christoph Angerer, Silvia Severini, Florian Matthes, and Burkhard Rost. Codetrans: Towards cracking the language of silicon’s code through self-supervised deep learning and high performance computing. *arXiv preprint arXiv:2104.02443*, 2021.
- [7] Cristian Vincenzini and Sandra Roger. Generación de comentarios a partir de código fuente utilizando transformers. *XXVIII Congreso Argentino de Ciencias de la Computación*, pages 973–977, 2022.
- [8] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318, 2002.
- [9] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81, 2004.

Deep Learning para Visión por Computadora.

Franco Ronchetti^{1,2}, Facundo Quiroga^{1,3}, Gastón Ríos^{1,3}, Pedro Dal Bianco^{1,3},
Santiago Ponte Ahon¹, Oscar Stanchi¹, Laura Lanzarini¹, Alejandro Rosete⁴, Waldo Hasperué¹

¹ Instituto de Investigación en Informática LIDI, Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina.*

² Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CICPBA)

³ Becario postgrado UNLP

⁴ Universidad Tecnológica de La Habana “José Antonio Echeverría” (CUJAE), La Habana, Cuba

* Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)

Contacto: fronchetti@lidi.info.unlp.edu.ar

CONTEXTO

Esta presentación corresponde a algunas de las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto F025 “Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data” perteneciente al Programa de Incentivos (2018-2023).

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de reconocimiento de patrones en imágenes, video y, utilizando técnicas de Aprendizaje Automático clásicas, y Aprendizaje profundo con Redes Neuronales Convolucionales, Recurrentes y Transformers. El trabajo presentado describe diferentes estrategias inteligentes para la traducción automática de la lengua de señas, particularmente enfocado en la Lengua de Señas Argentina (LSA), junto con herramientas para su aprendizaje e interpretación de resultados.

Este es un problema complejo y multidisciplinar, que presenta diversos subproblemas a resolver como el reconocimiento del intérprete que realiza una seña, la segmentación de manos, la clasificación de diferentes configuraciones y de un gesto dinámico, entre otros.

Por otro lado, la creación de un conjunto de datos apropiado para la LSA resulta esencial para la creación de modelos específicos. Uno de los subproblemas atacados en este artículo es la creación de un conjunto de datos apropiado.

Por otro lado, se está estudiando la forma de reconocer formas de mano de la Lengua de Señas con conjuntos de datos de tamaño reducido, dada la falta de datos de entrenamiento para este dominio.

Por último, se están utilizando Redes Generativas Adversarias (GANs) para aumentar bases de datos de formas de mano, con el objetivo de complementar desde otro enfoque el entrenamiento de modelos para su clasificación.

En otra línea de investigación, se está estudiando la forma en que las redes neuronales codifican la invarianza a las transformaciones y otras propiedades transformacionales, con el objetivo de poder analizar y comparar estos modelos. De esta forma se espera poder mejorar los modelos de clasificación de objetos transformados, en particular, de formas de mano. En el mismo sentido, se están estudiando técnicas de interpretabilidad de los modelos.

Palabras clave: Redes Neuronales, Redes Convolucionales, Redes Recurrentes, Visión por Computadoras, Lengua de Señas, Bases de datos, *Crowdsourcing*, Redes Generativas Adversarias, Invarianza, Equivarianza.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI (III-LIDI) tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos métodos de Aprendizaje Automático y Redes Neuronales.

Como resultado de estas investigaciones se han diseñado e implementado técnicas originales aplicables a la clasificación y el análisis de características de objetos en imágenes, generación de imágenes para aumentación de datos, y estudio del funcionamiento de las redes neuronales. En relación con esta línea, actualmente se están desarrollando los siguientes temas:

1.1. Reconocimiento y traducción de lengua de señas

La traducción de la lengua de señas (SLT) es un campo de estudio activo que abarca la interacción humano-computadora, la visión por computadora, el procesamiento de lenguaje natural y el aprendizaje automático. En esta línea de investigación, el objetivo es el reconocimiento de gestos en videos de Lengua de Señas y la traducción de esta al español. Para el reconocimiento se está trabajando con la base de datos LSA64 [11]. Esta consiste en un registro de 64 señas de la Lengua de Señas Argentina. Se compararon diversas técnicas en el estado del arte del Aprendizaje Automático basadas en Redes Neuronales. Específicamente, se compararon tanto arquitecturas basadas en Redes Recurrentes y Convolucionales, como distintas estrategias de preprocesamiento para optimizar la calidad del reconocimiento. También se están analizando los modelos entrenados para comprender mejor el impacto de estas estrategias de preprocesamiento y de la forma de representación lograda por cada tipo de arquitectura [2][11]. Para la traducción, se confeccionó LSA-T[10] una base de datos masiva y realista de la Lengua de Señas Argentina, que se detalla en la sección 1.2. Este problema representa una mayor complejidad que el reconocimiento ya que implica reconocer varias señas ejecutadas una tras otra en un mismo video y luego traducir las señas identificadas al español, cuya sintaxis y semántica son distintas.



Figura 1. Activación de las capas de una red recurrente para el reconocimiento de lengua de señas en video

1.2. Creación de una base de datos para traducción de Lengua de Señas

Esta línea de investigación tiene como objetivo generar y expandir el primer conjunto de datos continuo de Lengua de Señas Argentina (LSA). Esto es de particular relevancia ya que, en el campo de la traducción de lengua de señas, es necesario contar con bases de datos específicas de la lengua que se busca traducir. No es posible utilizar modelos de traducción entrenados sobre otra lengua de señas para traducir LSA. En este marco, se desarrolló la base de datos “LSA-T”, que contiene 14.880 videos a nivel de oración de LSA extraídos del canal de YouTube CN Sordos. Contiene, además, para cada video, su traducción al español e información posicional correspondiente a cada señante. También se presentó un método para inferir al señante entre varias personas que pudieran aparecer en un mismo video, un análisis detallado de las características del conjunto de datos, una herramienta de visualización para explorar el conjunto de datos y un modelo de SLT neuronal para servir como línea de base para futuros experimentos.

Este desarrollo se lleva a cabo en el marco de una tesis doctoral financiada por la UNLP a través de una beca de postgrado con el objetivo de contribuir al avance de la traducción de la lengua de señas y, en última instancia, mejorar la integración de las personas sordas en la sociedad.



Figura 2. Captura del canal CN Sordos, cuyos videos componen LSA-T

1.3. Generación de imágenes sintéticas

Adicionalmente, se están comenzando a utilizar Redes Generativas Adversarias (GANs) y Modelos de Difusión para generar imágenes y luego videos artificiales relacionadas con la lengua de señas. Este tipo de redes permitirá, por un lado, generar videos de LSA a partir de texto, a modo de traducción y por otro, aumentar las bases de datos de formas de mano, con el objetivo de complementar desde otro enfoque el estudio de modelos y algoritmos de clasificación para bases de datos con pocos datos etiquetados. Estas investigaciones son llevadas a cabo en marco de una tesis doctoral financiada por la UNLP a través de una beca de postgrado.

1.4. Aplicación web para reconocimiento de lengua de señas y crowdsourcing

Esta línea de investigación responde, a través del desarrollo de una aplicación web, a dos objetivos específicos. Por un lado, facilitar el acceso a los modelos de traducción para ayudar en la práctica y comunicación a través de lengua de señas. Por otro, la recopilación de videos para ampliar las bases de datos con los que se entrenan dichos modelos a través del *crowdsourcing*. Esto busca subsanar el hecho de que, al disponer de bases de datos obtenidas *in the wild* (por fuera del laboratorio), muchas veces la distribución de los datos es irregular, presentando muchas señas, frases o palabras que aparecen muy pocas veces a lo largo de toda la base de datos. En función de esto, la aplicación web sugiere a los usuarios frases puntuales a interpretar y permite la grabación de un video corto del usuario interpretando LSA a través de la

cámara de cualquier dispositivo electrónico. Finalmente, muestra la traducción de dicho video a texto en castellano generado por un modelo de SLT. Dicho modelo es ejecutado en el mismo navegador a través de Tensorflow-JS y funciona a través de la información posicional extraída de los videos a través de otro modelo de la librería Mediapipe [5]. Los videos grabados, con previa autorización, se almacenan con la intención de ser utilizados para extender dichas bases de datos. Según el video, esto puede requerir de un preprocesamiento o de la verificación de un intérprete experto.

1.5. Interpretabilidad de modelos de Deep Learning

La interpretabilidad es un campo de investigación que tiene como objetivo estudiar técnicas para comprender los estímulos por los cuales modelos de IA de caja negra generan determinadas salidas. Se utiliza con un rol diagnóstico para descubrir cómo es la contribución de las capas ocultas en los modelos.

En el marco de esta línea de investigación se realizó una implementación de RISE (Randomized Input Sampling for Explanation) para Captum. RISE es un método post-hoc y local, concentrado en modelos de caja negra ya que no requiere acceso a los parámetros internos de los modelos. Este genera un mapa de importancia que, al aplicarlo a imágenes, permite visualizar las regiones de la misma que fueron importantes para la predicción de determinada clase. Captum es una librería de interpretabilidad de modelos para PyTorch que permite implementar más fácilmente los algoritmos de interpretabilidad que pueden interactuar con los modelos de PyTorch.

Para probar esta implementación, se utilizaron los conjuntos de datos EyePACS y EyeQ cuyas imágenes sirven para evaluar retinopatía diabética y la calidad de las imágenes para esta patología, respectivamente. Se realizaron pruebas con la implementación de RISE para interpretar los resultados que arroja un modelo de evaluación de calidad de imagen.

1.6. Métricas de Equivarianza

Las redes neuronales son modelos tradicionalmente considerados como de caja negra. En los años recientes, se han realizado varios esfuerzos para comprender su funcionamiento de forma tal que el mismo sea más predecible o modulable.

La invarianza y equivarianza a las transformaciones son propiedades deseables en varios modelos de redes debido a que nos permiten razonar más fácilmente respecto a su funcionamiento.

En los últimos años, varios modelos fueron propuestos para añadir invarianza a la rotación y otras transformaciones en CNNs [3]. No obstante, no está claro cómo estos modelos impactan en el aprendizaje usual de los pesos de la red.

Por este motivo, se continúa con la utilización de las métricas previamente definidas [4] para caracterizar modelos de redes neuronales típicos, ya sea desde capas muy utilizadas como Batch Normalization, Dropout, Max-Pooling, arquitecturas completas como Residual Networks, VGG o AllConvolutional, y arquitecturas especializadas como TI-Pooling.

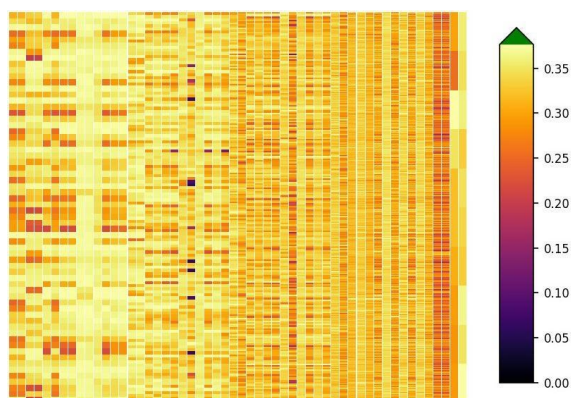


Figura 3. Invarianza por capas y unidades de una CNN con arquitectura ResNet.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Redes neuronales profundas, convolucionales y transformers para análisis de secuencia.
- Invarianza y auto-equivarianza en redes neuronales.
- Reconocimiento de lenguaje de señas.
- Generación de imágenes con GANs.
- *Crowdsourcing* para la extensión de bases de datos.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Comparación de modelos especializados para bases de datos con pocas muestras para la clasificación de formas de mano.
- Desarrollo de métricas de invarianza y auto-equivarianza para redes neuronales.
- Análisis de modelos para el reconocimiento y traducción de lengua de señas en video.
- Redes generativas para la creación de datos artificiales en la Lengua de Señas.
- LSA-T como primer base de datos continua de LSA.
- Sistema Web para Reconocimiento y *Crowdsourcing* de Lengua de Señas Argentina.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 3 profesores con dedicación exclusiva, 1 investigador CIC-PBA, 3 becarios de posgrado de la UNLP con dedicación docente, 2 becarios CIN y 1 profesor extranjero.

Dentro de los temas involucrados en esta línea de investigación, en los últimos dos años se han finalizado diversas tesinas de grado, como así también trabajos de final de carrera y de especialización.

Actualmente se están desarrollando 2 tesis de doctorado, 2 tesis de especialista, 1 de máster, 3 tesinas de grado de Licenciatura y 2 trabajos finales de Ingeniería en Computación. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Quiroga, F., Antonio, R., Ronchetti, R., Lanzarini, L., Rosete, A. A Study of Convolutional Architectures for Handshape Recognition applied to Sign Language, publicado en el XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017) (pp. 13-22). 2017
- [2] Cornejo Fandos, U., Ríos, G., Ronchetti, F., Quiroga, F., Hasperué, W., Lanzarini, L. Recognizing Handshapes using Small Datasets, publicado en el XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019, Rio Cuarto) (pp. 105-114). 2019.
- [3] Quiroga F., Ronchetti F., Lanzarini L., Fernandez-Bariviera A. Revisiting Data Augmentation for Rotational Invariance in Convolutional Neural Networks. International Conference on Modeling and Simulation in Engineering, Economics and Management (MS'2018 GIRONA). 2018.
- [4] Quiroga, F., Torrents-Barrena, J., Lanzarini, L., & Puig, D. Measuring (in) variances in Convolutional Networks. In Conference on Cloud Computing and Big Data (pp. 98-109). Springer, Cham. 2019.
- [5] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M.. Mediapipe: A framework for building perception pipelines. arXiv preprint arXiv:1906.08172. 2009.
- [6] Goodfellow I. J., Pouget-Abadie j., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative Adversarial Networks. NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems. v2. pp 2672-2680. 2014.
- [7] Jaschek M., Slettebak A., Jaschek C. *Be star terminology*. Be Star Newsletter. 1981.
- [8] Potash, P., Romanov, A., and Rumshisky, A. Ghostwriter: Using an lstm for automatic rap lyric generation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 1919-1924. 2015.
- [9] Young, T., Hazarika, D., Poria, S., and Cambria, E. Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, 13(3):55. 2018.
- [10] Dal Bianco, P., Ríos, G., Ronchetti, F., Quiroga, F., Stanchi, O., Hasperué, W., & Rosete, A. (2023, January). LSA-T: The First Continuous Argentinian Sign Language Dataset for Sign Language Translation. In *Advances in Artificial Intelligence—IBERAMIA 2022: 17th Ibero-American Conference on AI*, Cartagena de Indias, Colombia, November 23–25, 2022, Proceedings (pp. 293-304). Cham: Springer International Publishing.
- [11] Iván Mindlin, Facundo Quiroga, Franco Ronchetti, Pedro Dal Bianco, Gastón Ríos, Laura Lanzarini, Waldo Hasperué. “A Comparison of Neural Networks for Sign Language Recognition with LSA64”. JCC-BD&ET: Conference on Cloud Computing, Big Data & Emerging Topics. pp 104-117. Springer, Cham. Junio 2021.

INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA PARA INNOVAR Y DINAMIZAR PROCESOS

Esnaola L. M.¹, Tessore J. P.², Agesilao A. T.³, Citate Gómez I. M.³, Moyano M. E.³
Jaszczyszyn A. G.⁴, Ramón H. D.⁵

Instituto de Investigación y Transferencia en Tecnología (ITT)
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)
Sarmiento Nro. 1169 2° Piso, Junín (B) – TE: (0236) 4407750 INT 11610

{leonardo.esnaola, juanpablo.tessore, agustin.agesilao, ignacio.citate, emilia.moyano,
adrian.jaszczyszyn, hugo.ramon}@itt.unnoba.edu.ar

RESUMEN

En su definición más genérica, un proceso puede conceptualizarse como una serie de tareas interrelacionadas que, juntas, transforman las entradas en salidas. Estas tareas pueden ser llevadas a cabo por personas; por la naturaleza; por máquinas, o por combinaciones de ellas, y pueden ser simples o extremadamente complejas. La Inteligencia Artificial (IA), por su parte, se ocupa de construir máquinas o sistemas inteligentes capaces de realizar tareas que normalmente requieren inteligencia humana. Este proyecto propone identificar, analizar, seleccionar e intervenir procesos, utilizando herramientas y técnicas de la IA, para favorecer su optimización.

Las líneas de investigación que forman parte de este proyecto abarcan temáticas como la generación automática de código y su impacto en el proceso de desarrollo de software; la computación afectiva y sus aplicaciones; y los sistemas autónomos y su potencialidad.

Palabras clave: PROCESOS; IA; OPTIMIZACIÓN.

CONTEXTO

Las líneas de investigación presentadas a continuación se enmarcan en el proyecto de investigación "Inteligencia Artificial como herramienta para innovar y dinamizar procesos", con lugar de trabajo en el Instituto de Investigación y Transferencia en Tecnología (ITT), presentado en la convocatoria a Subsidios de Investigación Bianuales (SIB) 2022, aprobado y financiado por la Secretaría de Investigación, Desarrollo y Transferencia (SIDT) de la UNNOBA.

Como hipótesis de trabajo se plantea la existencia de actividades cuyos procesos presentan gran potencial para ser mejorados aplicando herramientas y técnicas de la IA. El objetivo general del proyecto consiste en identificar, analizar, seleccionar e intervenir procesos utilizando herramientas y técnicas de la IA para favorecer su optimización.

¹ Docente Investigador ITT / Doctorando UNLP

² Docente Investigador ITT / Doctorando UNLP / Becario CONICET

³ Becario ITT

⁴ Docente Investigador ITT / Maestrando UNLP

⁵ Docente Investigador ITT

1. INTRODUCCIÓN

El proyecto presentado propone el uso de la Inteligencia Artificial (IA) como una herramienta para innovar y dinamizar procesos. Cabe preguntarse entonces en qué tipo de procesos es posible utilizar IA para motorizar esta innovación.

Para abordar este interrogante, resulta conveniente desambiguar el significado del término “proceso”. Desde el punto de vista de la ingeniería, podría definirse a “proceso” [1] como “*una serie de tareas interrelacionadas que, juntas, transforman las entradas en salidas*”. Las tareas, a las que la definición hace referencia, podrían ser llevadas a cabo por personas; por la naturaleza; por máquinas, o por combinaciones de ellas, y pueden ser simples o muy complejas. Los procesos a los que se hace referencia en este trabajo son aquellos que transforman a las entradas en salidas, y que permiten resolver todo tipo de problemas en las más diversas industrias [2-3].

En tanto, la IA constituye un campo en permanente desarrollo, y cuya evolución parece darse cada vez en tiempos más acotados. Lo que hasta hace poco tiempo parecía imposible para los sistemas que emplean IA, hoy ya es posible. Por mencionar un ejemplo concreto, ¿quién hubiera pensado que un proceso tan creativo y humano como componer música o pintar obras de arte pudiera ser llevado a cabo por una computadora? Actualmente, la IA generativa permite llevar a cabo este tipo de tareas, pudiendo mencionar MuseNet [4] en el caso de la música, o Deep Dream Generator [5] en el caso de las pinturas, pero se trata sólo de dos ejemplos entre muchos otros.

En este sentido, resulta razonable pensar que la IA puede configurarse como una herramienta muy potente y versátil, capaz de facilitar la

innovación en los más diversos procesos, habida cuenta de que, en última instancia, todos los procesos que podamos imaginar, con excepción de los naturales, han sido diseñados por humanos y son ejecutados por ellos, o por máquinas que ellos mismos construyen.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se presentan las principales líneas de investigación del proyecto:

Línea 1: Estudiar el impacto que la generación automática de código provoca en el proceso de desarrollo de software.

A medida que la tecnología avanza, también lo hace la forma de programar. Es decir, el mecanismo a través del cual se le indica al hardware cómo debe operar para trabajar sobre los datos de entrada, procesarlos y producir una salida. Programar, o codificar, es sólo una de las etapas en el desarrollo de software, aunque es una etapa muy importante. En este sentido, la ingeniería de software es la disciplina que “[...] *comprende todos los aspectos de la producción de software desde las etapas iniciales de la especificación del sistema, hasta el mantenimiento de éste después de que se utiliza*” [6].

Cabe preguntarse si las técnicas de programación acompañan de forma directa al avance de la tecnología. Indudablemente ha habido cambios, en los 70s se abandonaba la programación a través de tarjetas perforadas y se pasaba a una programación textual, utilizando computadoras más modernas (para la época) y lenguajes de programación de bajo nivel. En las décadas o años posteriores, comenzaron a utilizarse lenguajes de programación de más alto nivel y otros paradigmas de la programación, como la programación orientada a objetos. Se comenzó a compartir y a reutilizar bibliotecas de código, *frameworks* y hubo una evolución en las interfaces de usuario, muchos de estos últimos

avances fueron motorizados por la aparición de Internet. Sin embargo, pese a todos estos avances, no se notó un gran salto en las técnicas de programación como el que ocurrió al abandonar las tarjetas perforadas. Más recientemente, ese cambio parece venir de la mano de la IA y el impacto que ésta puede tener en una etapa indispensable del desarrollo de software que parece algo estancada [7].

Otra pregunta relevante que surge a raíz de este impacto es, ¿dejarán de existir los programadores cuando la IA más evolucionada se ocupe de realizar esta tarea? Tal vez es una pregunta excesivamente taxativa, es difícil vaticinar esa extinción, pero sí es esperable que, así como ha ocurrido con otros empleos y ocupaciones, el impacto de la tecnología provoque cambios profundos en la forma de hacer las cosas y los puestos deban adaptarse para acompañar esos cambios. Sin embargo, muchos sostienen, y con razón, que desarrollar software es mucho más que sólo programar [8].

La línea presentada buscará estudiar ese impacto, pero no limitándose a la etapa de programación en sí, sino al soporte que la ingeniería de software y sus herramientas brinda (o no) a esta nueva forma de programar.

Línea 2: La Computación Afectiva (CA) y sus aplicaciones.

Cuando se habla de CA, se hace referencia al estudio y el desarrollo de sistemas y dispositivos que pueden reconocer, interpretar, procesar y estimular las emociones humanas [9]. Aquí se utiliza el término “afecto” como sinónimo de “emoción”. Las tecnologías de CA pueden percibir el estado emocional de los usuarios y actuar en consecuencia [10].

Cabe preguntarse, ¿qué aspectos en nuestras interacciones pueden “delatar” nuestro estado emocional? La respuesta a esta pregunta combina el aporte de varias disciplinas, como

por ejemplo la psicología, la sociología, la antropología, etc. Por mencionar algunos ejemplos cercanos pueden señalarse:

- El tono de voz de una persona no es el mismo cuando está enojada que cuando está alegre o triste [11]. Desde el punto de vista de la CA, un sistema podría procesar la señal de sonido correspondiente a nuestra voz, compararla con muestras base y determinar así nuestro estado de ánimo.
- Las expresiones faciales son también otro aspecto que puede permitir inferir el estado emocional. Ya en 1872 Darwin [12] concluía en sus análisis que la expresión de ciertas emociones es innata y universal en los seres humanos. Resulta posible entonces asociar las expresiones faciales con alguna de las emociones básicas que, según Ekman [13], son las siguientes: ira, asco, miedo, alegría, tristeza y sorpresa. Desde el punto de vista de la CA, un sistema podría procesar las imágenes de rostros, analizar sus expresiones y determinar así el estado de ánimo.
- La escritura también puede delatar nuestro estado emocional. Esto puede conseguirse a través de la grafología [14] (aunque se limita a textos manuscritos) o, más en general, analizando el contenido del texto escrito [15]. El procesamiento del lenguaje natural (PLN) [16], en general, y el análisis de sentimientos [17-19], en particular, se ocupan de esta tarea y permiten determinar la preeminencia de alguna de las emociones básicas subyacentes a partir del análisis del texto.

Sus aplicaciones son diversas y los sistemas de CA podrían integrarse en diversos procesos consiguiendo mejorarlos y potenciarlos. Por ejemplo, el proceso de enseñanza podría adaptarse al aprendizaje del estudiante analizando su estado emocional [20-21], procurando determinar la aprehensión de los contenidos tratados; en un canal de atención a

clientes [22-23], ya sea que esta interacción se dé de forma escrita, a través de voz o incluso de forma presencial, sistemas de CA podría determinar el estado emocional y propiciar acciones para actuar en consecuencia; en un aeropuerto, sistemas de CA podrían analizar en tiempo real tanto las expresiones faciales de los viajeros [24], como su comportamiento, buscando identificar sujetos de interés para un análisis más minucioso, por nombrar algunos ejemplos.

Línea 3: Los Sistemas Autónomos y su potencialidad.

Los sistemas autónomos [25], también conocidos como sistemas autónomos, conectan el mundo de los bits con el mundo de los átomos. La inteligencia es capaz de percibir, entender y tomar decisiones en el mundo real. Si se consideran los productos que potencian el mundo que nos rodea puede advertirse que están diseñados para ser operados por humanos, son básicos, y están codificados estrictamente para realizar una función rígida sin inteligencia innata. Como humanos, aprendemos cómo operarlos y tolerar sus limitaciones. Pero, ¿qué pasaría si pudiéramos enseñarles a modificar su comportamiento para adaptarse a nuestras necesidades? Como humanos, enseñamos (y aprendemos) todo el tiempo, y esta misma técnica puede utilizarse también para enseñarle a una máquina. Actualmente eso es posible [26], y los algoritmos inteligentes que operan estas máquinas podrían aprender a través de ejemplos, tal como nosotros solemos aprender.

Los sistemas autónomos utilizan técnicas de IA que les permiten aprender a través de la interacción con el mundo real. En este sentido, el aprendizaje por refuerzo resulta de gran relevancia. Bajo este enfoque, una máquina puede realizar una tarea intentando e intentando hasta conseguir dominarla. Los sistemas autónomos existen en el mundo físico

con nosotros, por lo que enseñarles puede ser costoso y potencialmente riesgoso. Generalmente se requieren millones de muestras de datos para poder entrenar a una IA a realizar una tarea. Entonces, ¿cómo puede resolverse esta tarea? Una aproximación consiste en crear una réplica digital del mundo y un agente inteligente capaz de aprender dentro de este entorno. Así, este agente puede intentar una tarea una y otra vez, y fallar, hasta que la domine, sin afectar al mundo real. Sin embargo, este enfoque sólo funciona si somos capaces de representar adecuadamente todos los aspectos del mundo físico y, dado que se trata sólo de software, es posible crear cientos o miles de instancias de esta realidad sintética simultáneamente. Es más, como esta realidad no está limitada por la restricción temporal que sí afecta a nuestra realidad, podemos acelerarla. Así, en relativamente poco tiempo humano, se podría entrenar una IA en un inmenso número de potenciales situaciones. Entonces, ¿qué sucedería si, antes de tomar una acción en el mundo real, los sistemas autónomos pudieran analizar miles de posibles cursos de acción en una realidad sintética paralela, examinar los resultados posibles y seleccionar la mejor alternativa para ejecutar en el mundo real? [27]

3. RESULTADOS OBTENIDOS/ESPERADOS

En la convocatoria SIB 2022 se presentaron y aprobaron diversos proyectos que abarcan diferentes disciplinas. En este sentido, este equipo de trabajo ha mantenido encuentros, y está previsto continuar manteniéndolos, con referentes de algunos de los demás proyectos, buscando definir puntos de contacto que permitan enriquecer los mismos, a través de un trabajo interdisciplinar.

En trabajos anteriores, que se vinculan con las líneas de trabajo del proyecto, se ha avanzado en Procesamiento de Lenguaje Natural (PLN) a

partir de textos informales extraídos de redes sociales [28-31].

Entre otras aplicaciones, se espera desarrollar un asistente virtual de interfaz de voz de usuario orientado al sector agropecuario. Además, se espera estudiar el impacto de la generación automática de código en el desarrollo de aplicaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por investigadores formados y en proceso de formación, becarios del Sistema Nacional de Ciencia y Tecnología, becarios alumnos, graduados e investigadores externos.

Se han presentado planes de trabajo para becas, las cuales se encuentran actualmente en proceso de evaluación.

Se esperan finalizar, además, dos tesis doctorales, una tesis de maestría y presentarse a una convocatoria a beca postdoctoral del CONICET.

5. BIBLIOGRAFÍA

1. EIA. (7 de enero de 1999). ANSI/EIA-632-1998 Processes for Engineering a System, Appendix A, p66. Accedido el 23 de febrero de 2023.
2. Yarali, A. (2022). Digital Transformation Trends in the Automotive Industry.
3. Singh, P., Manjunatha, A. S., Baig, A., Dhopeswar, P., Huo, H., Bharathy, G., & Prasad, M. (2022). Application of Artificial Intelligence in Healthcare by Industries in Australia: Opportunities and Challenges. En International Conference on Intelligent Vision and Computing (pp. 568-580). Springer, Cham.
4. OpenAI (s.f.). Musenet.
5. Deep Dream Generator (s.f.). Deep Dream Generator.
6. Sommerville, I. (2005). Ingeniería del software. Pearson educación.
7. Gorosterrazu, G. (3 de abril de 2022). Por qué es momento de abandonar la programación tradicional. ámbito.com.
8. Slotinsky, D. (31 de marzo de 2022). Por qué puede morir el puesto mejor pago de Argentina: el cambio que puede dejar a miles sin empleo. El Cronista.
9. Picard, R. W. (2000). Affective computing. MIT press.
10. Banafa, A. (1 de agosto de 2018). ¿Qué es la computación afectiva?
11. Cabrelles Sagredo, M. S. (s.f.). La Influencia de las Emociones en el sonido de la voz. Biblioteca Virtual Miguel de Cervantes.
12. Darwin, C. (1984). Trad. esp.: La Expresión de las emociones en los animales y en el hombre. Alianza.
13. Ekman P, Friesen W V. Constants across cultures in the face and emotion. J Pers Soc Psychol. 1971;17(2):124–129.
14. Roy, C., Dey, R., Chaudhuri, C., & Das, D. (2019, Diciembre). Emotion Predictor Using Social Media Text and Graphology. En 2019 IEEE 9th International Conference on Advanced Computing (IACC) (pp. 96-102). IEEE.
15. Munezero, M., Montero, C., Sutinen, E. and Pajunen, J., 2014. Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text. IEEE Transactions on Affective Computing, 5(2), pp.101-111.
16. Khan, M., Durrani, M., Ali, A., Inayat, I., Khalid, S. and Khan, K., 2016. Sentiment analysis and the complex natural language. Complex Adaptive Systems Modeling, 4(1).
17. Poria, S., Cambria, E., Bajpai, R. and Hussain, A., 2017. A review of affective computing: From unimodal analysis to

- multimodal fusion. *Information Fusion*, 37, pp.98-125.
18. Rouast, P., Adam, M. and Chiong, R., 2021. Deep Learning for Human Affect Recognition: Insights and New Developments. *IEEE Transactions on Affective Computing*, 12(2), pp.524-543.
 19. Shoumy, N., Ang, L., Seng, K., Rahaman, D. and Zia, T., 2020. Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals. *Journal of Network and Computer Applications*, 149, p.102447.
 20. Newman, H., & Joyner, D. (2018, Junio). Sentiment analysis of student evaluations of teaching. En *International conference on artificial intelligence in education* (pp. 246-250). Springer, Cham.
 21. Sangeetha, K., & Prabha, D. (2021). Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. *Journal of Ambient Intelligence and Humanized Computing*, 12(3), 4117-4126.
 22. Al-Otaibi, S., Alnassar, A., Alshahrani, A., Al-Mubarak, A., Albugami, S., Almutiri, N., & Albugami, A. (2018). Customer satisfaction measurement using sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 9(2), 106-117.
 23. Kumar, S., Yadava, M., & Roy, P. P. (2019). Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *information fusion*, 52, 41-52.
 24. Kim, C. M., Kim, K. H., Lee, Y. S., Chung, K., & Park, R. C. (2020). Real-time streaming image based PP2LFA-CRNN model for facial sentiment analysis. *IEEE Access*, 8, 199586-199602.
 25. Gartner. (2021). Las principales tendencias tecnológicas estratégicas de Gartner para 2022 [Ebook].
 26. Luna Macias, A. (25 de noviembre de 2020). Recicla-IA: Sistema para clasificación de Residuos Sólidos Urbanos e implementación con sistema robótico. Medium.
 27. Microsoft. (25 de enero de 2021). The future of autonomous systems. YouTube.com.
 28. Juan Pablo Tessore, Leonardo Martín Esnaola, Claudia Cecilia Russo, and Sandra Baldassarri. 2019. Comparative analysis of preprocessing tasks over social media texts in Spanish. In *Proceedings of the XX International Conference on Human Computer Interaction (Interacción '19)*. Association for Computing Machinery, New York, NY, USA, Article 27, 1–8. <https://doi.org/10.1145/3335595.3335632>
 29. L. Esnaola, J. P. Tessore, H. Ramón and C. Russo, "Effectiveness of preprocessing techniques over social media texts for the improvement of machine learning based classifiers," 2019 XLV Latin American Computing Conference (CLEI), Panama, Panama, 2019, pp. 1-10, doi: 10.1109/CLEI47609.2019.235076.
 30. Tessore, J.P., Esnaola, L.M., Lanzarini, L. et al. Distant Supervised Construction and Evaluation of a Novel Dataset of Emotion-Tagged Social Media Comments in Spanish. *Cogn Comput* 14, 407–424 (2022). <https://doi.org/10.1007/s12559-020-09800-x>
 31. Tessore, J.P., Esnaola, L.M., Ramón, H.D. et al. Contextual information usage for the enhancement of basic emotion classification in a weakly labelled social network dataset in Spanish. *Multimed Tools Appl* (2022). <https://doi.org/10.1007/s11042-022-13750-x>

Detección de *Ransomware* en *Blockchains*

Yanina Ditz, Gabriela Minetti

Laboratorio de Investigación en Sistemas Inteligentes (LISI)

Facultad de Ingeniería - Universidad Nacional de La Pampa

e-mail: {yaninaditz, minettig}@ing.unlpam.edu.ar

Resumen El *ransomware* es un código malicioso que bloquea el acceso a los datos de un dispositivo, exigiendo un rescate para permitir el acceso nuevamente. El rescate suele pagarse en *Bitcoin*. Según el FBI, a fines de 2015, las víctimas habían pagado US\$27 millones en rescates a los atacantes. A partir de este problema, surgen sistemas para la detección de *ransomware* conocido y desconocido, utilizando el modelo topológico del análisis de datos, y técnicas de minerías de datos como *Random Forest* (RF), *Support Vector Machine* (SVM), Regresión logística, KNN y Análisis Discriminante lineal (LDA), *Naive Bayes* (NB), árboles de decisión (CART), *Bagging* (BGC), Ascenso de gradiente (GGC), entre otros. A partir del estudio de estas herramientas de aprendizaje automático se profundiza en la implementación y análisis de las redes neuronales para aprender a reconocer patrones y comportamientos comunes en las transacciones de *Bitcoin*.

Palabras claves: *ransomware*, *Blockchain*, *Bitcoin*

CONTEXTO

Desde su creación en 1998, el Laboratorio de Investigación de Sistemas Inteligentes (LISI) de la Facultad de Ingeniería (FI) de la Universidad Nacional de La Pampa (UNLPam), se ha abocado al estudio de algoritmos cada vez más eficientes para la solución de problemas complejos, tanto de optimización como de diseño. Esta línea de investigación se desarrolla en el marco del proyecto de investigación: "Big data optimization con algoritmos metaheurísticos utilizando frameworks de computación distribuida", acreditado por la FI y dirigido por la Dra. Salto. Su objetivo principal consiste en obtener algoritmos inteligentes que permitan obtener soluciones de calidad a las pro-

blemas planteadas con bajo esfuerzo computacional.

I. INTRODUCCIÓN

El *ransomware* es un tipo de *malware* (código malicioso) que se propaga a sí mismo y causa serios problemas de seguridad mediante el uso de encriptación para almacenar o eliminar datos en una computadora o red. Es decir, bloquea la computadora o impide acceder a los datos, utilizando el cifrado de clave privada hasta que se pague un rescate. En la actualidad este rescate generalmente se paga en *Bitcoin*[1]. Si bien la extorsión basada en datos se ha propagado a partir del 2005, el desarrollo del software de cifrado de rescate y *Bitcoins* ha facilitado enormemente este esquema delictivo [2].

Los ataques de *ransomware* se producen tanto en computadoras personales, como en teléfonos celulares al cambiar el número PIN del dispositivo y luego requerir un rescate para obtener el nuevo PIN [2]. El *ransomware* se puede dividir en dos tipos básicos. El más común es el *ransomware* criptográfico, que encripta los archivos y datos. El segundo tipo es el *ransomware* del bloqueo. Esta versión bloquea la computadora u otro dispositivo, evitando que las víctimas la usen [3]; generalmente los datos almacenados en el dispositivo no se modifican.

A finales de los 90 y hasta 2005, los métodos de pago en línea no estaban disponibles tan fácilmente. Las víctimas recibieron instrucciones de pagar rescates a través de mensajes de texto SMS o enviando tarjetas prepagas por correo [3]. Todos estos métodos de pago eran riesgosos, ya que un investigador determinado podría rastrear al atacante. El *ransomware* realmente despegó cuando en 2008 *Bitcoin* entró en uso. *Bitcoin* es una moneda electrónica que es mucho más difícil

de rastrear y, por lo tanto, ayudó a anonimizar las transacciones. Eso hizo difícil o incluso imposible rastrear al atacante siguiendo el pago [4].

En el año 2005, aparece el primer *ransomware* moderno, conocido como GPCoder, fue creado usando el algoritmo RSA para encriptación y se difundió a través de un *e-mail spam*, en el cual ofrecían [1]. En el 2006 aparece el virus troyano Cryzip, que copiaba los archivos de datos a archivos de almacenamiento protegidos con contraseña, y una vez hecho esto se borraban los originales, que también utilizó RSA y encriptaba el directorio 'Mis documentos' de la víctima y le sugería que que compre productos de una página *online* determinada para obtener la clave. En el año 2008 aparece una versión de GPCoder, llamada GPCoder.AK, que dejaba un archivo de texto con instrucciones en cada subdirectorios encriptado [1].

A mitad del 2011, debido a la aparición de servicios de pagos anónimos, el *ransomware* dio el gran salto al duplicarse el número de delitos bajo esta modalidad. A comienzo de 2012, aparece Citadel, que es una herramienta que permitía la fácil distribución y el manejo de textitsoftbots, y así producir y distribuir de manera más simple el *ransomware*. En 2013, Kaspersky detectó un nuevo troyano móvil, Svpeng el cual se difundió por mensaje *spam*, afectando principalmente a clientes de bancos Rusos [5].

En agosto de 2013 aparece la pieza más famosa de *ransomware* creada por el *hacker* Slavik, denominada CryptoLocker, se difundía a través de archivos adjuntos de *e-mails* y descargas fraudulentas. Desde septiembre de 2013 a mayo de 2014, se estimó que más de 500000 víctimas fueron infectadas por CryptoLocker y que el 1.7% pagaron el rescate [1], [5]. Al finalizar el año 2015, el FBI estimó que las víctimas pagaron US\$27 millones a los atacantes detrás de CryptoLocker. También en el 2014 los ciber criminales introducen CryptoWall, que bloquea el sistema a través de un *e-mail* o una descarga fraudulenta, impidiendo el acceso a cualquier almacenamiento externo o unidades compartidas conectadas al objetivo.

En el 2016 se introdujo otra versión de *ransomware*, denominada Cerber, que afectó un gran número de computadoras empresariales [1], [5]. Actualmente, las 3 versiones de *ransomware* principales son CryptoWall, CTB-Locker (curve-Tor-

Bitcoin) y TorrentLocker. Tanto CryptoWall como CTB-Locker infectan programas de ventas. TorrentLocker recolecta direcciones de correo electrónico cuando infecta una computadora para enviar *spam* a otros usuarios [1], [5]. Todo lo expresado en esta sección motiva la línea de investigación orientada al análisis y desarrollo de *software* para la detección de *ransomware*.

II. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

La detección de *ransomware* requiere de métodos que permitan manejar, analizar y obtener información de la transacción de *Bitcoin*, para identificar direcciones fraudulentas. Antes de seguir con la detección de *ransomware* en *blockchains* de *bitcoins* es necesario explicar esto último. Una cadena de bloques (*blockchain*) es una lista de bloques en constante crecimiento que almacena una colección de registros (transacciones) [6]. Los bloques contienen un *hash*, un enlace al bloque anterior, una marca de tiempo (*timestamp*) y datos de transacciones, protegidos mediante criptografía. Los bloques se distribuyen en la red, ya que son ejecutados y controlados por muchos pares de puntos o nodos, y están descentralizados, ya que no hay una autoridad central que los controle. En cadenas de bloques públicas como *Bitcoin* y *Ethereum*, cualquier punto puede convertirse en validador o minero, y cualquiera puede descargar y almacenar una copia de su historial completo. El término *broadcasted* se refiere a que una vez minado un bloque, la información de este evento es enviada a todos los nodos de la red y que la(s) transacción(es) ha(n) sido confirmada(s), irreversible, visible para el público [7].

La detección de *ransomware* se lleva a cabo a través de diferentes técnicas realizan tareas como seleccionar y extraer datos, procesarlos, obtener su valor, visualizarlos y presentarlos. Una técnica es el modelo topológico del análisis de datos, cuya idea fundamental es extraer patrones ocultos de datos a través del análisis sistemático de formas de datos tales como, ciclos, cuantificados en varias escalas de resolución y, así, poder detectar y predecir transacciones de *bitcoins* con *ransomware* [8].

Si bien lo antes expuesto mejoró la detección de *ransomware*, todavía es necesario superar el enfoque de detección de *ransomware* estático y basado en firmas. Por ende, se ha propuesto el

sistema dinámico de detección de *ransomware* utilizando técnicas de minería de datos como *Random Forest* (RF), *Support Vector Machine* (SVM), Regresión logística, KNN y Análisis Discriminante lineal (LDA), *Naive Bayes* (NB), árboles de decisión (CART), *Bagging* (BGC), Ascenso de gradiente (GGC), entre otros algoritmos para detectar *ransomware* conocido y desconocido [9], [10], [11], [12], [13].

A partir del análisis de las herramientas mencionadas en el párrafo anterior, en esta línea de investigación, se analiza la utilización de las redes neuronales para aprender a reconocer patrones y comportamientos comunes en los datos de transacciones de *Bitcoin*, con la intención de detectar anomalías que pueden indicar la presencia de *ransomware*.

III. RESULTADOS ENCONTRADOS Y ESPERADOS

De esta línea de investigación se espera analizar y adaptar una red neuronal para la detección dinámica de *ransomware* en *blockchains* de *Bitcoins*. Inicialmente, el alcance del proyecto de investigación actual es proporcionar un prototipo rápido. Luego, para la próxima versión del modelo, se planea probar pasos adicionales en la metodología para mejorar el rendimiento del mismo al aumentar el conjunto de datos sin clasificar; resolver las limitaciones del conjunto de datos categorizados ya sea orgánicamente (encontrar más observaciones) o sintéticamente aumentando las categorías submuestreadas; introducir técnicas automáticas para la reducción de la dimensionalidad en lugar de extracción e ingeniería manual de características; normalización de los atributos; y ajuste de parámetros de los algoritmos. Esto implica (i) evaluar los modelos desarrollados tal que permitan comparar los resultados obtenidos con respecto a enfoques existentes, y (ii) documentar el trabajo realizado mediante la escritura de informes técnicos y/o la publicación de artículos de eventos científicos.

IV. FORMACIÓN DE RECURSOS HUMANOS

Con esta línea de investigación se propende a la formación, de la primera autora, en investigación

siguiendo el método científico. Además, este trabajo le permitirá llevar a cabo gran parte de la investigación requerida en su tesis de maestría.

REFERENCES

- [1] R. Richardson and M. M. North, "Ransomware: Evolution, mitigation and prevention," *International Management Review*, vol. 13, no. 1, p. 10, 2017.
- [2] N. Aldaraani and Z. Begum, "Understanding the impact of ransomware: A survey on its evolution, mitigation and prevention techniques," in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, 2018, pp. 1–5.
- [3] I. A. Chesti, M. Humayun, N. U. Sama, and N. Jhanjhi, "Evolution, mitigation, and prevention of ransomware," in *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*, 2020, pp. 1–6.
- [4] S. Maniath, P. Poornachandran, and V. Sujadevi, "Survey on prevention, mitigation and containment of ransomware attacks," in *Security in Computing and Communications: 6th International Symposium, SSCC 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers 6*. Springer, 2019, pp. 39–52.
- [5] M. N. Alenezi, H. K. Alabdulrazzaq, A. A. Alshaher, and M. M. Alkharang, "Evolution of malware threats and techniques: a review," *Int. j. commun. netw. inf. secur.*, vol. 12, no. 3, 2022.
- [6] H. Yin and R. Vatrapu, "A first estimation of the proportion of cybercriminal entities in the bitcoin ecosystem using supervised machine learning," 12 2017, pp. 3690–3699.
- [7] A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder, *Bitcoin and cryptocurrency technologies: a comprehensive introduction*. Princeton University Press, 2016.
- [8] C. G. Akcora, Y. Li, Y. R. Gel, and M. Kantarcioglu, "Bitcoinheist: Topological data analysis for ransomware detection on the bitcoin blockchain," *arXiv preprint arXiv:1906.07852*, 2019.
- [9] F. Karbalaie, A. Sami, and M. Ahmadi, "Semantic malware detection by deploying graph mining," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, no. 1, p. 373, 2012.
- [10] A. Kharraz, S. Arshad, C. Mulliner, W. K. Robertson, and E. Kirda, "Unveil: a large-scale, automated approach to detecting ransomware," in *USENIX Security symposium*, vol. 25. Austin, Texas, 2016.
- [11] Z.-G. Chen, H.-S. Kang, S.-N. Yin, and S.-R. Kim, "Automatic ransomware detection and analysis based on dynamic api calls flow graph," in *Proceedings of the International Conference on Research in Adaptive and Convergent Systems*, 2017, pp. 196–201.
- [12] H. Sun Yin and R. Vatrapu, "A first estimation of the proportion of cybercriminal entities in the bitcoin ecosystem using supervised machine learning," in *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 3690–3699.
- [13] P. Nerurkar, "Illegal activity detection on bitcoin transaction using deep learning," *Soft Computing*, pp. 1–18, 01 2023.

DESARROLLO DE SOLUCIONES TECNOLÓGICAS DE BAJO COSTO A PROBLEMÁTICAS EN EL SECTOR HORTÍCOLA

Mauro Salina¹, Lucía Osés¹, Braian Pezet¹, Patricio Medina¹, Matías Acciaio¹, Jérica Guzmán¹, Juan Salvatore¹, Martín Morales^{1,2}, Jorge Osio^{1,3}, Marcelo Cappelletti^{1,3}

¹ Programa TICAPPS, Univ. Nac. Arturo Jauretche, Florencio Varela (1888), Argentina.

² Centro UTN CODAPLI-FRLP, La Plata (1900), Argentina.

³ Grupo de Control Aplicado, Instituto LEICI (UNLP-CONICET), La Plata (1900), Argentina.

e-mail: mdsalina@unaj.edu.ar

Resumen

Las líneas de Investigación presentadas en este trabajo tienen como objetivo general el de generar herramientas tecnológicas (hardware-software) basadas en Internet de las Cosas (IoT) y aprendizaje automático, que permitan brindar soluciones innovadoras de avanzada y de bajo costo, y que contribuyan a remediar las principales problemáticas que posee el sector hortícola en el territorio de la Universidad Nacional Arturo Jauretche (UNAJ). Específicamente, se propone el desarrollo e implementación experimental de un sistema IoT de monitoreo y control escalable, completamente autónomo independiente del usuario, y de bajo costo de realización, para ser instalado en huertas de la región y que permita obtener información precisa en tiempo real de magnitudes climatológicas, del estado del suelo, del agua y del nivel de iluminación. Además, se planea incorporar al sistema modelos basados en aprendizaje automático para que, en base a la información recolectada por los sensores, el sistema pueda de manera inteligente tomar decisiones apropiadas en pos de obtener el rendimiento óptimo de los recursos existentes y mejorar el desarrollo y producción de los cultivos intensivos, así como también, posibilite detectar y/o predecir distintas plagas que pueden afectar a los cultivos. Adicionalmente, se pretende desarrollar un sistema basado en aprendizaje profundo aplicado al procesamiento de imágenes para la detección, reconocimiento y clasificación de diferentes malezas que pueden afectar a la producción y a los procesos de

cosecha de los cultivos de la región. Finalmente, se propone desarrollar un sistema de control y gestión de energía basado en recursos renovables (solar y eólico), que pueda ser aplicado en huertas de la zona de influencia de la Universidad Nacional Arturo Jauretche (UNAJ).

La tarea a desarrollar posee una demanda potencialmente alta, dado que entre los principales beneficiarios se encuentran todos aquellos pequeños y medianos productores, no solo de la región cercana a la UNAJ, sino también de localidades vecinas como La Plata, Brandsen, San Vicente, entre otras.

Se debe destacar además que estas líneas de investigación impulsan la formación de recursos humanos en el área de IoT y aprendizaje automático, tanto de jóvenes docentes investigadores como de estudiantes de la UNAJ.

Palabras clave: *Internet de las Cosas, Aprendizaje automático, sector hortícola.*

Contexto

Las líneas de I/D presentadas en este trabajo están incluidas dentro del Programa TICAPPS (TIC en aplicaciones de interés social) de la Universidad Nacional Arturo Jauretche (UNAJ), Resolución N° 064/17, bajo la dirección del Dr. Ing. Martín Morales.

1. Introducción

La Universidad Nacional Arturo Jauretche (UNAJ) se localiza en el partido de Florencio

Varela y tiene como área de influencia directa a los partidos de Berazategui, Quilmes y Almirante Brown. Estos partidos se encuentran en la denominada zona sur del Conurbano Bonaerense de la Provincia de Buenos Aires. El partido de Florencio Varela es uno de los municipios con mayor superficie rural dentro del Conurbano Bonaerense. Actualmente cuenta aproximadamente con unos 500 pequeños y medianos productores hortícolas, florícolas y frutícolas. En todos los casos, esta actividad está basada fundamentalmente en los cultivos intensivos, donde se busca maximizar la producción en espacios reducidos, utilizando un solo tipo de producto. En particular, dentro de la producción hortícola, las modalidades de producción son tanto a campo abierto como bajo cubierta (invernadero).

Una de las características sobresalientes de esta zona rural es que limita con áreas de similares características, con los partidos vecinos, lo que transforma a la subregión en estratégica para la producción de alimentos frescos de origen vegetal, tanto para el consumo directo como para la industrialización.

En los últimos años, la solidez del sistema productivo de los pequeños y medianos productores agroindustriales de la región, se ha visto resentida por los vaivenes de la economía y en algunos casos por condiciones climáticas adversas que afectaron en mayor proporción a los cultivos realizados a la intemperie, con el consiguiente aumento de los precios como consecuencia de la disminución de la oferta.

Esta actividad requiere, entre otras cosas, de la disponibilidad oportuna de datos, tanto del clima como del suelo y del agua, con el propósito de obtener cultivos sanos y buenos rendimientos. Sin embargo, la mayoría de los productores de esta región, presentan inconvenientes tales como:

- falta de información precisa, actual e histórica, de magnitudes climatológicas, del estado del suelo, del agua y del nivel de iluminación.
- insuficiente innovación tecnológica, dado que existe una carencia importante en cuanto a

la implementación de la tecnología informática de avanzada;

- mano de obra sin adecuada calificación, dado que muchas veces el problema no es la ausencia de información o de equipamiento tecnológico para este sector, sino el desconocimiento de la tecnología existente y de la utilización de la información brindada por ella;
- falta de una infraestructura adecuada;
- falta de capacitación en planificación y gestión;
- vulnerabilidad en el proceso de comercialización.

A su vez, los cultivos pueden ser afectados por diferentes amenazas, como por ejemplo las diversas malezas y/o plagas que pueden afectar a la producción y a los procesos de cosecha, ya que en algunos casos esto es un impedimento para acceder a los propios cultivos. Existe una gran diversidad de malezas que aparecen dependiendo principalmente de la zona geográfica y de la estación del año. La aparición de estos yuyos viene acompañada con la resistencia a herbicidas y la dificultad para controlarlos. Por lo tanto, es muy importante detectar rápidamente la manifestación de estas hierbas malas, para evitar su propagación y reducir su banco de semillas.

Por su parte, las plagas e insectos son otro de los factores que se deben tener en cuenta, también existe una gran diversidad de los mismos, poder lograr asociar las diferentes variables climáticas y ambientales a la aparición de estas plagas permitirá actuar a los productores logrando una menor pérdida de cultivos e incluso una menor utilización de productos químicos que podrían afectar las plantaciones y cultivos.

En general, los productores de dicha región no utilizan recursos renovables como fuente de suministro de energía. En el área de aplicación de las actividades propuestas, inmersa en el sistema interconectado, la energía eléctrica suministrada por la red puede minimizarse, y en el mejor de los casos anularse, mediante el aprovechamiento del recurso solar y/o eólico local.

Sin dudas, el extraordinario avance que viene teniendo en los últimos años la tecnología informática, permite pensar en diferentes estrategias que contrarresten estos inconvenientes. En este sentido, el uso de herramientas basadas en Internet de las Cosas (IoT: *Internet of Things*) y de algoritmos basados en aprendizaje automático (*Machine Learning*), representa una inmejorable oportunidad para:

- promover la vinculación tecnológica entre una Universidad Nacional y el sector socio-productivo agroalimentario de la región;
- generar conocimiento e innovación que permita incorporar la tecnología informática en el sector hortícola de la región;
- impulsar la formación de recursos humanos para el aporte de soluciones tecnológicas;
- aplicar los resultados obtenidos para la formación académica de los estudiantes, vinculando la teoría y la práctica, sobre una problemática concreta dentro del territorio de influencia de la Universidad;
- capacitar y acompañar a los pequeños y medianos productores de la región para el fortalecimiento de la actividad, mediante el uso de las herramientas de las Tecnología de la Información y las Comunicaciones (TIC), en cuanto a la producción y al cuidado del medio ambiente.

2. Líneas de Investigación y Desarrollo

Las líneas de I/D que se presentan en este trabajo están basadas en el estudio y desarrollo de herramientas tecnológicas (hardware-software) basadas en Internet de las Cosas (IoT) y aprendizaje automático, que permitan brindar soluciones innovadoras de avanzada y de bajo costo, y que contribuyan a remediar las principales problemáticas que posee el sector hortícola en el territorio de la UNAJ. Las herramientas desarrolladas serán aplicadas para:

(a) obtener información precisa en tiempo real de magnitudes climatológicas, del estado del suelo, del agua y del nivel de iluminación.

(b) tomar decisiones apropiadas en pos de obtener el rendimiento óptimo de los recursos existentes y mejorar el desarrollo y producción de los cultivos intensivos.

(c) detectar e identificar malezas, las cuales pueden afectar a la producción de los cultivos intensivos de la región.

(d) detectar y/o predecir plagas de forma temprana para minimizar la pérdida de cultivos.

(e) minimizar el consumo de energía proveniente de la red eléctrica, y promover la utilización de energías alternativas.

3. Resultados y Objetivos

Resultados alcanzados:

Con respecto al inciso (a), las tareas desarrolladas hasta el momento han consistido en:

- la planificación, armado y pruebas de funcionamiento del hardware del sistema de monitoreo y control (sensores, memorias, unidades de procesamiento, etc.), teniendo en cuenta las necesidades del sector hortícola. El sistema se dividió en bloques funcionales que se desarrollan independientemente hasta el momento de la implementación del prototipo final. Sobre este punto queda pendiente la adquisición de diversos actuadores, y el armado y pruebas del bloque funcional correspondiente a estos componentes.

- el desarrollo del software de diferentes partes del sistema de monitoreo y control. Específicamente, se llevó a cabo la programación del software embebido para la lectura de los sensores; la programación de la interfaz de comunicación para la transmisión de datos entre los diferentes bloques funcionales (en particular, se utilizó el módulo GSM/GPRS, que permite transmitir información mediante las antenas celulares o Wifi); la programación de una aplicación web

para visualizar datos actuales, alarmas y gráficos, y una base de datos para el almacenamiento de la información histórica, que resulte simple e intuitiva para cualquier usuario. Quedan pendientes aún la programación del software embebido para la activación de los actuadores, la programación del algoritmo de control que procesa los datos y activa los actuadores y la programación de una aplicación móvil, con similares prestaciones que la aplicación web ya desarrollada.

Finalmente, quedan pendientes las pruebas de funcionamiento del sistema completo para las tareas de ajuste y calibración del mismo, y la instalación del sistema desarrollado en huertas de la región, para validar y ajustar su funcionamiento teniendo en cuenta los factores reales. Esta etapa permitirá comprobar además si el entorno desarrollado es amigable al usuario final.

Por su parte, los incisos (b), (c), (d) y (e) son líneas de estudio comenzadas recientemente.

Objetivos esperados:

Se espera continuar con el desarrollo e implementación experimental de un sistema IoT de monitoreo y control escalable, completamente autónomo independiente del usuario, y de bajo costo de realización, para ser instalado en huertas de la región y que permita obtener información útil y precisa en tiempo real de magnitudes climatológicas, del estado del suelo, del agua y del nivel de iluminación. Se espera incorporar al sistema nuevas funcionalidades, tales como: (i) modelos basados en aprendizaje automático para que, en base a la información recolectada por los sensores, el sistema pueda de manera inteligente tomar decisiones apropiadas en pos de obtener el rendimiento óptimo de los recursos existentes y mejorar el desarrollo y producción de los cultivos intensivos; (ii) modelos basados en aprendizaje profundo aplicado al procesamiento de imágenes para la detección, reconocimiento y clasificación de diferentes malezas que pueden afectar a la producción y a los procesos de cosecha de los cultivos de la región.

Se espera a través del desarrollo del sistema de control y gestión de energía basado en recursos renovables, promover el uso racional y eficiente de la energía y la utilización de energías alternativas en la zona de influencia de la UNAJ.

4. Formación de Recursos Humanos

Uno de los principales objetivos del Programa TICAPPS, dentro de la temática de las líneas de I/D presentadas en este trabajo, es la formación de recursos humanos, tanto de docentes investigadores como de estudiantes. En este sentido, se pretende que se familiaricen de manera práctica y teórica, tanto con las herramientas de Internet de las Cosas (sistemas embebidos, redes de comunicación de datos, uso de sensores y actuadores y desarrollo de una aplicación web y una aplicación móvil), como de las nociones fundamentales de los algoritmos basados en aprendizaje automático y profundo.

El equipo de trabajo de las líneas de I/D que se presenta está conformado por tres docentes investigadores formados (dos Doctores y un Magister), tres docentes investigadores en formación, y dos estudiantes avanzados, todos autores de este artículo. Actualmente, se encuentran en curso una Tesis de Doctorado y una beca de Estímulo a las Vocaciones Científicas del Consejo Interuniversitario Nacional (Becas EVC – CIN), relacionadas directamente con las líneas de I/D presentadas. A su vez, todos los miembros docentes autores de este trabajo, participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

5. Bibliografía

- N. Nilsson. "Principles of Artificial Intelligence". Editorial Springer-Verlag. 1980.
- M. Jordan y T. Mitchell. "Machine learning: Trends, perspectives, and prospects", Science. Vol. 349, Issue 6245, pp. 255-260. 2015. DOI: 10.1126/science.aaa8415

- E. Alpaydin. "Machine Learning: The New AI". Editorial MIT Press. 2016.
- O'Reilly Radar Team. "Big Data Now". O'Reilly Media, 2016.
- A. Géron. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems". 2nd Edición. ". Editorial O'Reilly Media, 2019.
- S. Haykin. "Neural Networks. A Comprehensive Foundation". Editorial Prentice Hall. 1999.
- F. Chollet. "Deep Learning with Python". First edition. Manning Publications. 2017.
- I. Goodfellow, Y. Bengio y A. Courville. "Deep Learning". MIT Press, (2016).
- P. Waher. "Learning Internet of Things". Editorial Packt Publishing, 2015.
- Gubbi, J. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 2013.
- Listado de malezas resistentes en Argentina <http://www.senasa.gob.ar/casos-confirmados-de-malezas-resistentes-en-argentina>
- Asociación Argentina de Productores en Siembra Directa
<http://agrovoz.lavoz.com.ar/agricultura/en-cordoba-yuyo-colorado-ya-ocupa-mas-hectareas-que-soja-y-maiz>
- K. Taretto. "Celdas Solares, Teoría, Ensayo y Diseño". Ed. Nueva Librería. Ed. 1º. 2015.

Desarrollo de herramientas para generación de material de entrenamiento y análisis de textos estructurados

Marina Cardenas, Julio Castillo, Nicolás Hernández

Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/Facultad Regional Córdoba/Universidad Tecnológica Nacional
{ ing.marinacardenas, dr.jotacastillo, damiannicolas05}@gmail.com

RESUMEN

En este artículo se describen las herramientas desarrolladas para la investigación sobre técnicas de análisis de textos y minería de datos aplicando métodos y algoritmos de análisis de texto estructurado y no estructurado, estableciendo un conjunto de criterios y seleccionando la mejor alternativa en base a los mismos. En este contexto, se describen las herramientas desarrolladas que permiten abordar la problemática de la escasez de material de entrenamiento para el análisis de textos y herramientas que utilizadas para el análisis y procesamiento de grandes volúmenes de textos, con el objetivo de simplificar el trabajo necesario para analizar, crear, y modificar material de entrenamiento para sistemas basados en aprendizaje por computadora, a la vez que permiten ahorrar tiempo, y disminuir errores manuales cuando se trabaja con textos. Este trabajo se encuentra enmarcado dentro de un proyecto de investigación homologado por la Secretaría de Ciencia y Tecnología (SCyT) de la UTN que tiene por objetivo el crear herramientas que permitan analizar grandes volúmenes de datos que se encuentran en forma de archivos de textos estructurados o no estructurados, adecuando la información para su utilización en el entrenamiento de sistemas de análisis de texto. Dentro los archivos a analizar se pueden mencionar los archivos de código fuente de lenguajes de programación como Python, Java, C++ y C, con la reciente incorporación de archivos de word, pdf y txt.

Palabras Clave: análisis de texto, extracción de información, corpus, machine learning, redes neuronales artificiales, lenguaje formal, lenguaje natural, lenguaje estructurado, lenguaje no estructurado.

CONTEXTO

Este artículo presenta el proyecto denominado “Modelo para el procesamiento de textos estructurados Fase 2” (cód. SIECACO0008518), que es un proyecto homologado por la SCyT de la UTN, que se enmarca dentro del área de lingüística computacional, y que tiene como objetivo el desarrollo de herramientas para análisis textual de diversas fuentes de información en formato estructurado.

Actualmente, el proyecto se encuentra dentro del grupo de investigación denominado Grupo de Aprendizaje Automático, Lenguajes y Autómatas (GA²LA) en donde se desarrollan proyectos relacionados con autómatas y lenguajes formales, procesamiento del lenguaje natural, y aprendizaje automático, y en especial proyectos orientados a la aplicación de la inteligencia artificial para resolver problemas de las ciencias sociales. Físicamente, los integrantes del proyecto desarrollan sus actividades en el Laboratorio de Investigación de Software LIS del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC).

1. INTRODUCCIÓN

En este artículo se describen herramientas de software que han sido desarrolladas para el estudio e investigación sobre análisis de texto estructurado y el área de minería de datos. Las mismas, han sido construidas para abordar el problema del material de entrenamiento en los algoritmos basados en aprendizaje por computadora, disminuir el trabajo manual de clasificación, anotado y selección que suele ser necesario cuando se trabaja con orígenes de datos basados en texto, y brindar soporte en las prácticas de detección de código fuentes similares con distintos enfoques, desde detección de clones de código fuente que afectan al ciclo de vida y mantenimiento del software hasta detección de prácticas de plagio de software.

Los materiales de entrenamiento producidos por estas herramientas pueden ser utilizados para abordar problemas de extracción de información y minería de datos en textos estructurados [1, 2], tales como búsquedas de implicaciones en textos [3], implicación textual [4, 5], extracción de información [6, 7] o minería de datos [8]. En particular, fueron diseñadas para ser empleadas cuando se utilicen técnicas de aprendizaje supervisado de análisis de textos. La generación y análisis del material del entrenamiento para un sistema de análisis de texto estructurado es una tarea muy ardua y artesanal para ser realizada por un humano, lo cual motiva la construcción de programas que permita ayudar a los anotadores humanos en la construcción (semiautomática) de corpus [9, 10].

2. LÍNEAS DE INVESTIGACIÓN

Las líneas de investigación de este proyecto se basan en el uso de redes neuronales más precisamente en aprendizaje supervisado, semi supervisado y machine learning y su aplicación en el análisis de la Lingüística Computacional. Este campo tiene por objetivo lograr la capacidad de reconocer y comprender el lenguaje humano mediante la

creación de modelos computacionales. Una de las problemáticas encontradas para esta área de investigación, es la escasez de material para ser utilizado en el aprendizaje supervisado tanto en el entrenamiento como en las pruebas, de esto se deriva la creación y utilización de corpus, que es el estudio empírico de la lengua a partir de los datos que proporcionan ejemplos reales de producciones lingüísticas.

El equipo de investigación también trabaja en el desarrollo de modelos computacionales en general. Uno de los modelos desarrollados permite modelar la ocurrencia de incendios forestales, y forma parte de otro proyecto de investigación que se denomina Modelos para la Predicción de Incendios Forestales, un proyecto homologado por la SCyT de la UTN.

La línea de innovación está relacionada a la integración de los resultados obtenidos de la lingüística de corpus, de la construcción de modelos y del uso de técnicas de aprendizaje automático, en el marco del grupo GA²LA.

3. RESULTADOS OBTENIDOS/ESPERADOS

El desarrollo del presente proyecto nos ha dado como resultado un conjunto de herramientas que han sido desarrolladas para el estudio e investigación sobre análisis de texto estructurado y el área de minería de datos.

La primera herramienta se denomina Programa de Mapeo de Datos (PMD), que tiene como objetivo permitir el análisis y la manipulación de la información de diferentes fuentes y orígenes de datos y almacenarlos en una estructura estándar en una base de datos relacional SQL Server, permite tomar datos de orígenes de datos estructurados y registrarlos en una nueva base de datos que se encuentra normalizada para facilitar la búsqueda y análisis de textos.

El PMD está desarrollado tanto en versión de aplicación de escritorio (en el lenguaje de programación C#), como en versión web.

Ambas versiones permiten almacenar los datos normalizados en una estructura estándar de una base de datos SQL Server facilitando la búsqueda y análisis de textos.

En la Figura 1 se puede visualizar la pantalla donde el usuario establece la conexión con las bases de datos involucradas en el proceso de mapeo / transformación de datos desde un origen hacia un destino.



Figura 1. Interfaz principal de mapeo de datos

La segunda herramienta es un Banco de Prueba de Algoritmos de Semejanza (BPAS) que consiste en una aplicación de consola que utiliza el método de jerarquía analítica (AHP) para realizar comparaciones entre diversos algoritmos que realizan evaluación de la similitud de textos estructurados, aunque también pueden aplicarse sobre textos no estructurados. El objetivo de esta herramienta es el de evaluar la capacidad de la detección de similitudes de código de programas en los diferentes niveles establecidos para el método AHP. y se ha analizado código fuente escritos en diversos lenguajes de programación, tales como, Java, C, C#, Perl y Python.

La última herramienta mencionada, permite realizar el procesamiento de archivo de código fuente con diferentes algoritmos que permitan la detección de prácticas de reutilización para determinar aquellos con mayor similitud.

Este sistema tiene por objetivo determinar las similitudes de archivos de código fuente de lenguajes como Java, C++, C y Python, entre

otros, trabajando a diferentes niveles de granularidad como módulos, funciones y líneas de código [11]. Esta herramienta ha sido desarrollada con orientación web, para que su disponibilidad sea mayor, por lo que su interfaz se creó mediante la utilización de esas tecnologías tal como se muestra en la figura 2. El SDS permite cargar los archivos necesarios para trabajar en la detección de similitudes, en base a los algoritmos implementados por herramientas disponibles en la Web. Asimismo permite la configuración de los parámetros requeridos por cada algoritmo para poder realizar el análisis comparativo, como así también para almacenar el resultado en un nuevo archivo y generar reportes con las salidas de las ejecuciones realizadas, que permitirán al decisor identificar rápidamente un subconjunto de archivos que presentan gran similitud entre sí, dentro de un conjunto grande de archivos.

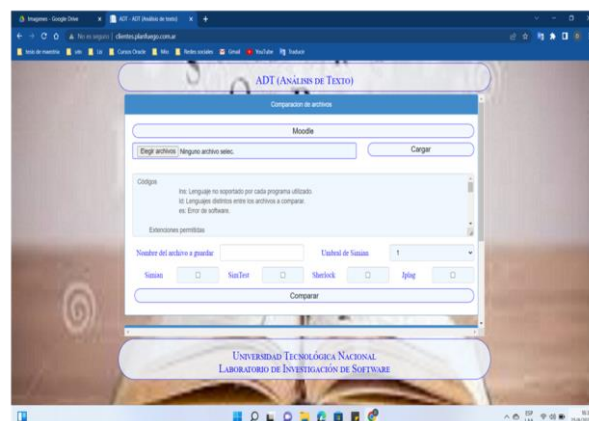


Figura 2. Interfaz principal del SDS.

4. FORMACIÓN DE RECURSOS HUMANOS

- Una doctoranda en ingeniería con mención en sistemas de información en la UTN-FRC, que está trabajando específicamente en el subsistema de detección de similitudes en archivos de código fuente (SDS). Además, realiza la dirección de becarios de posgrado y

- de becarios de grado en el contexto del proyecto.
- Un doctor en ciencias de la computación quien desarrolló su tesis en el área de investigación y realiza la dirección de becarios de posgrado y de becarios de grado en el contexto del proyecto.
 - Un maestrando en Ingeniería en Sistemas de Información de la UTN-FRC, que está desarrollando su tesis en el área de análisis de textos estructurados y estimaciones de software.
 - Ocasionalmente participan en el proyecto, alumnos que realizan su práctica supervisada. La cantidad de alumnos depende de las necesidades del proyecto y de los alumnos dispuestos a trabajar en el proyecto. Las prácticas supervisadas son un requisito necesario para la obtención del grado de Ingeniero. Las mismas pueden realizarse en la industria o en el ámbito académico en el contexto de proyectos de investigación homologados.
 - En el proyecto participan alumnos becarios anualmente, complementando así su formación curricular desde el punto de vista científico.
 - Además participan docentes investigadores ingenieros del departamento de Ingeniería en Sistemas de Información de la UTN FRC.
 - Adicionalmente, se realizan charlas de difusión y jornadas de capacitación a alumnos y a docentes de ingeniería en sistemas de información en las líneas temáticas enumeradas anteriormente.

5. BIBLIOGRAFÍA

- [1] Klavansy, J., Resnik P., “The Balancing Act. Combining Symbolic and Statistical Approaches to Language”, MIT Press, 1996.
- [2] Manning, C., Schutze H., “Foundations of Statistical Natural Language Processing”. The MIT Press, Cambridge, MA, 1999.
- [3] Castillo, J., “Sagan in TAC2009: Using Support Vector Machines in Recognizing Textual Entailment and TE Search Pilot task”. TAC, NIST, USA, 2009.
- [4] Castillo, J. y Cardenas. M., “Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment”. Iberamia 2010, LNCS, vol. 6433, 2010, pp. 366-375.
- [5] Castillo, J., “Using Machine Translation Systems to Expand a Corpus in Textual Entailment”, Proceedings of the Iccetal 2010, LNCS, vol. 6233, pp. 97-102.
- [6] Feldman, R. y Hirsh, H., “Exploiting Background Information in Knowledge Discovery from Text”, Journal of Intelligent Information Systems, 1996.
- [7] Lewis, D., “Evaluating and optimizing autonomous text classification systems”, Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval. Seattle, US, 1995, pp. 246-254.
- [8] Craven, M., Shavlik, J., “Using Neural Networks for Data Mining”, Future Generation Computer Systems, 13, 1997, pp. 211-229.
- [9] Castillo, J. y Cardenas, M., “Creación de Corpus para Aplicaciones de Análisis de Texto No Estructurado”. CONAIISI 2015. Buenos Aires.
- [10] Stefan, T. y Stefanowitsch, A., “Corporain in Cognitive Linguistics”, Corpus-Based Approaches to Syntax and Lexis, Berlin: Mouton, 2006, pp. 1-17.
- [11] Castillo, J. y Cardenas, M., Curti, A., Casco, O., Navarro, M., Hernández, N., y Velazco, M., “Desarrollo de sistemas de análisis de texto,” in XIX Workshop de

Investigadores en Ciencias de la
Computación, 2017, pp. 58–62.

SISTEMA DE PREDICCIÓN DE INCENDIOS FORESTALES

Julio Castillo, Marina Cardenas, Nicolás Hernández, Martín Gramática, Diego Serrano
Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/Facultad
Regional Córdoba/Universidad Tecnológica Nacional
{dr.jotacastillo, ing.marinacardenas, damiannicolas05, martingra, diegojserrano}@gmail.com

RESUMEN

En este artículo se describen las líneas de investigación relacionadas con el proyecto “Modelos de pronósticos para la detección de incendios forestales - Fase 2”, las cuales abarcan las siguientes áreas de conocimiento como, inteligencia artificial y aprendizaje supervisado, desarrollo de sistemas web, desarrollo de aplicaciones móviles, integración con sistemas de información geográfica (GIS). El presente proyecto inició con el desarrollo de un sistema de aprendizaje supervisado [1], el cual tuvo como resultado la creación de un modelo que permite predecir la ocurrencia de incendios forestales, mediante la utilización de ciertas variables climáticas. Estos datos son obtenidos por un sistema web que permite la recopilación y tratamiento de datos meteorológicos, para ser utilizados por el modelo de predicción. Tanto el modelo de predicción como el sistema de administración de datos meteorológicos, tiene por objetivo brindar apoyo al Plan Provincial de Manejo del Fuego de la Secretaría de Gestión de Riesgo y Catástrofes del Gobierno de la Provincia de Córdoba.

Palabras clave: Modelos de predicción, aprendizaje supervisado, IA, desarrollo web, incendios forestales, FWI, recolección automática, web scraping, WCF, GIS, Bing Maps, datos meteorológicos.

CONTEXTO

El presente proyecto se encuentra homologado por la Secretaría de Ciencia y Tecnología dependiente de la Universidad Tecnológica Nacional (UTN), con código de proyecto: SITBCO0008196TC, el cual es llevado a cabo en el Laboratorio de

Investigación de Software (LIS)¹ del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN - FRC).

Para dar soporte en la investigación de estas áreas de conocimiento, se formalizó la creación de un grupo de investigación nombrado Grupo de Aprendizaje Automático, Lenguajes y Autómatas (GA²LA) [2], el mismo tiene por objetivo el aprendizaje y aplicación práctica de los conocimientos adquiridos en problemas reales, y articular diversos proyectos dentro de las líneas de investigación de teoría de autómatas y lenguajes, IA, desarrollo de sistemas de información.

1. INTRODUCCIÓN

Uno de los fenómenos climáticos que causan mayor impacto en el medio ambiente y genera mayores pérdidas y contaminación, son los incendios forestales [3][4]. Actualmente han aumentado ya que se producen las condiciones propicias debido a la contaminación ambiental [5][6].

El desarrollo de este proyecto se inició hace más de 10 años atrás y se encuentra enmarcado dentro del grupo GA²LA en el Laboratorio de Investigación de Software de la UTN-FRC para adquirir conocimientos en sistemas de aprendizaje supervisado y brindar una herramienta que ayude a combatir focos de incendios forestales a través de la detección temprana de los mismos. Conforme el proyecto fue avanzando, surgió la necesidad de contar con información meteorológica actualizada para poder realizar las predicciones de focos ígneos, para lo cual se desarrollaron módulos que permiten

¹ www.investigacion.frc.utn.edu.ar/mslabs/

recolectar, de manera automática, la información meteorológica de diferentes sitios web que es procesada y transformada para alimentar el modelo de predicción y en la confección de los índices necesarios para poder crear el FWI (Fire Weather Canadian Forest o Índice Canadiense de Incendios Forestales) [7]. Toda esta información puede ser visualizada de diferentes formas, tales como informes, estadísticas y mapas digitales.

A continuación se describen los subsistemas que se desarrollaron para dar soporte en el desarrollo del proyecto:

1.1. Subsistema de Entrenamiento y Predicción

Este sistema tiene por objetivo predecir la ocurrencia de focos de incendios forestales para así poder alertar de manera temprana a la población para evitar o minimizar las pérdidas de vidas, económicas y ecológicas.

Este subsistema se muestra en la figura 1 y se divide en dos partes, la primera es la etapa de entrenamiento, donde se utilizan registros para crear un modelo matemático /computacional. La segunda parte toma el modelo producido y, mediante el ingreso de nuevos registros, se pueden inferir nuevas ocurrencias de siniestros.

Fig 1.Subsistema de predicción

1.2. Subsistema de Recopilación de Información

Este subsistema, que se muestra en la figura 2, tiene por objetivo la administración de los datos meteorológicos [8] para alimentar el subsistema de entrenamiento y predicción antes mencionado, también permite visualizar datos de predicciones, siniestros, mediciones e índices. Este sistema está orientado a web, por lo que se desarrolló con las tecnologías ASP.Net y C# que se incluyen en la plataforma de Microsoft.

Fig 2. Subsistema de Recopilación de Información

1.3. Subsistema Chat

Este subsistema permite establecer la comunicación entre el testigo presencial (es quien, ante una ocurrencia de un principio de incendio o un incendio forestal, da aviso) y el cuartel de bomberos para lograr que se combata el fuego de manera efectiva y eficiente. Este chat está en desarrollo mediante la tecnología provista por Microsoft, SignalR que se encarga de abstraer al programador de los protocolos de comunicación, permitiendo una comunicación en tiempo real.

1.4. Subsistema Twitter

Este subsistema es una alternativa de comunicación al subsistema Chat, que utiliza la popular red social twitter para que los testigos presenciales, puedan dar aviso. Este subsistema cuenta con la particularidad de que se implementó la detección de falsos positivos, mediante un sistema de palabras claves para filtrar aquellos mensajes que no están dando una alerta sobre un incendio.

1.5. Subsistema de Recolección Automática de Información Meteorológica

Este sistema recolecta la información automáticamente de diversos orígenes de datos web [8], tales como:

- Servicio Meteorológico Nacional².
- WunderGround³.
- Freemeteo⁴.
- Ventusky⁵.
- Base de Datos del Observatorio Hidrometeorológico de la Provincia de Córdoba⁶.

Dichas fuentes presentan los datos de distintas maneras, por lo que se utilizaron técnicas como el procesamiento de archivos XML y JSON, que son conocidos formatos para el intercambio de datos, y en otros casos se aplicó la técnica de web scraping a través de la implementación de un servicio web con la tecnología WCF de Microsoft. Dado que los datos meteorológicos deben ser medidos en ciertos horarios, se posibilitó la programación de alarmas en los horarios relevantes y se permite la modificación, si cambian los requerimientos de las mediciones.

1.6. Subsistema de Mapas Digitales Interactivos

Este subsistema permite visualizar la información en mapas digitales, incorporando información espacial a los datos que ya disponemos. Para poder realizar esto, se utilizó Bings Maps como proveedor de mapas digitales y, a través de la API, mediante las tecnologías de JavaScript, GeoJSON, ASP.Net se definieron cuatro opciones de mapas electrónicos: mapas de predicciones, de índices de riesgo, de siniestros y por último de mediciones meteorológicas.

Este subsistema provee información visual acerca de los datos recolectados por el Subsistema de Recolección Automática de Información Meteorológica y su visualización

se basa en un cálculo de índices de riesgo propios de cada estación meteorológica y son calculados en base a la información meteorológica.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Como se menciona anteriormente, este proyecto inicia con la línea de investigación de aprendizaje supervisado, que tiene por objetivo que una computadora, en base a datos de entrenamiento y algoritmos, aprenda a inferir la ocurrencia de un evento, en este caso focos de incendio forestales. Dicho proyecto es la continuación de un proyecto previo, también homologado en UTN denominado “Modelos de pronósticos para la detección de incendios forestales”. Se ha continuado trabajando sobre los resultados obtenidos del proyecto anterior, lo cual se puede ver reflejados en los diferentes subsistemas descritos en la sección anterior.

El proyecto involucra las áreas conocimientos de inteligencia artificial, en particular aquellos relacionados con aprendizaje automático por computadoras de tipo supervisado, a la vez que involucra la construcción de modelos de pronósticos. Se trata de un problema complejo y no lineal, el cual se ve afectado en gran medida por la mano del hombre, puesto que la mayor parte de los incendios son intencionales.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como resultados obtenidos se menciona que se logró el desarrollo e integración de varios subsistemas, que ahora conforman un sistema de administración de datos, que se encarga de la recolección de forma automática de datos meteorológicos, creación de índices de riesgo de incendios,

² <https://www.smn.gob.ar/>

³ <https://www.wunderground.com/>

⁴ <https://freemeteo.com.ar/>

⁵ <https://www.ventusky.com/>

⁶ <https://bdhm.ohmc.ar/>

gestión de predicciones y siniestros. Este sistema brinda soporte al modelo de predicción al proporcionar datos de entrenamiento y de prueba, además brinda diferentes opciones para la visualización de la información, mediante estadísticas, informes y mapas digitales.

Por último se espera poder mejorar los modelos de predicción mediante la integración con sistemas GIS y la incorporación de datos satelitales mediante la incorporación de tecnologías como Google Engine.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación y desarrollo de software está formado principalmente por docentes investigadores de la Universidad Tecnológica Nacional, Facultad Regional Córdoba. Además, anualmente se incorporan al proyecto, becarios alumnos, y eventualmente un becario de posgrado.

El presente proyecto ha servido como marco para la realización de un tesis de maestría por parte de la ingeniera Marina Cardenas, la realización de una tesis de especialidad por parte del ingeniero Nicolás Hernández, y también permitió que estudiantes realizaran sus prácticas supervisadas con el objetivo de completar sus estudios.

6. BIBLIOGRAFÍA

[1] Cardenas M., Medel R., Castillo J., Vázquez J. y Casco O. (2015) Modelos de aprendizaje supervisados: aplicaciones para la predicción de incendios forestales en la provincia de Córdoba. XVII Workshop de Investigadores en Ciencias de la Computación. Salta, Argentina.

[2] Vázquez J., Castillo J., Constable L., y Cardenas M. (2018) GA²LA: Grupo de

Aprendizaje Automático, Lenguajes y Autómatas. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).

[3] Calderón M., Gavarrete G. y Guzmán Navarrete R.(2013). La prevención y manejo de desastres ambientales que afectan el medio ambiente, el patrimonio y la vida de los salvadoreños. Tesis de Licenciatura. Universidad de El Salvador.

[4] Cruz, A. (2005) Causa calentamiento global desastres. Diario: El Universal. México; 2005, septiembre 14.

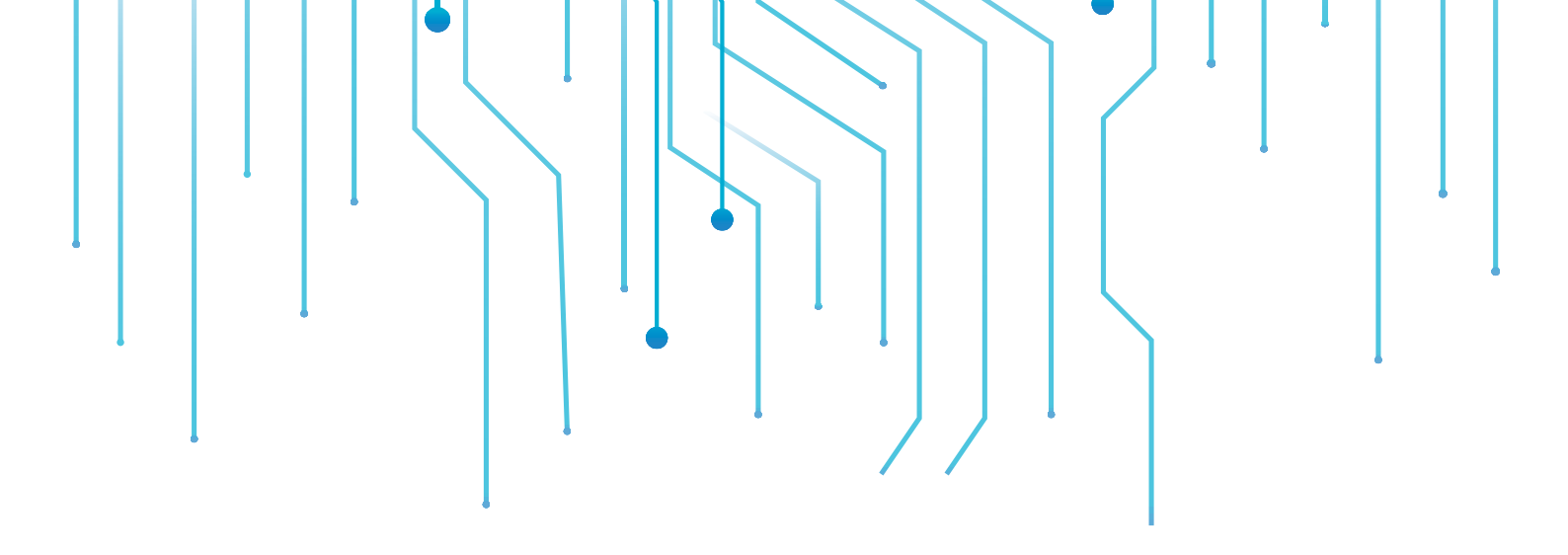
[5] Sanchez S., Zanvettor R., Grilli M. y Ravelo A. (2021) Impacto de la sequía en los incendios forestales en las Sierras de Córdoba, Argentina; Asociación Argentina de Agrometeorología; Revista Argentina de Agrometeorología; 12; 9-2021; 37-45.

[6] Luna C., Fontana, M., Kurtz, D. y Saucedo G. (2022) Estado de situación de daños por incendios forestales registrados en Argentina. Innova Biology Sciences, 2(1), 23–40.

[7] Lawson B. y Armitage O. (2008) Weather Guide for the Canadian forest fire danger rating system. Natural Resources Canada, Canadian Forest Service, Northern Forestry Center. Edmonton, Canadá.

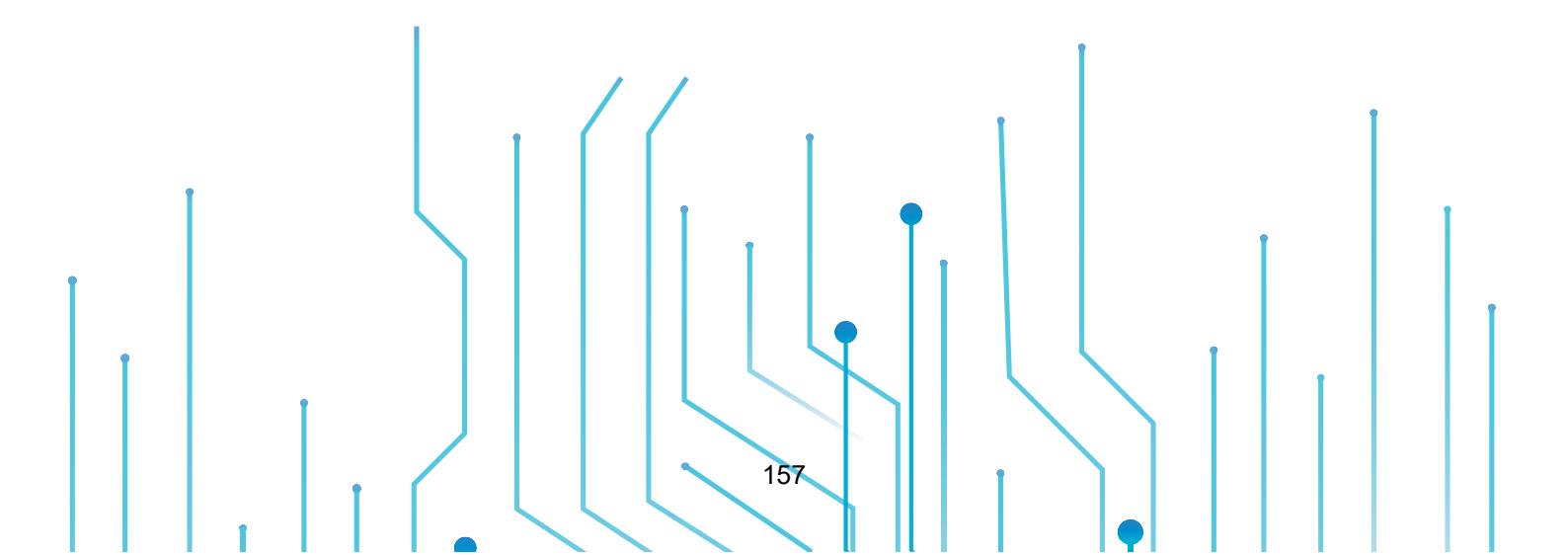
[8] Cardenas M., Castillo J., Medel R., Casco O., Navarro M., Gutierrez S. y Curti A. (2016) Sistema de Predicción de Incendios Forestales para la Provincia de Córdoba. CONAIIISI 2016.

[9] Cardenas M., Hernandez N., Navarro Mugas M. y Castillo J.. (2019) Herramientas de Recolección, Visualización y Análisis de Incendios Forestales, 7mo CONAIIISI. 14-15 de Nov. de 2019 en la Universidad Nacional de La Matanza – San Justo, Buenos Aires, Argentina.



BDDM

Bases de Datos y Minería de Datos



Propuesta de técnicas de validación para la calidad de datos abiertos e identificación de patrones para predicciones con Machine Learning:

Segunda Parte

Roxana Martínez, Pablo Vilaboa, Agustín Simón, Juan Iannini, Eugenio Clavijo, German Colato, Nelson Catala, Carlos Bateca, Serenela Lopumo, Hernán Zabala, Gonzalo Parada, Ayelén Chaglasian Sgang, Rocío Vargas, Sebastián Tsiro

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{roxana.martinez, pablo.vilaboa, rocio.vargas}@uai.edu.ar

{MiguelAgustin.Simon, Juanmanuel.GarciaIannini1, Eugenio.Clavijo, GermanDario.Colato, Nelson.Catala, CarlosAlberto.BatecaCaicedo, Serenela.Lopumo, HernanGuillermo.Zabala, Gonzalo.Parada, AyelenMalena.Chaglasian, SebastianMartinDemetrio.Tsiros}@alumnos.uai.edu.ar

RESUMEN

Dentro de los varios focos del paradigma de Gobierno Abierto, existen los conjuntos de datos abiertos públicos disponibilizados en los sitios web gubernamentales, éstos permiten no sólo una mejor transparencia en las actividades públicas que se realizan, sino que, además, pueden ser reutilizados a través de estudios para brindar mejores servicios y obtener beneficios para la sociedad. Mantener estos datos con calidad, permite que se lleven a cabo, análisis de escenarios para predecir y evitar posibles falencias o errores en la toma de decisiones políticas que tenga un alto impacto en los ciudadanos. Actualmente, no existen demasiados trabajos de investigación que se centren en el contenido de los datasets, en aspectos de: análisis de contenido interno, validación, herramientas de verificación, cotejo en las estructuras de datos, guías de buenas prácticas, o bien reconocimiento de patrones estructurales, entre otros temas. Es por esto, por lo que, este trabajo de investigación se centra en dichos puntos mencionados para validar la calidad de los datos públicos abiertos y que éstos puedan ser utilizados con un fin positivo en la comunidad. Este trabajo, no solo trabaja en prototipos de software en validaciones de

datos específicos, sino que, además, permite estudiar modelos predictivos, que a través del aprendizaje automático pueden brindar un mejor análisis para la toma de decisiones.

Palabras clave: Datos Abiertos, Gobierno Abierto, Métricas de Calidad de Datos, Machine Learning, Predicciones a partir de Patrones.

CONTEXTO

El presente trabajo es la segunda parte del proyecto denominado “Investigación y desarrollo de software para la validación de la calidad de datos abiertos e identificación de patrones para predicciones”, que tiene inicio en el mes de marzo 2022 y con fecha de culminación en el mes de marzo 2024. Este proyecto pertenece a la línea de investigación de Ingeniería de Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI), el cual contribuye al desarrollo de las Tecnologías de la Información y las Comunicaciones (TIC) en Argentina y en el mundo, llevando adelante la investigación en diversas áreas a través de tratamiento y manipulación de datos en diversas temáticas, tales como: aspectos gubernamentales,

gobierno de datos, economía, salud, videojuegos, multimedia entre otras. El proyecto es financiado y evaluado por la Secretaría de Investigación de la Universidad, tiene una duración de 2 años, y cuenta con la participación de docentes y estudiantes de grado y posgrado en diversas carreras de la Facultad de Tecnología Informática.

1. INTRODUCCIÓN

Gobierno Abierto se encuentra en pleno auge tanto a nivel nacional como internacional, haciendo partícipe a distintos gobiernos y ciudadanos en este contexto colaborativo y participativo de datos y servicios, en conjunto con el objetivo de brindar una mayor transparencia en las acciones gubernamentales realizadas por parte del Estado Nacional. Si bien Gobierno Abierto implica varios enfoques, este trabajo se centra en la implementación de nuevas técnicas de gestión de datos y en el estudio de la disponibilización de datos públicos a través de un diseño tecnológico, que se encuentra centrado en el ciudadano promedio para fomentar la inclusión social. “Al ampliar el acceso a la información pública se fortalece la rendición de cuentas y se enriquece el debate público, a la vez que se crean nuevas oportunidades para generar valor agregado” [1]. Por otra parte, es necesario mencionar que “los datos que se pueden reutilizar y redistribuir sin ninguna restricción se denominan datos abiertos” [2]. Estos se encuentran en los portales gubernamentales o bien en sitios web que poseen licencias específicas que permiten trabajar con formatos digitales concretos, sobre un posible modelo estándar abierto.

Es preciso conocer que “la cultura de datos abiertos trata de obtener un valor añadido de la información [...], la información no genera sólo su valor por estar reservada a unos pocos, sino que lo hace por su disponibilidad para ser interpretada y traducida por cualquier actor interesado en trabajar con ella” [3], es decir, un dato abierto puede ser accedido a través de un formato especial que permite la interoperabilidad con otros softwares.

Como parte de la problemática de este tema, cabe mencionar que “no toda la información pública disponible o publicada en la web es información abierta válida para su reutilización. No sólo se trata de publicar los datos, sino que hay que garantizar el acceso a ellos, razón por la que debe recurrirse a formatos digitales, estandarizados y abiertos, siguiendo una estructura clara que permita su comprensión y reutilización” [4].

Existen algunos trabajos enfocados a las mediciones en aspectos de calidad de datos públicos [5] [6] [7], por otro lado, otras investigaciones apuntan a medidas otorgadas a la calidad de los datos abiertos, este es el caso del esquema modelo de cinco estrellas de Berners-Lee [8], o bien trabajos realizados con informes sobre el estado de los datos elaborados por fundaciones comprometidas con el impacto de estos datos públicos [9]. Es importante que se trabaje de manera colaborativa sobre algunos puntos FODA (Fortalezas, Oportunidades, Debilidades y Amenazas) con respecto a varias situaciones a nivel Estado Nacional, con el fin que tanto las entidades gubernamentales como los ciudadanos, puedan brindar soluciones específicas a determinados problemas trabajando en conjunto [10].

Gracias al gran impulso de la tecnología en el área de datos, el concepto de Machine Learning (ML) cada vez se encuentra más en auge, también conocido en español como aprendizaje automático, es una línea que deriva de la inteligencia artificial que hace foco en estudiar y crear algoritmos que permiten entrenar los procesos de decisión de las computadoras. En el contexto del aprendizaje automático, los datos son un insumo importante, ya que son el punto de inicio desde el cual se puede aprender y entrenar cada uno de los modelos de Machine Learning [11]. En lo que respecta al proceso de entrenamiento, implica linealmente el uso de datos, los cuales pueden ser datos estructurados o no estructurados [...]. La elección de usar un algoritmo depende del tipo de problema que se esté abordando. Esta decisión, es parte del proceso de selección del ingeniero de datos, que tiene como tarea,

identificar y trabajar con el algoritmo que mejor se adapte a cada problema [12].

El aprendizaje automatizado, tiene una amplia variedad de usos en diferentes ámbitos. Algunos de los beneficios de su utilización, se muestran a continuación:

- En el ámbito Financiero, ML se usa para la detección de fraudes, la evasión, el análisis de perfil para otorgar créditos, entre otros aspectos. Los modelos, analizan grandes volúmenes de datos de fuentes diferentes. Por ejemplo, los algoritmos cruzan los datos financieros buscando patrones o anomalías que pudiesen indicar una o varias actividades de interés para la investigación [13].
- Otro ámbito es el contexto de la salud: el aprendizaje automático es usado en varios puntos, por ejemplo: puede utilizarse para predecir enfermedades y analizar imágenes médicas. Del mismo modo, estos modelos analizan grandes volúmenes de datos médicos para interpretar el comportamiento de algunas variables en la búsqueda de nuevo conocimiento con mayor rapidez y precisión [14].
- El sector del marketing utiliza machine learning para personalizar anuncios y para establecer una segmentación adecuada del público objetivo. Los modelos de aprendizaje automatizado pueden analizar los datos de cada cliente identificando comportamientos genéricos y preferencias según su segmentación. Esto permite que los profesionales puedan crear anuncios dirigidos a un cliente potencial y con ello construir un canal de comunicación más efectivo [15] [16].
- El sector de la producción utiliza aprendizaje automatizado para diferentes tareas dentro de la empresa como el mantenimiento predictivo, la optimización de los procesos de producción, entre otros. Los modelos de ML analizan datos de los sensores de la planta de producción, los datos se recopilan durante el proceso y se almacenan. Con estos datos, se pueden predecir fallos, o se puede mejorar el proceso para hacer un mejor uso de los recursos. El efecto de utilizar proceso de

aprendizaje automatizado permite reducir el tiempo de inactividad y los costos que estas fallas producen al realizar tareas de mantenimiento predictivo [17].

En conclusión, la utilización del aprendizaje automatizado permite trabajar con los datos para entrenar a diferentes modelos, que luego, éstos a través de algoritmos, colaboran en la toma de decisiones en el estudio de escenarios de análisis. Su aplicación ha conquistado muchos ámbitos, ofreciendo beneficios y usos variados. Esto último, hace de este tema un campo de estudio muy interesante. Por otra parte, es fundamental comprender que para que esto se consiga, es necesario contar con datos de calidad en los portales de datos abiertos que serán utilizados en contextos de ML. Sin esta calidad, los resultados estudiados podrían verse sesgado o con falencias. Hoy por hoy, sigue siendo un gran desafío para las políticas públicas, ya que muchos de los conjuntos de datos públicos no se encuentran en formatos estandarizados o bien se presentan incompletos.

2. LÍNEAS DE INVESTIGACIÓN DESARROLLO

Este proyecto pertenece a la línea de investigación de Ingeniería de Software del Centro de Altos Estudios en Tecnología Informática (CAETI). Para esta segunda parte del proyecto, se hace enfoque en los siguientes ejes del tema:

- Continuar con el diseño y desarrollo de algoritmos para identificar patrones en datos públicos abiertos.
- Estudiar modelos predictivos orientados a contextos gubernamentales, como así también, analizar posibles predicciones sobre nuevos datos encontrados con técnicas de machine learning.
- Analizar herramientas de manejo con Machine Learning que permitan análisis de patrones con enfoque predictivo.
- Establecer criterios para mantener estándares de calidad de datos;
- Desarrollar una herramienta de validación de datasets gubernamentales basada en

métricas de calidad de datos estándares como geoespaciales;

- Generar una guía de buenas prácticas para las técnicas de Machine Learning en la utilización de datos abiertos.
- Definir y recolectar la muestra de datasets para ser testeados por la herramienta propuesta como validadora de calidad;
- Análisis de casos predictivos con casos reales.
- Análisis de casos de aplicación con los prototipos desarrollados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Enfoque del proyecto:

Los conjuntos de datos públicos del contexto ofrecido por el nuevo paradigma de Gobierno Abierto, si bien son una fuente útil que se disponibiliza a los distintos organismos, como ciudadanos, tienen falencias en aspectos de calidad de los datos. Es por lo que resulta necesario, que éstas puedan ser detectadas y analizadas a tiempo, para evitar errores en el estudio de éstos, y mitigar posibles interpretaciones erróneas. Para ello, la utilización de herramientas de software para realizar la verificación y medición de este aspecto son fundamentales. Mantener la calidad de datos, permitirá trabajar con algoritmos en los que se pueda detectar patrones que fomenten el estudio de predicciones en escenarios posibles, y que esto traerá aparejado, beneficios sociales a las personas.

Objetivos principales:

Realizar el análisis, diseño y desarrollo de herramientas de software para la gestión y validación de la calidad de los datos públicos en el contexto de Gobierno Abierto. Detectando el “estado de salud” de las diversas fuentes de datos provenientes de casos de aplicación gubernamentales con los prototipos desarrollados, incorporando algoritmos para identificar patrones que logren predicciones sobre nuevos datos.

Metodología y Técnicas:

El proceso metodológico de investigación utilizado para esta segunda parte del proyecto, continúa con un proceso sistemático cualitativo, que implica las siguientes etapas: estudio del problema; revisión teórica; recolección y revisión de datos, clasificación y análisis de datos; estudio de escenarios de identificación de patrones sobre éstos con enfoques predictivos sobre nuevos datos; desarrollo de prototipos (análisis cuantitativo de datos); visualización de la predicción de patrones; validación de la solución propuesta; identificación de las limitaciones del trabajo.

Resultados obtenidos:

En lo que fue la primera parte del proyecto (primer año), se analizaron las falencias en aspectos de calidad de datos públicos abiertos, se lograron identificar técnicas de validación de calidad de datos abiertos en formatos estándares y se llevó a cabo el estudio de datos geolocalizados que se encuentran disponibilizados en los portales gubernamentales, como así también, se trabajó en el análisis de propuestas de diseño y desarrollo de algoritmos con la finalidad de descubrir patrones en datos públicos. Además, se estudiaron posibles modelos predictivos orientados a este contexto (herramientas de software y demás), para efectuar predicciones y mejorar los resultados de análisis en resultados finales con técnicas de Machine Learning. Finalmente, se lanzó un primer prototipo de software, que tiene como objetivo la validación de calidad en datos públicos abiertos geolocalizados.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto se compone por una docente investigadora, Doctora en Ciencias Informáticas de la Universidad Nacional de La Plata (UNLP), directora del Proyecto de Investigación, 2 (dos) docentes tesistas y 2 (dos) estudiantes tesistas de la carrera de Maestría en Tecnología Informática en la Universidad Abierta Interamericana (UAI). También cuenta con un docente tesista de la

carrera del Doctorado en Ciencias Informáticas de UAI, un estudiante tesista de la carrera de la Especialización en Ingeniería de Softwares de UAI, 5 (cinco) estudiantes colaboradores de la carrera de Ingeniería en Sistemas Informáticos de UAI, y 2 (dos) tesistas de la carrera Licenciatura en Gestión de Tecnología Informática en UAI. En relación directa con la línea de I/D presentada para el proyecto, los miembros del equipo se encuentran en realización de: 1 tesis doctoral, 4 tesis de maestría, 1 tesis de especialización y 2 tesinas de grado de la UAI.

5. BIBLIOGRAFÍA

- [1] Buenos Aires provincia (2017). “Kit de Apertura Municipal”. Disponible en: <http://escueladefiscales.com/Kit%20de%20Apertura%20Municipal%202017%20-%20provincia%20de%20buenos%20aires.pdf>
- [2] Oviedo, E., Mazón, J. N., & Zubcoff, J. J. (2013). Hacia un modelo de calidad de datos para portales de datos abiertos. In XXXIX Latin American Computing Conference (CLEI), Nanguata (pp. 1-8).
- [3] Magallón Rosa, R. (2017). Datos abiertos y acceso a la información pública en la reconstrucción de la historia digital.
- [4] Garriga-Portolà, M. (2011). ¿Datos abiertos? Sí, pero de forma sostenible. *Profesional de la Información*, 20(3), 298-303.
- [5] ISO 25012 (2008). “Ingeniería de software - Requisitos de calidad y evaluación de productos de software (SQuaRE) - Modelo de calidad de datos”. Disponible en: <https://www.iso.org/obp/ui/es/#iso:std:iso-iec:25012:ed-1:v1:en>
- [6] Martínez, R. et al. (2021). Metrics proposal to measure the quality of governmental datasets. *IEEE Latin America Transactions*, Vol. 100. ISSN 1548-0992.
- [7] de España, G. (2017). Manual práctico para mejorar la calidad de los datos abiertos. Madrid. Disponible en: https://datos.gob.es/sites/default/files/doc/file/manual_practico_para_mejorar_la_calidad_de_los_datos_abiertos_1.pdf
- [8] Open Data (2012). “5 Open Data”. Disponible en: <https://5stardata.info/en/>
- [9] ODI Open Data Institute. “The 2019 Data Skills Framework”. Disponible en: <https://theodi.org/article/open-data-skills-framework/>
- [10] Martínez, R. (2022). “Métricas de calidad para validar los conjuntos de datos abiertos públicos gubernamentales” (Doctoral dissertation, Universidad Nacional de La Plata).
- [11] Baloglu, O., Latifi, S. Q., & Nazha, A. (2022). “What is machine learning?”. *Archives of Disease in Childhood-Education and Practice*, 107(5), 386-388.
- [12] Mahesh, B. (2020). “Machine Learning Algorithms - A Review”. *International Journal of Science and Research (IJSR)*, 9, 7.426. doi:10.21275/ART20203995.
- [13] Caneiro, L. P., & Ortega, E. R. (2021). “Predicción de series temporales basada en Machine Learning: aplicaciones económicas y financieras. In *Nuevos métodos de predicción económica con datos masivos*” (pp. 189-214). Fundación de las Cajas de Ahorros (FUNCAS).
- [14] Oala, L., Murchison, A. G., Balachandran, P., Choudhary, S., Fehr, J., Leite, A. W., ... & Wiegand, T. (2021). “Machine learning for health: algorithm auditing & quality control”. *Journal of medical systems*, 45, 1-8.
- [15] Brei, V. A. (2020). “Machine learning in marketing: Overview, learning strategies, applications, and future developments”. *Foundations and Trends® in Marketing*, 14(3), 173-236.
- [16] Mithun S, U., Thonse, H., Rashmi, S., & hamed, N. (2021). “El papel del aprendizaje automático en el marketing digital”. *SAGE Open*.
- [17] Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., & Wrobel, S. (2019). “A review of machine learning for the optimization of production processes”. *The International Journal of Advanced Manufacturing Technology*, 104(5-8), 1889-1902.

Knowledge Graphs para accesibilidad y reutilización de activos digitales del Mar Argentino

Marcos Zárate^{1,2}, Carlos Buckle¹, Renato Mazzanti¹, Mirtha Lewis^{2,3}, Claudio Delrieux⁴,
Gustavo Nuñez¹, Darío Ceballos²

¹ Laboratorio de Investigación en Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco (LINVI-FI-UNPSJB), Puerto Madryn, Argentina.

² Centro para el Estudio de Sistemas Marinos, Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas (CESIMAR-CONICET), Puerto Madryn, Argentina.

³ Centro de Investigaciones y Transferencia Golfo San Jorge, (CIT-GSJ-CONICET), Comodoro Rivadavia, Argentina.

⁴ Departamento de Ing. Eléctrica y Computadoras, Universidad Nacional del Sur (DIEC-UNS), Bahía Blanca, Argentina

zarate@cenpat-conicet.gob.ar, cbuckle@unpata.edu.ar, renato@cenpat-conicet.gob.ar, mirtha@cenpat-conicet.gob.ar, cad@uns.edu.ar, guscostaf@gmail.com, disenodc@gmail.com

Resumen

El propósito de este proyecto es comprobar que la publicación de contenido como grafos de conocimiento (KG¹) mejorará la capacidad de localización, accesibilidad, interoperabilidad y reutilización de activos digitales (principios FAIR²) y vinculación con otros conjuntos de datos generados en la iniciativa Pampa Azul³. La disponibilidad de grafos y su posterior consulta permitirá reemplazar y potenciar la estructura actual de los datos (tablas, archivos, etc.) con mayor flexibilidad y conocimiento más preciso del dominio. En consecuencia, este enfoque facilitará que los científicos, técnicos y usuarios no expertos, descubran, analicen y utilicen estos datos de manera consistente para investigaciones interdisciplinarias realizadas en el marco del segundo lanzamiento de Pampa Azul (2020) en el cual, la gestión y modelado de datos en línea se considera fundamental para abordar la complejidad y el alcance de las temáticas con proyección amplia en el uso de la información.

En el proyecto participan docentes investigadores de la carrera de Licenciatura en Informática, especialistas del dominio de Ciencias del Mar, un becario doctoral CONICET y un becario EVC-CIN que desarrollará en este contexto su tesina de grado.

Palabras clave: Grafos de Conocimiento, Web Semántica, Datos Oceanográficos, Big Data.

Contexto

Este proyecto se inicia dentro del Grupo de Estudios de Macroecología Marina (GEMM@), perteneciente al CESIMAR-CONICET y el Laboratorio de Investigación en Informática de la Universidad Nacional de la Patagonia (LINVI-UNPSJB). Como se comentó previamente, el contexto de trabajo es el programa Pampa Azul. Este proyecto fue evaluado por FONCyT en el área Tecnología Informática de las Comunicaciones y Electrónica, con identificación PICT-2021-I-INVI-00245 y será financiado por la Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación para desarrollar entre 2023 y 2024.

Dado que este tipo de proyectos tienen una concepción interdisciplinaria, el grupo de

¹ KG: Knowledge Graphs, por sus siglas en inglés

² FAIR: Findable, Accessible, Interoperable, and Reusable, por sus siglas en inglés.

³ <https://www.pampazul.gob.ar/>

trabajo incluye: investigadores de las ciencias de la computación de diferentes áreas e investigadores de las ciencias biológicas, expertos en el dominio de aplicación.

1. Introducción

Para el nuevo enfoque de Pampa Azul es necesario desarrollar sistemas capaces de gestionar integración y comunicación de los datos capturados en campañas oceanográficas, tanto para un aprovechamiento integral de los grupos e instituciones participantes como para usuarios externos que requieran de información. El volumen y la heterogeneidad de activos digitales recolectados en estas campañas es muy amplio. Uno de los enfoques prometedores para abordar los problemas asociados con Big Data es almacenar la información con una estructura (real o latente) de representación de conjuntos de datos en forma de grafos [1]. Esto se desarrolla en estrecha relación con tecnologías de la Web Semántica [2], datos enlazados (en inglés Linked Data) [3], análisis de datos a gran escala y computación en la nube. El desarrollo de grafos de conocimiento, implica la publicación, el intercambio y la conexión de datos en la Web y ofrece una nueva forma de integración e interoperabilidad. Para ello es necesario gestionar conocimiento interdisciplinar y formar recursos humanos en los diferentes dominios. En el contexto internacional, encontramos diferentes iniciativas que utilizan grafos de conocimiento, la más destacada es GeoLink [4], proyecto financiado por la iniciativa EarthCube, la cual ha aprovechado los principios de datos enlazados para crear un conjunto de datos que permite a los usuarios consultar y razonar en algunos de los repositorios de metadatos de geociencia más destacados de los Estados Unidos.

2. Motivación

Previo a Pampa Azul (año 2009), el Ministerio de Ciencia, Tecnología e Innovación Productiva, desarrolló el sistema nacional de Datos Biológicos (SNDB) sobre

una plataforma denominada Integrated Publishing Toolkit (IPT) [5] desarrollada por Global Biodiversity Information Facility (GBIF) para datos biológicos basados en el estándar Darwin Core [6] el cual es ampliamente adoptado por la comunidad internacional. Para datos marinos se creó el sistema nacional de datos del mar (SNDM) sobre la plataforma desarrollada por International Oceanographic Data and Information Exchange (IODE), con el fin de visualizar la información de los centros nacionales productores de datos oceanográficos de Argentina de tipo satelitales, geológicos, batimétricos, biológicos y físico-químicos.

En la actualidad este portal de datos cayó en desuso, entre otras cosas porque no permite visualizar de manera interactiva los datos u otro tipo de información, convirtiéndose en una herramienta poco atractiva para los investigadores o público en general interesado en la oceanografía. Además, no permite ser integrado con otros repositorios de datos científicos. Esto generó un desinterés en contribuir con conjuntos de datos por parte de la comunidad científica. Los principales problemas en la gestión y explotación de datos del mar es que involucran marcos conceptuales de diferentes disciplinas, como ser la Oceanografía (física, química y biológica), las Geociencias y la Meteorología, entre las cuales existe una gran diversidad en los tipos y formatos de datos que deben gestionarse. En el contexto del nuevo lanzamiento de Pampa Azul, la gestión y modelado de datos en línea se considera fundamental para abordar la complejidad y el alcance de las temáticas que exigen una aproximación interdisciplinaria, de allí la fuerte motivación para este tipo de proyectos de investigación.

3. Líneas de Investigación y Desarrollo

Las líneas de investigación principales están ampliamente vinculadas con el Modelado Conceptual en la Web Semántica y la

construcción de grafos de conocimiento oceanográfico [7,8] mediante datos enlazados [9, 10] para la integración de datos científicos y su explotación [11, 12]. Pero además, será necesario abordar teorías, técnicas y herramientas para la visualización de datos científicos oceanográficos [13, 14].

4. Resultados esperados

Los resultados que se esperan obtener del proyecto son los siguientes:

1. Integrar una extensión del modelo conceptual BiGe-Onto [15]
2. Generar una versión estable del grafo de conocimiento OceanGraph [16] con ontologías propias del dominio de datos del mar y repositorios de datos enlazados.
3. Desarrollar una plataforma de visualización interactiva en línea para consultar y analizar la información proveniente de los grafos de conocimiento.
4. Desarrollar e incorporar sistemas de razonamiento para la explotación de conocimiento implícito en el grafo.
5. Capacitar recursos humanos para asegurar la sustentabilidad, incrementar la cantidad de conjuntos de datos provenientes de ciencias del mar.

Como resultado, se obtendrá una versión estable y mejorada de BiGe-Onto y OceanGraph, en la cual se utilizarán tecnologías open sources y otras off-the-shelf como razonadores y motores de consultas, y las requeridas por la nueva versión. Para continuar con el proceso de publicación de grafos de conocimiento seguiremos adoptando la metodología propuesta en [17] para gestionar las diferentes etapas del desarrollo y actualización del mismo. Este proceso tiene un modelo de ciclo de vida incremental iterativo, que se basa en la mejora continua y la extensión de los datos resultantes de la realización de varias iteraciones.

Durante el desarrollo de la investigación, se analizarán las tecnologías subyacentes a los grafos de conocimiento, por ejemplo,

ontologías propias del dominio, repositorios de datos enlazados, y cómo estos pueden ser integrados en nuestro grafo de conocimiento. Como resultado, se obtendrá una versión estable y mejorada de OceanGraph.

5. Formación de recursos humanos

Este proyecto de investigación da continuidad a una línea de trabajo que se viene realizando desde hace cinco años en conjunto entre el laboratorio LINVI-UNPSJB y el CESIMAR-CONICET en Puerto Madryn, que ha definido una política de formación de recursos humanos sostenida en esta área de la ciencia. En este proyecto en particular, se suma un Becario doctoral de CONICET y un becario EVC-CIN que realizará su tesina de grado estudiando y desarrollando plataformas de visualización para explotar los datos del grafo de conocimiento. Los directores de estas tesis y tesina son investigadores de la UNPSJB, de la UNS y del CESIMAR-CONICET.

Referencias

- [1] Ehrlinger, L., and W. Wöß. 2016. "Towards a Definition of Knowledge Graphs." SEMANTiCS (Posters, Demos, SuCESS). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1054.8298&rep=rep1&type=pdf>
- [2] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific American*, 284(5):34–43, May 2001.
- [3] Tom Heath and Christian Bizer. Linked data: Evolving the web into a global data space. *Synthesis lectures on the semantic web: theory and technology*, 1(1):1–136, 2011.
- [4] Cheatham, Michelle, Adila Krisnadhi, Reihaneh Amini, Pascal Hitzler, Krzysztof Janowicz, Adam Shepherd, Tom Narock, Matt Jones, and Peng Ji. 2018. "The GeoLink Knowledge Graph." *Big Earth Data* 2 (2): 131–43.

- [5] Robertson, Tim, Markus Döring, Robert Guralnick, David Bloom, John Wiecek, Kyle Braak, Javier Otegui, Laura Russell, and Peter Desmet. 2014. "The GBIF Integrated Publishing Toolkit: Facilitating the Efficient Publishing of Biodiversity Data on the Internet." *PloS One* 9 (8): e102623.
- [6] Wiecek, John, David Bloom, Robert Guralnick, Stan Blum, Markus Döring, Renato Giovanni, Tim Robertson, and David Vieglais. 2012. "Darwin Core: An Evolving Community-Developed Biodiversity Data Standard." *PloS One* 7 (1): e29715.
- [7] Marcos Zárate, Pablo Rosales, Pablo Fillottrani, Claudio Delrieux, and Mirtha Lewis. Oceanographic data management: Towards the publishing of pampa azul oceanographic campaigns as linked data. In *Proceedings of the 12th Alberto Mendelzon International Workshop on Foundations of Data Management (AMW 2018)*, 2018.
- [8] Marcos Zárate, Pablo Rosales, Germán Braun, Mirtha Lewis, Pablo Rubén Fillottrani, and Claudio Delrieux. OceanGraph: Some initial steps toward a oceanographic knowledge graph. In *Iberoamerican Knowledge Graphs and Semantic Web Conference*, pages 33–40. Springer, 2019.
- [9] Krzysztof Janowicz, Pascal Hitzler, Benjamin Adams, Dave Kolas, II Vardeman, et al. Five stars of linked data vocabulary use. *Semantic Web*, 5(3):173–176, 2014.
- [10] Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts*, pages 205–227. IGI Global, 2011.
- [11] Judith A Blake and Carol J Bult. Beyond the data deluge: data integration and bioontologies. *Journal of biomedical informatics*, 39(3):314–320, 2006.
- [12] Giuseppe Andronico, Valeria Ardizzone, Roberto Barbera, Bruce Becker, Riccardo Bruno, Antonio Calanducci, Diego Carvalho, Leandro Ciuffo, Marco Fargetta, Emidio Giorgio, Giuseppe Rocca, Alberto Masoni, Marco Paganoni, Federico Ruggieri, and Diego Scardaci. e-infrastructures for e-science: A global view. *Journal of Grid Computing*, 9:155–184, 06 2011.
- [13] Tanu Malik and Ian Foster. Addressing data access needs of the long-tail distribution of geoscientists. In *Geoscience and Remote Sensing Symposium (IGARSS)*, 2012 IEEE International, pages 5348–5351. IEEE, 2012.
- [14] Peter Fox and James Hendler. Changing the equation on scientific data visualization. *Science*, 331(6018):705–708, 2011.
- [15] Zárate, M., Braun, G., Fillottrani, P., Delrieux, C., & Lewis, M. (2020). BiGe-Onto: an ontology-based system for managing biodiversity and biogeography data. *Applied Ontology*, 15(4), 411–437.
- [16] Zárate, M., Rosales, P., Braun, G., Lewis, M., Fillottrani, P. R., & Delrieux, C. (2019). OceanGraph: Some initial steps toward a oceanographic knowledge graph. In *Knowledge Graphs and Semantic Web* (pp. 33–40). Springer International Publishing. https://doi.org/10.1007/978-3-030-21395-4_3
- [17] Fensel, Dieter, Umutcan Şimşek, Kevin Angele, Elwin Huaman, Elias Kärle, Oleksandra Panasiuk, Ioan Toma, Jürgen Umbrich, and Alexander Wahler. 2020. *Knowledge Graphs: Methodology, Tools and Selected Use Cases*. Springer Nature.

Descubrimiento de patrones de comportamiento vinculados al abandono en la Universidad Nacional de Luján mediante la aplicación de técnicas de aprendizaje automático

Mario G. Oloriz¹, Juan M. Fernandez¹, Carlos Jara², Carla Martinez²,
Ruben Baquel², Sebastián Bertoglio², Hugo Delfino²

{moloriz, jmfernandez, cjara, cmartinez, rbaquel, sbertoglio, hdelfino}@unlu.edu.ar,

¹División Computación, Departamento de Ciencias Básicas, Universidad Nacional de Luján

²División Estadística, Departamento de Ciencias Básicas, Universidad Nacional de Luján

Resumen

El estudio del abandono en la educación superior resulta un tema de relevancia dado el continuo debate que se propicia respecto de los recursos públicos que las naciones afectan a su sostenimiento, principalmente en Argentina, donde el subsistema de gestión pública abarca más del 75 % de la matrícula total del sistema. Las altas tasas de abandono ponen en tensión la política de gratuidad e ingreso irrestricto que se aplica históricamente y que es reconocido como uno de los medios para la movilidad social.

La multicausalidad del fenómeno requiere de su estudio y análisis desde diferentes perspectivas así como de la definición de políticas académicas para tratar de disminuir el fracaso educativo en el nivel superior.

Por su parte, en la actualidad, existen numerosas técnicas de ciencia de datos mediante las cuales es posible extraer conocimiento a partir de bases de datos, las cuales están siendo aplicadas a áreas muy disímiles. Estas técnicas permiten encontrar patrones que no podrían ser descubiertos mediante las técnicas convencionales de análisis de datos multivariantes.

En este trabajo se describen las acciones emprendidas en el marco del proyecto de investigación “Descubrimiento de patrones de comportamiento vinculados al abandono en la Universidad Nacional de Luján mediante la aplicación de técnicas de aprendizaje automático”, así como las líneas de I+D comprendidas en el mismo.

Contexto

Este trabajo se encuentra en el marco del Proyecto de Investigación “Descubrimiento de patrones de comportamiento vinculados al abandono en la Universidad Nacional de Luján mediante la aplicación de técnicas de aprendizaje automático”, aprobado mediante Disposición PCDD-CB N°027/23 en el marco de los Proyectos de Investigación PI2+, Disposición

CDD-CB N°139/19, del Departamento de Ciencias Básicas de la Universidad Nacional de Luján.

El objetivo principal de este proyecto consiste en identificar comportamientos vinculados al abandono en el ámbito de la Universidad Nacional de Luján, así como también a nivel nacional, a efectos de generar modelos y acciones que permitan disminuir este fenómeno negativo a partir de la aplicación de técnicas y estrategias propias de aprendizaje automático.

Introducción

Con la masificación de internet y la digitalización de los procesos organizacionales, surge una creciente necesidad de desarrollar herramientas que permitan extraer información útil (conocimiento) a partir de grandes volúmenes de datos, los cuales están en constante crecimiento. De la mano de esta necesidad, aparece a mediados de la década del 90, un campo emergente denominado descubrimiento de conocimiento en bases de datos (KDD) [13].

Históricamente, la noción de encontrar patrones útiles en los datos ha recibido diversos nombres, incluyendo minería de datos, extracción de conocimiento, descubrimiento de información, recolección de información, arqueología de datos o procesamiento de patrones. Actualmente, se ha acuñado el término ciencia de datos para denominar esta disciplina, la cual se ubica en la intersección entre las disciplinas de ciencias de la computación, estadística y sus ámbitos de aplicación [50].

Existen numerosas técnicas de ciencia de datos mediante las cuales es posible realizar ese proceso de extracción de conocimiento a partir de bases de datos [22]. En este sentido, abordando la problemática desde el punto de vista del objetivo del descubrimiento de conocimiento, existe cierto consenso en categorizar en cuatro tipos principales a estas tareas: clasificación, predicción numérica, asociación y agrupamiento [8].

A su vez, la ciencia de datos es aplicada para la

generación de conocimiento en infinidad de áreas totalmente disímiles, permitiendo encontrar patrones que no podrían haber sido descubiertos mediante las técnicas convencionales de análisis de datos multivariantes. Ejemplos de ello son mejoras de procesos de eficiencia energética [25], detección y tratamiento de enfermedades como COVID-19 [4], medicina del deporte [29], preferencias de investigación en sociología [20], entre otros.

Por su parte, el estudio del abandono en la educación superior es un tema de relevancia dado el continuo debate que se propicia respecto de los recursos públicos que las naciones afectan al sostenimiento de este nivel educativo, principalmente en Argentina, donde el subsistema de gestión pública abarca casi el 75 % de la matrícula total del sistema [12]. Las altas tasas de abandono ponen en tensión la política de gratuidad e ingreso irrestricto que se aplica históricamente y que es reconocido como uno de los medios para la movilidad social [17].

En el contexto descripto, esta investigación se orienta a estudiar la aplicación de técnicas de ciencia de datos para la explotación de información de bases de datos pertenecientes al ámbito de la Educación Superior a efectos de extraer conocimiento relacionado con comportamientos poblacionales, en particular el abandono de los estudios y su estrecha relación con el rendimiento académico, fundamentalmente en la Universidad Nacional de Luján (UNLu).

Antecedentes

En los últimos años, se han desarrollado diversas investigaciones que aplican técnicas de aprendizaje automático al ámbito educativo, entre las cuales se puede mencionar la detección de fraude en exámenes [24], la clasificación automática de consultas académicas [14], la predicción de preferencias sobre el uso de plataformas de aprendizaje móvil [2] así como la predicción del rendimiento académico en los estudios [54, 55] e incluso estudios del mismo fenómeno a partir del impacto producido por la pandemia originada por COVID-19 [52], entre muchos otros. A su vez existen diversas experiencias alrededor del mundo sobre la aplicación de un enfoque de ciencia de datos para la detección del abandono en educación así como la formulación de modelos que permitan predecir este fenómeno [3, 7, 26, 47, 49].

Así como en la UNLu, en la mayoría de las Universidades públicas de Argentina, el ingreso es irrestricto a las Carreras de grado y pregrado. Las publicaciones oficiales, correspondientes al año 2016, indican que existen alrededor de 1.940.000 estudiantes universitarios de grado y pregrado y 161.000 de carreras de postgrado. A su vez, algo más del 78 % de estas personas cursan carreras de grado y pregrado en instituciones de gestión pública y poco más del 76 % carreras de posgrado en el mismo tipo de institucio-

nes. La misma publicación, da cuenta que ingresan al nivel universitario alrededor de 450.000 estudiantes por año y egresan por año, aproximadamente, 120.000 (tasa de egreso del 26 %).

Un problema señalado de manera recurrente en toda Latinoamérica son los comportamientos poblacionales de abandono y el rendimiento académico decreciente como problemáticas que deben ser estudiadas.

El concepto de abandono ha sido muy discutido académicamente, no obstante resulta claro que para los Estados implica un enorme costo, mientras que para el estudiante puede significar algún tipo de frustración y requiere de políticas de estado que posibiliten anticipar este fenómeno y mitigarlo.

Los especialistas presentan diversas hipótesis que intentan explicar la naturaleza de esta problemática. Algunos encuentran explicación en dos teorías sociológicas; por un lado “El modelo de integración del estudiante” [53] donde la integración del estudiante al mundo académico afecta en forma directa a la determinación de abandonar o no los estudios, y por el otro es el “Modelo de desgaste del estudiante” [6] que otorga relevancia a los factores externos a la institución universitaria.

Por su parte, Hackman y Dysinger plantean que el problema fundamental del abandono tiene que ver con la ausencia de interés y no con la imposibilidad por parte del estudiante de cumplir con los requisitos que la Universidad exige [19].

Mientras tanto, otros autores plantean el abordaje desde distintas dimensiones de análisis (Psicológicos, Económicos, Sociológicos, Organizacionales y de Interacción) de acuerdo a la relevancia que se le otorga a las variables que intentan explicar el fenómeno del abandono y retención, sean familiares, individuales o institucionales [9].

Asimismo, existen trabajos en los cuales se concluye que una de las causales principales del abandono tienen origen en temas relacionados con procesos vocacionales [5].

Como se observa, existe una variada cantidad de interpretaciones que intentan explicar el fenómeno del abandono aunque sin embargo, al tratarse de una problemática muy compleja y multicausal, ninguna da por agotado el tema.

En estudios recientes, realizados en el ámbito de la Universidad Nacional de Luján (UNLu), aplicando técnicas estadísticas multivariantes tradicionales, se han observado una serie de cuestiones que sirven de indicio para otras investigaciones. Por un lado, se ha logrado demostrar, parcialmente, que la institución cumple el fin social para el cual fue creada dado que quienes transitan por una de las Carreras que allí se dictan muestran un mejor desempeño en el ámbito laboral que quienes no lo hicieron. [36, 15].

En relación a la cuantificación de la problemática de abandono, objeto de este estudio, se ha encontrado, en un estudio longitudinal de once cohortes de

la UNLu, que el 71.3 % de quienes ingresan terminan abandonando la carrera una vez transcurrida su duración teórica [37]. Al mismo tiempo, en el estudio citado, se comprobó que el 65.5 % del abandono se produce de manera temprana (llegando al 46.7 % de la matrícula) mientras que el 34.5 % restante se produce de manera tardía (el 24.6 % de la matrícula).

En el mismo sentido, intentando generar una taxonomía de abandono para la Universidad Nacional de Luján, las investigaciones previas permiten sostener las siguientes conclusiones:

- La tasa de abandono aumenta en función del aumento de la edad del estudiante al momento de iniciar los estudios universitarios.
- No se han encontrado evidencias sobre que el tipo de gestión del establecimiento secundario del cual egresó el estudiante influya respecto del abandono de los estudios universitarios.
- Tal como lo verificaron otros autores, se encontró que la probabilidad de abandonar los estudios crece en función de la cantidad de horas que trabaja el estudiante [18].
- No se observan diferencias significativas respecto de la relación entre el abandono de los estudios y el género del estudiante.
- Se observa que el estado civil podría influir en la probabilidad que el estudiante abandone la carrera dado que para los estudiantes solteros la tasa de abandono es sustantivamente inferior que la observada para las restantes categorías de la variable estado civil.
- La tasa de abandono aumenta en función de la cantidad de tiempo transcurrido desde el momento en que el estudiante termina el nivel secundario y empieza a cursar el nivel universitario.
- La tasa de abandono es menor para quienes viven en la misma ciudad en que se encuentra la Sede en que cursan o en ciudades que se encuentran a una distancia entre 61 y 200 km.
- Se observó que, para la cohorte 2013, no se verifica que el promedio alcanzado en el nivel secundario pueda utilizarse como un predictor directo del abandono durante el primer año [44].

En cuanto al rendimiento académico, si bien algunos autores sostienen que el concepto de rendimiento nace en las sociedades industriales para referirse a normas, criterios y procedimientos de medida en el ámbito laboral y relacionado más específicamente a la productividad del trabajo [10]. Este concepto, aplicado a quienes cursan estudios universitarios, es lo que se define como rendimiento académico.

En el ámbito de la UNLu, se ha estudiado este fenómeno encontrando que existen algunas modalidades de dictado de las actividades académicas que favorecen el rendimiento académico de los estudiantes [32]. A su vez, se ha constatado en que no existen prácticamente diferencias significativas en el desempeño académico de los estudiantes que ingresan con título de nivel medio de aquellos que lo hacen mediante la modalidad prevista por el artículo 7° de la Ley de Educación Superior, mayores de 25 años de edad sin estudios secundarios completos [35].

Otro de los aspectos estudiados con técnicas analíticas tradicionales, es el impacto de los programas de becas en el abandono y el rendimiento académico, encontrando que las mismas permiten una disminución del abandono (de hasta el 36 %) y la mejora del rendimiento académico si bien el nivel de efectividad varía de acuerdo al campo disciplinar de la Carrera que cursa el estudiante y el monto que percibe como estipendio incide en la efectividad a medida que avanza en la carrera [45, 39, 43, 38, 33].

También se encontraron efectos positivos de los programas de becas a nivel nacional al correlacionar los fondos que cada institución universitaria aplica a esos programas con el abandono interanual, para el período 2013-2014 [40].

Otro de los aspectos que se abordaron, es la incidencia entre la cantidad de estudiantes por docente en el abandono de los estudios superiores, concluyendo que la relación entre esas dos variables solo pudo validarse en las instituciones clasificadas como nuevas, o sea aquellas que no tienen aún 12 años de funcionamiento. En los demás casos, ya sea agrupando por tamaño o agrupando por región, no se observó correlación estadísticamente significativa entre estos dos indicadores [34].

En las carreras de postgrado, también el abandono es un problema de relevancia dada la baja tasa de graduación que se da en este nivel. En un estudio de las cohortes 2010 a 2014, de la Universidad Nacional de Luján, se calculó un abandono superior al 50 % de la matrícula, dejando de lado el problema de la elaboración de tesis o trabajos finales [41].

Finalmente, dada la sustitución de las clases presenciales por clases remotas o mediadas por las TIC, que se produjo durante los años 2020 y 2021 debido a la pandemia por COVID-19, se abordó el estudio del impacto en el abandono y el rendimiento académico de quienes se encontraban cursando carreras de pregrado y grado [42, 16, 46].

Estos hallazgos sobre abandono y rendimiento académico, obtenidos mediante el abordaje tradicional basado en métodos estadísticos pueden ser enriquecidos con el abordaje que propone la explotación de información para encontrar piezas de conocimiento sin ninguna presunción previa a partir de la masa de información disponible en los sistemas de gestión académica universitaria. Desde esta hipóte-

sis, se plantean las líneas de I+D que se explicitan y abordan a continuación.

Líneas de I+D

A continuación se presentan y describen brevemente las principales líneas de I+D en las cuales se trabaja actualmente en el marco de este proyecto.

a. Determinación de las variables más importantes para el análisis del abandono en los estudios

La selección de características, en inglés *feature selection*, tiene como objetivo elegir un subconjunto de características relevantes a partir de las características originales de un conjunto de datos, eliminando las características irrelevantes, redundantes o ruidosas y maximizando la relevancia para la variable objetivo, como las etiquetas de clase en la clasificación. [31]. Las técnicas más representativas de este enfoque incluyen la ganancia de información, relieve, puntuación de Fisher y Lazo [51].

La selección de características está comprendida dentro de las técnicas de reducción de dimensionalidad que también comprenden a las técnicas de extracción de características [51]. Complementariamente, estos enfoques de extracción de características proyectan características en un nuevo espacio de características con menor dimensionalidad y las nuevas características construidas suelen ser combinaciones de las originales. Los ejemplos de técnicas de extracción de características incluyen el análisis de componentes principales (PCA), el análisis discriminante lineal (LDA) y el análisis de correlación canónica (CCA).

Esta línea de investigación tiene como objetivo identificar, a partir de los repositorios de datos disponibles relacionados con el abandono y el rendimiento académico en la educación superior, el conjunto de atributos o características más representativas para la predicción del abandono y los grupos de riesgo en la educación superior.

b. Caracterización de poblaciones estudiantiles

El *clustering* o agrupamiento es el proceso de agrupar objetos similares en diferentes grupos de acuerdo a alguna medida de distancia definida [27]. La agrupación puede considerarse el problema de aprendizaje no supervisado más importante y como tal, se trata de encontrar una estructura en una colección de datos no etiquetados. Por lo tanto, un clúster es una colección de objetos que son similares entre sí y son diferentes a los objetos pertenecientes a otros clústeres.

Dado su potencial, un importante número de estudios de investigación se centran en varios aspectos del análisis de *clustering* por ejemplo, algoritmos de agrupamiento y extensiones para un tipo de datos particular, métricas de disimilitud (o distancia), número óptimo de conglomerados, relevancia de los atributos de datos por conglomerado, evaluación de los resultados del agrupamiento, y conjuntos de grupos [21].

En esta línea de I+D se propone la utilización de técnicas de aprendizaje no supervisado, particularmente de *clustering*, para la caracterización de la población estudiantil a partir de las fuentes de datos disponibles.

c. Generación de modelos para la predicción del abandono

La predicción del abandono puede identificarse como un problema supervisado, ya sea de predicción o clasificación, de acuerdo al tipo de variable objetivo utilizada.

En un modelo de clasificación, luego de obtener los datos, realizar el preprocesamiento de los mismos para la extracción de características, realizar el etiquetado y avanzar en un esquema de representación, se entrena el clasificador utilizando distintos enfoques o algoritmos [48], como el aprendizaje bayesiano, la regresión logística, redes neuronales, árboles de decisión y máquinas de vectores soporte.

El modelo generado a partir del entrenamiento debe ser capaz de capturar las características distintivas del conjunto de entrenamiento para luego poder analizar otras instancias no observadas previamente, lográndose así la capacidad de generalización del clasificador que se suele evaluar sobre otro conjunto de prueba separado [28].

A la fecha, y debido a la cantidad de algoritmos de aprendizaje existentes, resulta muy complejo sistematizar todos los abordajes posibles. De acuerdo a los antecedentes estudiados, los algoritmos escogidos que se están utilizando en el marco de esta investigación son el clasificador de Naive Bayes [30], la regresión logística [50], las máquinas de vector soporte [23], las redes neuronales recurrentes LSTM (*Long short-term memory*) [1] y XGBoost [11].

Las aspiraciones en torno a esta línea de estudio son generar diferentes modelos de predicción del abandono teniendo en cuenta las estrategias de caracterización de poblaciones estudiantiles y las características más importantes emergentes de la etapa de extracción de características.

d. Definición de políticas de mitigación del abandono en los estudios

El desarrollo de las líneas de I+D introducidas hasta aquí debieran favorecer la generación de cono-

cimiento asociado al descubrimiento de patrones de comportamiento vinculados al abandono en la Universidad Nacional de Luján, así como en el Sistema de Educación Superior Argentino.

A partir de ese conocimiento, se plantea el objetivo de proponer estrategias que mitiguen el abandono y permitan brindar acompañamiento institucional para mejorar el rendimiento académico de los grupos de riesgo.

Tanto las Instituciones de Educación Superior como los organismos nacionales vinculados con la coordinación del sistema educativo, demandan la generación de información científica que permita la generación de proyectos y la fundamentación de diferentes líneas de acción que propendan a la mitigación del abandono en la educación superior.

Objetivos

A partir de las líneas de investigación detalladas antes, en el marco de este proyecto se persiguen los siguientes objetivos específicos:

- Identificar, mediante técnicas de ciencia de datos, en el conjunto de variables descriptivas, el subconjunto de variables relacionadas con comportamientos poblacionales en educación superior.
- Identificar patrones relacionados con causales de abandono y disminución en el rendimiento académico de estudiantes de la Universidad Nacional de Luján.
- Proponer una configuración ideal para la aplicación de las técnicas de ciencias de datos a este tipo de problemáticas.
- Desarrollar modelos que permitan modificar estos patrones de comportamiento poblacional vinculados al abandono a efectos de revertir estas situaciones.

Complementariamente, su transferencia a la comunidad universitaria para la generación de políticas académicas que busquen la reducción del abandono resulta de alto interés. Las herramientas a desarrollar pueden ser aplicadas a diversos espacios académicos, lo cual plantea una oportunidad concreta de transferencia tanto en el sector público como en el privado.

Recursos Humanos

Se espera que este proyecto contribuya a consolidar un grupo de investigación en la temática y brindar un ámbito adecuado para la formación de recursos humanos, en el marco de la Universidad Nacional de

Luján, que incorporen saberes y competencias provenientes de la participación en actividades de investigación.

Concretamente, se ha incluido como integrantes a docentes auxiliares y se convocará a estudiantes que participen como pasantes esperando que el proyecto sea el ámbito para la realización de, al menos, una Tesis de Licenciatura en Sistemas de Información y una tesis de Licenciatura en Gestión Universitaria por año así como la culminación de al menos una Tesis de Maestría en Generación y Análisis de Información Estadística.

Referencias

- [1] AGGARWAL, C. C., ET AL. Neural networks and deep learning. *Springer 10* (2018), 978–3.
- [2] AKOUR, I., ALSHURIDEH, M., AL KURDI, B., AL ALI, A., SALLOUM, S., ET AL. Using machine learning algorithms to predict people's intention to use mobile learning platforms during the covid-19 pandemic: machine learning approach. *JMIR Medical Education* 7, 1 (2021), e24032.
- [3] ALI, M., AND ABDEL-HAQ, M. K. Bibliographical analysis of artificial intelligence learning in higher education: is the role of the human educator and educated a thing of the past? In *Fostering Communication and Learning With Underutilized Technologies in Higher Education*. IGI Global, 2021, pp. 36–52.
- [4] ALYASSERI, Z. A. A., AL-BETAR, M. A., DOUSH, I. A., AWADALLAH, M. A., ABASI, A. K., MAKHADMEH, S. N., ALOMARI, O. A., ABDULKAREEM, K. H., ADAM, A., DAMASEVICIUS, R., ET AL. Review on covid-19 diagnosis models based on machine learning and deep learning approaches. *Expert systems* 39, 3 (2022), e12759.
- [5] APARICIO, M., AND GARZUZI, V. Dinámicas identitarias, procesos vocacionales y su relación con el abandono de los estudios: Un análisis en alumnos ingresantes a la universidad. *Revista de Orientación Educacional* 20, 37 (2006), 15–36.
- [6] BEAN, J. P. Student attrition, intentions, and confidence: Interaction effects in a path model. *Research in higher education* (1982), 291–320.
- [7] BILQUISE, G., ABDALLAH, S., AND KOBBAEY, T. Predicting student retention among a homogeneous population using data mining. In *Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges*. Springer, 2020, pp. 243–260.
- [8] BRAMER, M., AND BRAMER, M. Data for data mining. *Principles of data mining* (2016), 9–19.

- [9] BRAXTON, J. M. *Reworking the student departure puzzle*. Vanderbilt University Press, 2000.
- [10] CAMARENA, R., CHÁVEZ, A. M., AND GÓMEZ, J. Reflexiones en torno al rendimiento escolar ya la eficiencia terminal. *Revista de la educación superior* 53, 3 (1985).
- [11] CHEN, T., AND GUESTRIN, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (2016), pp. 785–794.
- [12] DE EDUCACIÓN DE LA NACIÓN, M. Anuarios estadísticos universitarios, 2020.
- [13] FAYYAD, U., PIATETSKY-SHAPIO, G., AND SMYTH, P. From data mining to knowledge discovery in databases. *AI magazine* 17, 3 (1996), 37–37.
- [14] FERNÁNDEZ, J. M., AND ERRECALDE, M. Multi-class e-mail classification with a semi-supervised approach based on automatic feature selection and information retrieval. In *Cloud Computing, Big Data & Emerging Topics: 10th Conference, JCC-BD&ET 2022, La Plata, Argentina, June 28–30, 2022, Proceedings* (2022), Springer, pp. 75–90.
- [15] FERNÁNDEZ, J. M., AND OLORIZ, M. G. Relación entre el cursado de la tecnicatura en administración y gestión universitaria en la universidad nacional de luján y el avance en la carrera administrativa.
- [16] FERNÁNDEZ, J. M., OLORIZ, M. G., AND PUGGIONI, N. A. Efectos de la modalidad remota en el desempeño académico según campo disciplinar. In *Congresos CLABES* (2021).
- [17] GARCÍA DE FANELLI, A. Inclusión social en la educación superior argentina: Indicadores y políticas en torno al acceso ya la graduación. *Páginas de Educación* 7, 2 (2014), 124–151.
- [18] GIOVAGNOLI, P. I. Determinantes de la deserción y graduación universitaria: una aplicación utilizando modelos de duración. *Documentos de trabajo* (2002).
- [19] HACKMAN, J. R., AND DYSINGER, W. S. Commitment to college as a factor in student attrition. *Sociology of education* (1970), 311–324.
- [20] HEIBERGER, R. H. Applying machine learning in sociology: How to predict gender and reveal research preferences. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 74, Suppl 1 (2022), 383–406.
- [21] IAM-ON, N., AND BOONGOEN, T. Generating descriptive model for student dropout: a review of clustering approach. *Human-centric Computing and Information Sciences* 7 (2017), 1–24.
- [22] IGUAL, L., SEGUÍ, S., IGUAL, L., AND SEGUÍ, S. *Introduction to data science*. Springer, 2017.
- [23] JOACHIMS, T., ET AL. Transductive inference for text classification using support vector machines. In *Icml* (1999), vol. 99, pp. 200–209.
- [24] KAMALOV, F., SULIEMAN, H., AND SANTANDREU CALONGE, D. Machine learning based approach to exam cheating detection. *Plos one* 16, 8 (2021), e0254340.
- [25] KHAYYAM, H., NAEBE, M., MILANI, A. S., FAKHRHOSEINI, S. M., DATE, A., SHABANI, B., ATKISS, S., RAMAKRISHNA, S., FOX, B., AND JAZAR, R. N. Improving energy efficiency of carbon fiber manufacturing through waste heat recovery: A circular economy approach with machine learning. *Energy* 225 (2021), 120113.
- [26] LUAN, H., AND TSAI, C.-C. A review of using machine learning approaches for precision education. *Educational Technology & Society* 24, 1 (2021), 250–266.
- [27] MADHULATHA, T. S. An overview on clustering methods. *arXiv preprint arXiv:1205.1117* (2012).
- [28] MARIÑELARENA-DONDENA, L., ERRECALDE, M. L., AND SOLANO, A. C. Extracción de conocimiento con técnicas de minería de textos aplicadas a la psicología. *Revista Argentina de Ciencias del Comportamiento* 9, 2 (2017), 65–76.
- [29] MARTIN, R. K., PAREEK, A., KRYCH, A. J., KREMERS, H. M., AND ENGBRETSSEN, L. Machine learning in sports medicine: need for improvement. *Journal of ISAKOS* 6, 1 (2021), 1–2.
- [30] MCCALLUM, A., NIGAM, K., ET AL. A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization* (1998), Citeseer, pp. 41–48.
- [31] MIAO, J., AND NIU, L. A survey on feature selection. *Procedia Computer Science* 91 (2016), 919–926.
- [32] OLORIZ, M., AND FERNÁNDEZ, J. M. La acreditación de competencias en reemplazo de actividades curriculares: ¿mejora el tránsito de los estudiantes por la propuesta formativa?. el caso de la carrera de ingeniería industrial de la unlu.

- [33] OLORIZ, M., FERNÁNDEZ, J. M., AND AMADO, V. Impacto del programa de becas de la universidad nacional de luján en la disminución del abandono. In *III Conferencia Latinoamericana sobre el Abandono en la Educación Superior. CLABES* (2013).
- [34] OLORIZ, M., FERNANDEZ, J. M., AND BATTO, M. Relación entre la planta docente y el abandono interanual en las universidades argentinas.
- [35] OLORIZ, M. G., AND FERNANDEZ, J. Rendimiento académico de los estudiantes que ingresaron sin estudios de nivel medio completos en la universidad nacional de luján. *XII Coloquio Internacional sobre Gestión Universitaria de América en Sur. Veracruz, México* (2012).
- [36] OLORIZ, M. G., AND FERNÁNDEZ, J. M. El impacto de transitar por una carrera universitaria en la carrera laboral. el caso de la tecnicatura en administración y gestión universitaria en la universidad nacional de luján.
- [37] OLORIZ, M. G., AND FERNÁNDEZ, J. M. Relación entre las características del estudiante al momento de iniciar estudios superiores y el abandono en la universidad nacional de luján durante el período 2000-2010. In *Congresos CLABES* (2013).
- [38] OLORIZ, M. G., AND FERNANDEZ, J. M. Impacto del programa de becas estudiantiles de la universidad nacional de luján por campo disciplinar. In *Congresos CLABES* (2014).
- [39] OLORIZ, M. G., AND FERNÁNDEZ, J. M. Rendimiento académico de los estudiantes de la universidad nacional de luján que participan en los programas de becas internos y nacionales.
- [40] OLORIZ, M. G., AND FERNÁNDEZ, J. M. Estudio del abandono interanual en el sistema universitario argentino y su relación con los recursos aplicados a los programas de becas. In *Congresos CLABES* (2016).
- [41] OLORIZ, M. G., AND FERNÁNDEZ, J. M. El abandono en las carreras de postgrado de la universidad nacional de luján. In *Congresos CLABES* (2018).
- [42] OLORIZ, M. G., AND FERNÁNDEZ, J. M. El impacto de la pandemia por covid-19 en el abandono temprano de los estudios superiores. In *Congresos CLABES* (2021).
- [43] OLORIZ, M. G., FERNANDEZ, J. M., AND AMADO, M. V. Comparación entre el programa de becas estudiantiles de la universidad nacional de luján y los programas de becas nacionales. In *Congresos CLABES* (2014).
- [44] OLORIZ, M. G., FERNÁNDEZ, J. M., AND RODRÍGUEZ, L. S. Relación entre promedio final en el nivel secundario y el abandono de los estudios superiores. In *Congresos CLABES* (2014).
- [45] OLORIZ, M. G., FERNÁNDEZ, J. M., AND RODRIGUEZ, R. Impacto del programa de “respaldo a estudiantes argentinos” en la disminución del abandono en la universidad nacional de luján. In *Congresos CLABES* (2015).
- [46] OLORIZ, M. G., FERRERO, E. L., AND LUCHINI, M. L. Impacto de la suspensión de correlatividades en el cursado de actividades académicas. In *Congresos CLABES* (2021).
- [47] PALACIOS, C. A., REYES-SUÁREZ, J. A., BEARZOTTI, L. A., LEIVA, V., AND MARCHANT, C. Knowledge discovery for higher education student retention based on data mining: Machine learning algorithms and case study in chile. *Entropy* 23, 4 (2021), 485.
- [48] RUSSELL STUART, J., AND NORVIG, P. *Artificial intelligence: a modern approach*. Prentice Hall, 2009.
- [49] SHAFIQ, D. A., MARJANI, M., HABEEB, R. A. A., AND ASIRVATHAM, D. Student retention using educational data mining and predictive analytics: A systematic literature review. *IEEE Access* (2022).
- [50] SKIENA, S. S. *The data science design manual*. Springer, 2017.
- [51] TANG, J., ALELYANI, S., AND LIU, H. Feature selection for classification: A review. *Data classification: Algorithms and applications* (2014), 37.
- [52] TARIK, A., AISSA, H., AND YOUSEF, F. Artificial intelligence and machine learning to predict student performance during the covid-19. *Procedia Computer Science* 184 (2021), 835–840.
- [53] TINTO, V. Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research* 45, 1 (1975), 89–125.
- [54] TJANDRA, E., KUSUMAWARDANI, S. S., AND FERDIANA, R. Student performance prediction in higher education: A comprehensive review. In *AIP Conference Proceedings* (2022), vol. 2470, AIP Publishing LLC, p. 050005.
- [55] YAĞCI, M. Educational data mining: prediction of students’ academic performance using machine learning algorithms. *Smart Learning Environments* 9, 1 (2022), 11.

DESARROLLO DE SOLUCIONES DE INTELIGENCIA Y ANALITICA DE NEGOCIOS

María Inés Lund¹, Silvina I. Migani¹, Cristina Vera¹, Silvina Balmaceda¹, Cintia Ferrarini O¹., Esteban Gil M.^{1*}, Martín Manassero^{1*}, María Laura Molina^{1*}, Juan Capdevila^{1*}, Leandro Drazic^{1*}, Daniel Tapella², Fernando Más², Ramiro Savoie³

¹Departamento de Informática, FCEFN, Universidad Nacional de San Juan

^{1*}Alumno Departamento de Informática, FCEFN, Universidad Nacional de San Juan

²Gobierno de la Provincia de San Juan

³Instituto Tecnológico Buenos Aires

mlund@unsj-cuim.edu.ar, {silvina.migani, civerados, balmaceda.silvina, ferrarinicintia, manasseromm, lauramolinaf86, juan.capdevila27, leandrodrazic, danieltapella, famascba}@gmail.com, ramiro@deployr.ai

RESUMEN

Este trabajo tiene como propósito la definición de un conjunto ordenado de procesos que conduzcan al desarrollo de soluciones analíticas, donde los objetivos estratégicos del negocio sean contemplados en fases tempranas, de manera que se logren sistemas analíticos valiosos y oportunos para los tomadores de decisión de cualquier organización. Asimismo, se pretende aplicarlo en la construcción de un sistema analítico que brinde apoyo a la toma de decisiones de la Secretaría de Tránsito y Transporte (STyT) del Gobierno de la Provincia de San Juan (GPSJ), en temas relativos al Sistema Único de Boleto Electrónico (SUBE).

Palabras claves: Analítica e Inteligencia Empresarial - Estrategia Organizacional.

CONTEXTO

Esta presentación se enmarca dentro del proyecto “ANALÍTICA E INTELIGENCIA DE NEGOCIO. DISEÑO DE UN MODELO DE PROCESOS DE NEGOCIO”, el cual toma como base antecedentes y resultados alcanzados en proyectos anteriores:

- 21/E1122. “Aplicación de la ciencia de los datos para determinar la calidad de vida de la población del gran san juan en relación con la nación”. Financiamiento UNSJ. Res N 0591-20/R.
- 21/E1131. “Inteligencia de Negocios. Tecnologías, aplicaciones y tendencias”. Financiamiento UNSJ. Res N 0591-20/R.

El equipo de trabajo es variado, está constituido por personas que trabajan dentro del ámbito de la UNSJ (docentes/investigadores y alumnos de grado y postgrado), un profesor consultor del Instituto Tecnológico Buenos Aires (científico y analista de datos), y otras que se desempeñan en el gobierno provincial.

1. INTRODUCCIÓN

Hoy en día, la Analítica e Inteligencia de Negocios o Empresarial (ABI, por sus siglas en inglés de Analytics and Business Intelligence) constituye un factor estratégico para las organizaciones, ya que hace posible que los directivos tengan una visión clara y precisa tanto del estado actual como a futuro, y tomar decisiones en base a ella. Incluye un conjunto de aplicaciones, infraestructuras, herramientas

y prácticas que permiten el acceso y el análisis de la información con el propósito de mejorar y optimizar las decisiones y el rendimiento de una organización [1].

Los proyectos analíticos avanzados requieren afrontar retos como la obtención de los requisitos y su alineación a las estrategias empresariales para lograr auténticos beneficios en las organizaciones donde se desarrollan. Además, necesitan la colaboración de diferentes miembros del equipo de desarrollo, para cubrir las tres áreas principales: la del negocio, del diseño de la solución analítica y de la preparación de los datos que constituyen los repositorios o almacenes de datos (AD) [2], [3], [4]. En [5] se presenta un enfoque de ingeniería de requerimientos para proyectos ABI orientados al negocio que, entre otras cosas, incluye el análisis VMOST (por sus siglas en inglés de Vision, Mission, Objectives, Strategy, Tactics). Se propone un relevamiento sistemático que permita entender la visión, objetivos y procesos de una empresa o negocio desde la perspectiva de los stakeholders, y así alinear las soluciones analíticas a su plan estratégico.

Por otra parte, este tipo de proyectos posee un componente esencial, el repositorio o almacén de datos. Éstos contienen los datos fuente luego de haberlos pre procesados, consolidados y organizados para poder ser usados eficientemente por las herramientas de extracción de información/conocimiento. Por ello, su construcción constituye una etapa relevante dentro del desarrollo de esta clase de soluciones. Existen varios trabajos [6], [7], entre otros, que abordan esta temática, desde el diseño conceptual hasta el lógico y el físico. La implementación se concreta generalmente en una base de datos, y si bien los Sistemas de Gestión de Bases de Datos (SGBDs) Relacionales son los más utilizadas

actualmente [8] no sólo en contextos transaccionales sino también en los analíticos, es bastante conocido que los NoSQL están avanzando considerablemente. Prometen alta disponibilidad, eficiencia en el manejo de grandes volúmenes de datos, escalabilidad horizontal y flexibilidad de estructuras [9], aspectos significativos dentro de este tipo de proyectos. Consecuentemente, resulta de interés utilizar el paradigma de almacenamiento NoSQL como soporte de los AD [10], [11].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo está alineado al proyecto marco, el cual pretende establecer un conjunto de procesos que conduzcan al desarrollo de soluciones analíticas, donde los objetivos estratégicos del negocio sean contemplados en fases tempranas, de manera que se logren sistemas analíticos valiosos y oportunos para los tomadores de decisión de cualquier organización. Además, se propone desarrollar un sistema analítico que brinde apoyo a la toma de decisiones para la STyT.

La propuesta cuenta con antecedentes logrados en el tema en proyectos aprobados y subsidiados por CICITCA - UNSJ en los que el grupo ha trabajado. Entre las líneas de investigación y desarrollo abordadas se encuentran:

- Innovación en el área educativa.
- Captura de datos para generar información analítica en la pequeña empresa.
- Revisión sistemática de la literatura sobre el uso de inteligencia y analítica de negocios en el ámbito gubernamental.

- Análisis actual y simulaciones en el mundo financiero de inversión en criptomonedas.

A su vez, se ha trabajado en:

- La elaboración de un modelo ágil para generar Almacenes de Datos incrementales.
- La implementación de Almacenes de Datos en bases de datos relacionales e iniciado el estudio en no relacionales.

3. RESULTADOS

El grupo de investigación viene trabajando en esta temática específica desde hace algunos años y desde el 2020 más formalmente en proyectos de investigación, cuyos resultados fueron expuestos en los Workshops anteriores [12], [13]. En el transcurso del año 2022 los resultados más importantes fueron:

- “Aplicación de Inteligencia y Analítica de Negocios en diferentes contextos”. XXIV Workshop de Investigadores en Ciencias de la Computación – XXIV WICC 2022 (UChampagnat, 28 y 29 de Abril) [14].
- Concreción y firma de un Acta Complementaria entre la FCEFN de la UNSJ y La STyT del GPSJ, para llevar a cabo el proyecto “Solución de Inteligencia de Negocios a partir de datos provenientes de la STyT”. Res N° 113/2022-CD-FCEFN.
- Culminación de dos tesis de grado y la inminente finalización de una tesis de maestría.

4. FORMACIÓN DE RECURSOS HUMANOS

En el año 2022 el acento estuvo puesto en apoyar a los alumnos de grado y posgrado. Con muchos se lograron grandes avances, y con otros no tanto, dado que estaban simultáneamente cursando algunas materias. En consecuencia, obtuvieron sus títulos de Licenciado en Sistemas de Información los siguientes alumnos integrantes del proyecto:

- Martín Iván Manassero. Tesis: “Sistema Analítico y Gestor de Cartera de Inversiones en criptomonedas”. Calificación: Distinguido.
- María Laura Molina Frías. Tesis: “Aplicación de Inteligencia de Negocios para analizar aspirantes a las carreras de la FCEFN-UNSJ”. Caso de Estudio: Carreras del Departamento de Informática - años 2019, 2020, 2021. Calificación: Sobresaliente.

En cuanto a las tesis de posgrado:

- La Lic. Cristina Vera estará presentando en marzo el primer borrador de su tesis (Maestría en Informática). La temática abordada refiere al uso y comparación de gestores de bases de datos orientados a grafos y relacionales, tanto en contextos transaccionales como de análisis.
- El Lic. Fernando Emanuel Más, que se incorporó a mediados del año 2022, está cursando la Maestría en Inteligencia de Negocios y Explotación de Datos. Su tema de tesis es “Analítica de contribuyentes, en relación a su capacidad contributiva en el ámbito del estado provincial”.

Además, dentro del planteamiento del nuevo proyecto, se ha determinado sumar a las tareas de investigación y explotación de datos, a integrantes de la STyT, a fin de involucrarlos y capacitarlos en la temática.

5. BIBLIOGRAFÍA

- [1] “Definition of Analytics and Business Intelligence (ABI) - Gartner Information Technology Glossary.” [Online]. Available: <https://www.gartner.com/en/information-technology/glossary/business-intelligence-bi>. [Accessed: 25-Nov-2022].
- [2] N. Rosasco and J. Dehlinger, “A case study investigation of a lightweight, systematic elicitation approach for enterprise architecture requirements,” in *Communications in Computer and Information Science*, 2015, vol. 551, pp. 33–45.
- [3] S. Nalchigar and E. Yu, “Business-driven data analytics: A conceptual modeling framework,” *Data Knowl. Eng.*, vol. 117, pp. 359–372, Sep. 2018.
- [4] S. Nalchigar and E. Yu, “Designing Business Analytics Solutions. A Model-Driven Approach,” *Bus. Inf. Syst. Eng. 2018*, vol. 62, no. 1, pp. 61–75, Aug. 2018.
- [5] A. L. Cravero, S. E. Sepulveda, J. N. Mazón, and J. C. Trujillo, “Un enfoque de ingeniería de requerimientos basada en el alineamiento de almacenes de datos y la estrategia del negocio,” *Ingeniare. Rev. Chil. Ing.*, vol. 21, no. 3, pp. 314–327, Dec. 2013.
- [6] P. Ponniah, *Data Warehousing Fundamentals for it Professionals*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2010.
- [7] A. Vaisman and E. Zimányi, *Data Warehouse Systems - Design and Implementation*. Springer-Verlag Berlin Heidelberg, 2014.
- [8] “DB-Engines Ranking - popularity ranking of graph DBMS,” *DB-ENGINES*, 2022. [Online]. Available: <https://db-engines.com/en/ranking/graph+dbms>. [Accessed: 25-Nov-2022].
- [9] J. R. Lourenço, B. Cabral, P. Carreiro, M. Vieira, and J. Bernardino, “Choosing the right NoSQL database for the job: a quality attribute evaluation,” *J. Big Data*, vol. 2, no. 1, pp. 1–26, Dec. 2015.
- [10] E. Gallinucci, M. Golfarelli, and S. Rizzi, “Approximate OLAP of document-oriented databases: A variety-aware approach,” *Inf. Syst.*, vol. 85, pp. 114–130, Nov. 2019.
- [11] H. Akid and M. Ben Ayed, “Towards NoSQL graph data warehouse for big social data analysis,” in *Advances in Intelligent Systems and Computing*, 2017, vol. 557, pp. 965–973.
- [12] M. I. Lund *et al.*, “Inteligencia y analítica de negocios para la toma de decisiones en diferentes contextos,” in *XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja)*, 2021, no. August 2021, pp. 963–968.
- [13] M. Herrera *et al.*, “Análisis de calidad de vida de la población sanjuanina aplicando ciencia de datos,” in *XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja)*, 2021, no. August 2021, pp. 963–968.
- [14] S. I. Migani *et al.*, “Aplicación de inteligencia y analítica de negocios en diferentes contextos,” in *XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza)*, 2022, pp. 148–152.

Rediseño de la Plataforma Web MultiDB, Cuyo Propósito es el Aprendizaje de Lenguaje SQL, con Acceso a Distintos Motores de Base de Datos

Romero María Soledad, Damiano Luis Esteban, Maldonado Calixto, Giro Juan Matías, Guevara Andrea, Romero Rubén, Fernández Erika, Ríos Gastón Ignacio

Grupo de Investigación, Desarrollo y transferencia de Sistemas de Información
Departamento Ingeniería en Sistemas de Información

Facultad Regional Córdoba / Universidad Tecnológica Nacional

Maestro Marcelo López esq. Cruz Roja Argentina – Ciudad Universitaria - Córdoba
0351 - 4686385

{msromero, ldamiano, calixtomaldonado, mguevara, rromero, enfernandez}@frc.utn.edu.ar
{gastonriosc, matiasgiro}@gmail.com,

RESUMEN

El proyecto que se pone en consideración es un replanteo tecnológico de una aplicación que sirve de plataforma de aprendizaje del lenguaje de manipulación de bases de datos SQL, esta herramienta cuya denominación es MultiDB fue creada en un proyecto anterior entre los años 2015 y 2017, su denominación es MultiDb.

Su funcionalidad original es el poder realizar sólo consultas SQL (select) sobre distintos modelos de datos que se encuentran contruidos y poblados en distintos motores de base de datos relacionales. La modernización del producto permitirá asegurar el crecimiento de funcionalidades en beneficio del aprendizaje con participación activa del estudiante de la cátedra Base de Datos (antes Gestión de Datos) de la Carrera Ingeniería en Sistemas de Información. La concepción inicial de la herramienta ha sido facilitar y amenizar la práctica del alumno.

Los cambios serán por un lado de tipo tecnológico enfocados en actualizar los programas a los estándares más modernos y actuales, y por otro lado de tipo funcional, mejorando el comportamiento del producto, donde se agrega la posibilidad de utilizar otros comandos SQL, tales como: insert, update, delete, create.

La plataforma continuará siendo Web.

Seguirá teniendo un área administrativa para que los responsables de sostener los modelos de datos y las consultas SQL programadas puedan realizar cambios, como también una área de consumo público para que los estudiantes puedan acceder con facilidad a la utilización de la herramienta.

Palabras clave: SQL, Base de datos relacionales, Plataforma Web, Herramienta educativa.

CONTEXTO

El proyecto de investigación denominado “Aplicativo Web Para Interactuar Con Distintos Motores De Bases De Datos Relacionales” está homologado en la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, con el código de proyecto SIPPACO0008530, para un plazo de 24 meses, desde el 1 de enero de 2022. El proyecto funciona dentro del Grupo CIDS (Centro de Investigación, Desarrollo y transferencia de Sistemas de Información) , de la Universidad Tecnológica Nacional, Facultad Regional Córdoba.

El origen del proyecto está basado en la necesaria reingeniería de un proyecto anterior, MultiDb. Esta reingeniería se impone como una necesidad para poder disponer una programación más flexible y ágil al momento de incrementar prácticas interactivas para el estudiante, puesto

que el modelo de programación original, si bien da un resultado esperado para el proyecto no brinda la posibilidad de incorporar modificaciones en la forma y tiempo requeridos a fin de garantizar un aprendizaje dinámico.

El grupo de trabajo está conformado inicialmente por siete docentes investigadores de la carrera de Ingeniería en Sistemas de Información, de la Universidad Tecnológica Nacional - Facultad Regional Córdoba, con la colaboración de un graduado.

Los docentes se desempeñan en cátedras relacionadas a las siguientes temáticas: Programación, Base de Datos y Sistemas de Información.

El equipo de investigación tiene experiencia acabada en la administración de bases de datos, en capacitación y dictado de cursos, experiencia en desarrollo de sistemas de información con uso de bases de datos relacionales y pericia en tareas de arquitectura en empresas privadas. Pero es de destacar que este grupo ha participado en forma continua en distintos proyectos de investigación, tales como:

1. TecnoDB - Administrador de Base de Datos Relacional (2007) [1]
2. PROMETEO - Desarrollo de un método y una herramienta para el aprovechamiento de Metadatos de Base de Datos Relacionales (2010) [2]
3. Análisis y aplicación de metodologías para la generación de consultas complejas utilizando esquemas OLAP (2010) [3]
4. Generador Automático de Modelos de Datos Normalizados en Bases de Datos Relacionales (2014-2015) [4]
5. MultiDB. Plataforma Web para acceder a diferentes Sistemas de Gestión de Bases de Datos Relacionales (2016-2017) [5]
6. Buenas prácticas en el diseño de estructuras de datos en bases de datos relacionales (2018-2019) [6]

1. INTRODUCCIÓN

En la actualidad la plataforma es utilizada por los estudiantes de la cátedra de Base de Datos, del tercer nivel de la Carrera de Ingeniería en Sistemas de Información, en el proceso de aprendizaje. Mediante el feedback de los alumnos y docentes, frecuentemente se proponen mejoras sobre la plataforma. Ya que este proyecto es una evolución, vale la pena hacer una breve recapitulación: Nuestra plataforma no es la primera propuesta en el sentido de desear colaborar con tecnología en los procesos de enseñanza y de aprendizaje. Entonces referido al segundo interés, *“Desarrollos tecnológicos equivalentes o similares, con una fuerte inclinación hacia la enseñanza de origen externo”*, podemos mencionar los siguientes, que destacan sobre los encontrados:

1. Software comercial denominado Navicat, que permite administrar una base de datos a usuarios remotos. Actualmente promociona acceso a: MySQL, MariaDB, SQL Server, SQLite, Oracle y PostgreSQL. Este producto se podría utilizar para la enseñanza pero el objetivo es netamente comercial, por lo cual no es accesible si no se compra y está principalmente orientado a Administradores de Bases de Datos que acceden a las bases de datos de manera remota. Es un producto registrado por la empresa PremiumSoft. [7]

2. En ambientes educativos se encontró un software denominado RAT (Relational Algebra Translator), que traduce operaciones de Álgebra Relacional a SQL. El RAT fue desarrollado en el Laboratorio de Bases de Datos, de la Facultad de Ciencias Exactas y Naturales-Escuela de Informática, de la Universidad Nacional de Costa Rica. Implementa las operaciones originales del Álgebra Relacional y las de conjuntos. El objetivo central del RAT no es el lenguaje SQL, sino en cómo traducir desde un lenguaje abstracto, como el propuesto por el Álgebra Relacional, a un código ejecutable en SQL. Un aspecto importante es que el RAT demanda la instalación de un software cliente en el equipo y la instalación es dependiente del

Sistema Operativo, ya que requiere: “Windows 2000 SP4, Windows Xp, Windows Vista o Windows 7. Para poder ejecutarse la aplicación el sistema debe tener instalado el Framework 2.0 o superior el cual se puede descargar del sitio Web en el caso que la computadora no lo tenga ya instalado”. Otra característica es que el usuario debe poseer instalado en su equipo el ODBC, como estándar establecido por Microsoft, para permitir tener comunicación con diversas bases de datos. Estos aspectos siguen haciendo que los usuarios no se alejen de aspectos técnicos para la ejercitación y dependen de la instalación de sistema operativo o software cliente en la PC. El software cliente puede descargarse e instalarse libremente [8].

3. SQL Online es un sitio que provee infraestructura como servicio pero carece de seguimiento en el aprendizaje, tampoco es gradual el nivel de complejidad [9].

4. SQL Fiddle es un sitio que permite escribir y ejecutar código SQL y cuenta con opción para elegir entre varios motores de bases de datos. Carece de lineamientos pedagógicos y devoluciones constructivas al aprendizaje [10].

También se realizaron consultas sobre software que permita mejorar el aprendizaje del lenguaje cuando participamos en diversos eventos o grupos de trabajo, como por ejemplo: Como docentes en las reuniones de cátedras Inter facultades de las 12 regionales de UTN, que actualmente poseen la Carrera de Ingeniería en Sistemas de Información. En la Red de Carreras de Ingeniería en Informática/Sistemas de Información de CONFEDI (RIISIC) no tenemos información de algo que se asemeje.

Nuestra participación en congresos, workshops, jornadas y seminarios relacionados con la temática, así como la interconsulta con docentes de otras universidades que contemplan el dictado de la temática, nos permitió poder esclarecer la ausencia de una herramienta educativa con estas características. Por ello es de sumo interés aportar una herramienta novedosa, con características no habituales y que facilite el proceso de aprendizaje, donde también fue preciso cambiar el enfoque para adaptarlo al

aprendizaje basado en resultados de aprendizaje (modelo de aprendizaje por competencias) donde el foco está en cómo combinar integralmente el saber y saber hacer en pos de lograr los resultados esperados. En este sentido los cambios en la mirada metodológica, mediación pedagógica, estrategias, instrumentos y criterios de evaluación se convierten en requisitos de la herramienta MultiDB, para que el estudiante logre escribir correctamente sentencias SQL durante el cursado de la asignatura.

La interconsulta con profesionales en bases de datos, de empresas que desarrollan software en el medio local, permitió verificar que es un problema que necesita tratarse, dado el impacto del uso de las bases de datos en cualquier proyecto de software.

La temática por abordar está íntimamente relacionada con las bases de datos relacionales, que si bien surgieron hace ya 45 años y que inicialmente sólo se las utilizaba para persistir datos desde las aplicaciones de las empresas [11], hoy son utilizadas para la toma de decisiones a partir de la exploración de las bases de datos, como lo es con la Minería de Datos definida como “el descubrimiento eficiente de información valiosa, no obvia de una gran colección de datos” [12] [13].

Se trata de un problema relevante dado su impacto directo en la cantidad de horas hombre destinadas a la resolución de temas inherentes al manejo de datos, en el entorno de bases de datos relacionales.

Más aún si se considera que es una tarea componente de la mayor parte de los proyectos de software y que de ella dependen otras tareas fundamentales, como por ejemplo el desarrollo y el testing.

Todo desarrollador de aplicaciones debe conocer SQL, para embeber el código necesario en sus programas y así acceder/manipular los datos que solicita un usuario final [14] [15].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el Grupo CIDS se integran distintas líneas de investigación en la Ingeniería en Sistemas de Información y de Software, albergando proyectos que trabajan sobre calidad, métricas, mantenimiento, trazabilidad, requerimientos, bases de datos, seguridad, relacionados con las áreas de conocimiento: Sistemas de Información, Gestión Ingenieril y Programación.

El proyecto trabaja sobre conocimiento y desarrollos de programación sobre las temáticas: base de datos relacionales, programación Web, Dockerizar (empaquetar una aplicación).

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera evolucionar los resultados ya obtenidos en el producto anterior, que fueron los siguientes:

- Acceder a un intérprete SQL, para ejecutar sentencias SQL, accediendo a una estructura de datos ya conocida en la clase y poblada de datos.
- Ejecutar sentencias tipificadas, para que los estudiantes elijan según sus necesidades de ejecución.
- Observar el resultado de la ejecución, pudiendo listar, imprimir y copiar el conjunto de filas resultante, cuando fuera una consulta.
- Recibir mensajes de la plataforma cuando haya un error en la sintaxis.
- Validar el resultado a través de mensajes sobre la cantidad de filas y columnas devueltas por una consulta.
- Recibir mensajes sobre si el resultado es el esperado o qué error se detecta contra el resultado esperado.

Los resultados esperados, como consecuencia de las mejoras tecnológicas y funcionales realizadas al producto, son:

- Adaptar la herramienta al modelo de aprendizaje basado en competencias.
- Lograr mejores posibilidades de rediseño del producto en tiempos aceptables.
- Incorporar la posibilidad de ejecutar mayor cantidad de comandos SQL dentro de la aplicación.
- Construir mejores metodologías que permitan una medición de mayor calidad de los eventos que se producen en la manipulación de la herramienta por parte de los estudiantes.
- Establecer procesos que permitan establecer métricas sobre los resultados de los comandos ejecutados por los estudiantes.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está constituido por: a) una directora de proyecto, b) un codirector, c) 4 docentes investigadores de apoyo, e) un graduado.

Los 4 docentes junto a la directora y codirector que integran el equipo de investigación comparten su trabajo en la cátedra de Base de Datos, de tercer nivel de la Carrera de Ingeniería en Sistemas de Información, donde se trata la temática de Bases de Datos y específicamente el Lenguaje SQL. El ingeniero Calixto Maldonado se encuentra Doctorando.

BIBLIOGRAFÍA

- [1] Gastañaga, I., Maldonado, C., Martínez Spessot, C. I., & Hintermeister, E. (2006). TecnoDB una Base de Datos Relacional y Prometeo un método de aprovechamiento de Metadatos y Generador de Consultas. In VIII WICC.
- [2] Marciszack, M., Maldonado, C., Martínez Spessot, C. I., Muñoz, R., Navarro, A., Peretti, J. P., & Roggero, L. (2009). Prometeo: una herramienta para el aprovechamiento de

metadatos de base de datos relacionales. In XI WICC.

[3] Vaca, P. A., & Abrutsky, M. (2009). Herramienta para consultas complejas orientada a usuarios finales. In XI WICC.

[4] Paz Menvielle, M. A., Cuevas, J. C., Damiano, L. E., Muñoz, R., & Quinteros, S. (2013, June). Generador automático de modelos de datos normalizados en bases de datos relacionales. In XV WICC.

[5] Muñoz Roberto Miguel, Maldonado Calixto, Damiano Luis Esteban, Romero María Soledad, Cuevas Juan Carlos, Quinteros Sergio Ramón, Guevara Andrea, Carrasco Agustín. MultiDB, Plataforma Web para acceder a diferentes Sistemas de Gestión de Bases de Datos Relacionales. XVIII WICC (2016)

[6] Muñoz Roberto Miguel, Maldonado Calixto, Damiano Luis Esteban, Romero María Soledad, Bueno Matías, Quinteros Sergio Ramón, Guevara Andrea, Peretti Juan Pablo, Carrasco Agustín, Urbano Barbara, Arguello Santiago. Buenas Prácticas En El Diseño De Estructuras De Datos En Bases De Datos Relacionales. XX Workshop de Investigadores en Ciencias de la Computación (2018)

[7] <http://www.navicat.com/es/>. Último acceso: 16 de julio de 2021, a las 12:30 hs.

[8] <http://www.slinfo.una.ac.cr> Último acceso: 16 de julio de 2021, a las 17:45 hs.

[9] <https://sqliteonline.com/> Último acceso: 10 de Julio de 2021, a las 8:00 hs.

[10] <https://sqlfiddle.com/> Último acceso: 10 de Julio de 2021, a las 8:00 hs.

[11] Date, C.J. Introducción a los Sistemas de Bases de Datos. Volumen 1. Séptima Edición. México, 2001. Edit. Addison Wesley Longman, Inc. ISBN: 0-201-38590-2.

[12] Bigus, J. P. Data Mining with Neural Networks: Solving Business Problems from

Application Development to Decision Support. 1996. McGraw-Hill. ISBN:0-07-005779-6

[13] Riquelme, J., Ruiz, R y Gilbert, K. Minería de Datos: Conceptos y Tendencias. Artículo publicado por Revista Iberoamericana de Inteligencia Artificial. ISSN: 1137-3601. <http://www.redalyc.org/articulo.oa?id=92502902>. Último acceso: 27 de febrero de 2023, a las 11:35 hs.

[14] Introducción a los Sistemas de Bases de Datos. Volumen 1. Séptima Edición. México, 2001. Edit. Addison Wesley Longman, Inc. ISBN: 0-201-38590-2

Sistemas inteligentes en el uso de aplicaciones de bioinformática y sistemas embebidos

W. Hasperué¹ , C. Estrebou¹ , G. Camele^{1,3} , E. Rucci¹ , F. Ronchetti¹ , D. Castillo^{2,4} ,
G. Reyes Zambrano^{2,5} , L. Lanzarini¹ , A. Fernandez Bariviera⁶ 

¹ Instituto de Investigación en Informática LIDI*, Facultad de Informática, UNLP, La Plata, Argentina

² Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina

³ Becario postgrado UNLP.

⁴ Facultad de Ciencias de la Educación, Universidad Tecnológica Indoamérica, Ambato, Ecuador

⁵ Facultad de Ciencias Físicas y Matemáticas, Universidad de Guayaquil, Guayaquil, Ecuador

⁶ Dpto. de Economía, Universitat Rovira i Virgili, Reus, España

* Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)

{whasperue, cesarest, gcamele, erucci, fronchetti, laural}@lidi.info.unlp.edu.ar,
{david.castillos, gary.reyesz}@info.unlp.edu.ar, aurelio.fernandez@urv.net

CONTEXTO

Esta presentación corresponde a las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto “Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data” perteneciente al Programa de Incentivos (2018-2022).

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de Big Data y Minería de Datos utilizando técnicas de Aprendizaje Automático. Los sistemas desarrollados se aplican particularmente al procesamiento de grandes volúmenes de información y al procesamiento de flujo de datos.

Las investigaciones correspondientes al procesamiento de datos masivos están enfocadas en el estudio y desarrollo de técnicas de selección de características, donde el foco está puesto en la reducción de los tiempos de cómputo. La optimización puede realizarse tanto en la mejora de la ejecución en un entorno distribuido, como en la propuesta de técnicas metaheurísticas que permitan obtener un subconjunto óptimo de atributos.

Por otro lado, y desde el punto de vista de la salud, se está trabajando con investigadores del

CENEXA (CONICET-UNLP-CIC) en la obtención de modelos de predicción de diabetes y prediabetes.

Las investigaciones relacionadas con el análisis de flujos de datos se centran en la construcción de modelos dinámicos descriptivos. En particular el énfasis está puesto en la resolución de dos problemas de sumo interés en distintas áreas: el análisis de trayectorias GPS a fin de identificar congestiones en el tránsito y la identificación temprana de patrones de movimiento en pacientes con Alzheimer.

Palabras clave: Big Data, Minería de Datos, Diabetes, Análisis de flujos de datos, Reducción de características, Tiny ML, GPS.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos tipos de estrategias adaptativas. Los resultados obtenidos han sido aplicados en la solución de problemas de distintas áreas. A continuación, se detallan las investigaciones realizadas durante el último año.

1.1. BIG DATA

Selección de características

El objetivo de los algoritmos de selección de características es el de reducir las entradas a un tamaño apropiado para su procesamiento y

análisis. La selección de características implica la elección de ciertos atributos, tal que, con ese subconjunto, las “propiedades naturales” de los datos no sean alteradas.

Actualmente, en el III LIDI se están realizando tareas de investigación que incluyen el desarrollo de algoritmos de selección de características que puedan ser utilizados en bases de datos con información genética. Un objetivo de la medicina genómica es identificar un grupo de genes, cuyo patrón de expresión se encuentre asociado a un fenotipo en particular: concepto conocido como *gene signature* (biomarcador diagnóstico, pronóstico o predictivo de una patología en estudio).

Cuando el volumen de información a procesar crece, la ejecución de los algoritmos de selección de atributos convencionales incrementa notablemente su tiempo de procesamiento. En la actualidad se cuenta con herramientas que, al distribuir el cómputo entre diferentes nodos que conforman un cluster de computadoras, hacen posible el procesamiento de grandes volúmenes de datos de una manera eficiente. En este aspecto Apache Spark es uno de los frameworks más utilizados. En particular, su librería Spark ML contiene la implementación de muchos algoritmos de machine learning.

En [1] y en [2] se llevó a cabo una comparación entre cuatro algoritmos de clasificación implementados en MLlib: Random Forest, Support Vector Machine, Naïve Bayes y MultiLayer Perceptrón. Se analizaron cuatro métricas de poder pronóstico de los modelos junto con los tiempos de ejecución requeridos. Los experimentos están enfocados a comparar las métricas de los cuatro algoritmos estudiados en función del número de atributos seleccionado de una base de datos.

Identificación de Personas con Riesgo de Diabetes y Prediabetes

La Diabetes Tipo 2 (DT2) es una enfermedad crónica caracterizada por una disminución precoz y progresiva de la masa y de la función de las células beta del páncreas. Debido a su creciente prevalencia en combinación con su elevado costo de atención, constituye un serio

problema de salud pública, por lo que se han realizado grandes esfuerzos por desarrollar estrategias efectivas para su prevención y tratamiento, así como para evitar sus complicaciones crónicas. En ese sentido, resulta importante reconocer que las consecuencias negativas de esta enfermedad comienzan en una etapa previa conocida como prediabetes, la cual implica un riesgo elevado de desarrollar DT2 en los siguientes años.

El desarrollo de la DT2 es un proceso lento y progresivo condicionado por factores genéticos, ambientales y de comportamiento. Aunque no existe una cura definitiva para esta enfermedad, varios estudios han demostrado que se puede prevenir o demorar su aparición en personas con prediabetes a través de la adopción de un estilo de vida saludable y/o asociado con la ingesta de diversos fármacos. En Argentina, existe la iniciativa Programa PPDBA desarrollado por el CENEXA (CONICET-UNLP-CIC) [4].

La detección de DT2 y prediabetes representa un verdadero desafío para la medicina debido a la ausencia de síntomas patogenómicos y/o la falta de conocimiento de los factores de riesgo asociados. Los modelos existentes para predicción de diabetes y prediabetes no necesariamente aplican a la población argentina [3] y, hasta donde llega nuestro conocimiento, no existe modelo ni herramienta similar disponible en nuestro medio que permita identificar personas con alta probabilidad de tener estas enfermedades. Brevemente, esta línea propone desarrollar y validar modelos predictivos de diabetes y prediabetes específicos para la población argentina utilizando técnicas de Aprendizaje Automático. Se cuenta con acceso a la base de datos del PPDBA y apoyo del equipo médico de CENEXA. La concreción de esta línea representaría un avance en el conocimiento y un instrumento útil para los sistemas de salud de Argentina.

1.2. ANALISIS DE FLUJOS DE DATOS

Patrones de tránsito vehicular

El volumen de tráfico vehicular de las grandes ciudades se ha incrementado en los últimos años originando problemas de movilidad; por

ello el análisis de los datos del flujo vehicular son de suma importancia. Los Sistemas Inteligentes de transporte realizan el monitoreo y control vehicular recolectando trayectorias GPS. Las técnicas de agrupamiento permiten identificar patrones sobre el flujo vehicular.

En este sentido se ha desarrollado en el III-LIDI una metodología capaz de identificar, de manera dinámica, las velocidades más representativas del flujo vehicular en un período de tiempo. Para ello, el flujo de datos GPS entrada es representado utilizando un reticulado de celdas correspondientes a pequeñas zonas geográficas. Luego, las características de las celdas son agrupadas dinámicamente. La repetición de estos pasos en forma periódica utilizando períodos de tiempo cortos fueron representados en un mapa interactivo facilitando de esta manera la interpretación del desplazamiento de los vehículos a distintas velocidades. La validación del método propuesto fue realizada con dos conjuntos de datos de Roma y Guayaquil y los resultados obtenidos fueron satisfactorios. Los resultados de estas investigaciones se publicaron en [5].

Actualmente se está completando este procesamiento con la incorporación de índices de congestión con el objetivo de identificar los cambios en el flujo vehicular, analizando las zonas geográficas en un período dado de tiempo, y estimar si se trata de una congestión o no.

Detección de poses en pacientes con Alzheimer

Según la OMS, la demencia es una de las principales causas de discapacidad y dependencia entre las personas mayores. En todo el mundo, más de 55 millones de personas viven con demencia y se calcula que esta cifra aumentará a 78 millones para 2030. Demencia es un término general para varias enfermedades que generalmente son de naturaleza crónica, que resultan en deterioros cognitivos e interfieren con la capacidad para realizar las actividades de la vida diaria. La enfermedad de Alzheimer es la forma más común de demencia contribuyendo al 60-70% de los casos. Está comprobado que un diagnóstico precoz mejora la calidad de vida del paciente y del familiar, aumenta o mantiene su autonomía personal y

mantiene sus capacidades cognitivas. Las personas que padecen Alzheimer suelen manifestar poco interés en las cosas o en el medio que lo rodea y presentan cambios conductuales variables desarrollando movimientos corporales relacionados de forma directa con su estado anímico. Esta línea de investigación se centra en la identificación automática de los movimientos corporales a partir de imágenes y videos.

En [6] se comenzó trabajando con el análisis de la marcha de pacientes con Alzheimer. Luego en [7] se trabajó sobre las poses: "De pie" y "Sentado" en pacientes con Alzheimer a partir de imágenes obtenidas de centros de atención al adulto mayor del cantón Ambato, Ecuador. Se trabajó con una población de 45 personas de ambos sexos diagnosticadas con Alzheimer con edades entre 75 y 89 años. Estas poses identificadas fueron utilizadas posteriormente en un análisis exploratorio relacionado con las categorías de deambulación, nervioso, deprimido, desorientado o aburrido arrojando resultados satisfactorios.

En [8] se buscó clasificar automáticamente el estado de ánimo de un paciente con Alzheimer en una de las siguientes categorías: deambulante, nervioso, deprimido, desorientado, aburrido o normal a partir de videos obtenidos en hogares de ancianos del cantón Ambato, Ecuador. Se trabajó con una población de 39 personas de ambos sexos diagnosticadas con Alzheimer y cuyas edades oscilaban entre los 75 y 89 años. Los métodos utilizados fueron detección de poses, extracción de características y clasificación de poses. Se utilizaron redes neuronales, el clasificador walk y la métrica de distancia Levenshtein. Como resultado del análisis de cada video se generó automáticamente una secuencia de estados de ánimo interpretable por el experto humano. Se pudo afirmar satisfactoriamente que el software de visión artificial facilita el reconocimiento de los estados de ánimo de los enfermos de Alzheimer durante los cambios de pose a lo largo del tiempo ayudando al profesional médico en su diagnóstico.

1.3. TINYML

La combinación de las áreas del aprendizaje automático y de los sistemas embebidos da lugar al campo de estudio conocido como TinyML. En esta línea de investigación, desde 2021 se lleva adelante un proyecto de investigación [9, 10] con el objetivo de explorar y adaptar técnicas y modelos de aprendizaje automático para dispositivos con importantes limitaciones de hardware como los microcontroladores (MCU).

Uno de los principales aportes del proyecto es el desarrollo de EmbedIA, un framework que transforma modelos de redes neuronales Tensorflow/Keras (TF/Keras) en código compatible con C/C++/Arduino. Esta característica permite realizar inferencia de modelos en cualquier MCU sin requerimientos específicos de hardware, a diferencia de otros frameworks o bibliotecas que demandan dispositivos de 32 bits (mayoritariamente basados en arquitectura ARM) o soporte para instrucciones DSP o SIMD. Además, se han generado y adaptado modelos TF/Keras para su ejecución en diversos MCUs. De la comparación de éstos con otros frameworks [10,11,12], se concluyó que, en general, los modelos ocupan menos memoria y se ejecutan más rápido en EmbedIA.

Entre las tareas que se están desarrollando actualmente, se encuentra la integración a EmbedIA de nuevas funcionalidades. Por un lado, se está trabajando en la integración de soporte para modelos de aprendizaje automático Scikit-Learn basados en árboles de decisión. Por otro lado, se está trabajando en la compatibilidad con modelos de redes neuronales en formatos como tflite, onnx y torch, además del ya soportado por TF/Keras. También se están desarrollando y evaluando modelos convolucionales reducidos para detección de fruta en mal estado, detección de personas e identificación de palabras clave en audio para placas de desarrollo ESP32-CAM, Raspberry Pi Pico y ARM Stm32f411.

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Medición de performance de algoritmos de machine learning en entornos distribuidos.

- Desarrollo y mantenimiento de una herramienta para el análisis de progenie, basada en Spark.
- Identificación de Personas con Riesgo de Diabetes y Prediabetes
- Preprocesamiento de trayectorias vehiculares. Técnicas de segmentación.
- Agrupamiento dinámico de flujos de datos.
- Reconocimiento de poses humanas en imágenes y en videos
- Desarrollo de una BBDD de poses de movilidad.
- Estudio de técnicas de compresión de modelos para microcontroladores.
- Análisis de bibliotecas y frameworks de aprendizaje automático para microcontroladores.
- Desarrollo de un Framework de código abierto que transforma modelos desarrollados en Tensorflow/Keras y Scikit-Learn para ejecutarlos en microcontroladores.

3. RESULTADOS OBTENIDOS

- Medición de tiempos de ejecución de algoritmos de ML en un entorno Spark.
- Modelos de predicción de diabetes y prediabetes
- Diseño e implementación de un nuevo método de detección de rangos de velocidad agrupamiento de trayectorias GPS aplicable a la predicción de congestiones vehiculares.
- Desarrollo de una metodología capaz de clasificar automáticamente el estado de ánimo de un paciente con Alzheimer.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 4 profesores doctores con dedicación exclusiva, un profesor con dedicación exclusiva, 3 tesis de Doctorado en Cs. Informáticas (1 con beca de postgrado de la UNLP) y un profesor extranjero.

Dentro de los temas involucrados en esta línea de investigación, en los últimos 3 años se han finalizado 4 tesis de doctorado, 3 tesis de especialista, 5 tesinas de grado de Licenciatura y 7 prácticas profesionales supervisadas.

Actualmente se están desarrollando 5 tesis de doctorado, 3 tesis de maestría y 5 prácticas profesionales supervisadas. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Camele, G.; Hasperué, W.; Ronchetti, F.; Quiroga, F.; A comparative study of the performance of four classification algorithms from the Apache Spark ML library. CACIC. Salta. 2021.
- [2] Camele, G.; Hasperué, W.; Ronchetti, F.; Quiroga, F.M. Statistical analysis of the performance of four Apache Spark ML algorithms. *Journal of Computer Science and Technology* 22 (2), e14-e14, 2022.
- [3] Choudhury A., Gupta D. A Survey on Medical Diagnosis of Diabetes Using Machine Learning Techniques. In: *Recent Developments in Machine Learning and Data Analytics. Advances in Intelligent Systems and Computing*, vol 740. Springer, Singapore. 2019.
- [4] Gagliardino J. J., Etchegoyen G., Bourgeois M., Fantuzzi G., García S., González L., Elgart J. F., Ré M., Ricart A., Ricart J. P., Spinedi E., “Prevención primaria de diabetes tipo 2 en argentina: estudio piloto en la provincia de buenos aires,” *Revista Argentina de Endocrinología y Metabolismo*, vol. 53, no. 4, pp. 135 – 141, 2016.
- [5] Data stream processing method for clustering of trajectories. Reyes G., Lanzarini L., Estrebou C., Fernández-Bariviera A. *Communications in Computer and Information Science. Serie CCIS. Springer International Publishing. ISSN 1865-0929. Pags.151-163. 2022.*
- [6] Using Kinect to Detect Gait Movement in Alzheimer Patients. Castillo-Salazar D., Lanzarini L., Guevara C., Alvarado H.G. In: *Trends and Applications in Information Systems and Technologies. WorldCIST 2021. Advances in Intelligent Systems and Computing*, vol 1365, pp.14-28. ISBN 978-3-030-72656-0. Springer. 2021.
- [7] Castillo Salazar, D.R., Lanzarini, L., Gómez Alvarado, H.F., Cabrera López, J. The Detection, Extraction, and Classification of Human Pose in Alzheimer's Patients. In: Troiano, L., Vaccaro, A., Kesswani, N., Díaz Rodríguez, I., Brigui, I. (eds) *Progresses in Artificial Intelligence & Robotics: Algorithms & Applications. Lecture Notes in Networks and Systems*, vol 441. Springer, Cham. ISBN 978-3-030-98531-8. 2022.
- [8] Castillo Salazar, D., Lanzarini, L., Alvarado, H., Varela-Aldás, J. Artificial vision system to detect the mood of an Alzheimer's patient. (2022). In: Tareq Ahram, Jay Kalra and Waldemar Karwowski (eds) *Artificial Intelligence and Social Computing (AHFE 2022) International Conference. AHFE Open Access*, vol 28. USA. ISSN 2771-0718.
- [9] Estrebou C., Feming M., Saavedra M. D., Adra F. MbedML: A Machine Learning Project for Embedded Systems. IX Jornadas de Cloud Computing, Big Data & Emerging Topics. ISBN: 978-950-34-2016-4 Pp. 25-28. Facultad de Informática. UNLP. Junio 2021.
- [10] Estrebou, C., Saavedra, M. D., Adra, F., Fleming, M. TinyML for Small Microcontrollers. X Jornadas de Cloud Computing, Big Data & Emerging Topics. ISBN 978-950-34-2126-0. Pp 42-46. Facultad de Informática. UNLP. Mayo 2022.
- [11] Estrebou C., Feming M., Saavedra M. D., Adra F. A Neural Network Framework for Small Microcontrollers. XXVII Congreso Argentino de Ciencias de la Computación. ISBN: 978 -987-633-574-4. Pp. 51-60. Univ. Nacional de Salta. Octubre 2021.
- [12] Estrebou C., Feming M., Saavedra M. D., Adra F., De Giusti, A. E. Lightweight Convolutional Neural Networks Framework for Really Small TinyML Devices. Second International Conference on Smart Technologies, Systems and Applications. SmartTech-IC 2021. Quito, Ecuador. Diciembre 2021.

Datos no Estructurados: Indexación, Búsquedas y Aplicaciones

Darío Ruano, Paola Azar, Andrea Maldonado , Norma Herrera

Laboratorio de Investigación y Desarrollo en Bases de Datos

Universidad Nacional de San Luis, San Luis, Argentina

dmruano@email.unsl.edu.ar, epazar@unsl.edu.ar, andreamaldonadoma@gmail.com, nherrera@unsl.edu.ar

Andrés Pascal

Departamento Ingeniería en Sistemas de Información

FRCU, Universidad Tecnológica Nacional, Entre Ríos, Argentina

andrespascal22@gmail.com

Resumen

Los avances en las tecnologías de la información y la comunicación permitieron que las bases de datos crezcan de manera exponencial en tamaño y se diversifiquen en contenido. Un gran porcentaje de los datos disponibles son no estructurados, por lo que los mismos no son asimilables con los modelos clásicos de bases de datos, como el modelo relacional. Por esta razón, se necesita contar con nuevos modelos de datos y nuevas herramientas para el manejo de los mismos. En este proyecto nuestro interés está centrado al estudio de modelos para bases de datos no estructurados y técnicas de indexación y búsqueda asociados a los mismos.

Contexto

El presente trabajo se desarrolla en el ámbito de la línea Técnicas de Indexación para Datos no Estructurados del Proyecto Tecnologías Avanzadas de Bases de Datos (22/F814), cuyo objetivo es realizar investigación sobre manejo y recuperación eficiente de información no

tradicional que implique el manejo de datos no estructurados. Este proyecto forma parte de Laboratorio de Investigación y Desarrollo en Bases de Datos (LaBDa) de la Universidad Nacional de San Luis.

1 Introducción

El crecimiento exponencial de las fuentes de información disponibles, ha provocado que las bases de datos crezcan en volumen y en tipo de datos almacenados. En la actualidad, gran parte de la información disponible involucra el uso de datos no estructurados (sonidos, imágenes, video, huellas digitales, etc) que no admiten la representación clásica de bases de datos relacionales, donde la información se organiza como nuplas (filas) en relaciones (tablas). Los procesos de búsqueda se enfrentan a nuevos desafíos relacionado al rendimiento y al tipo de respuesta esperada.

Es en este contexto donde surgieron nuevos modelos de bases de datos capaces de cubrir las necesidades de las aplicaciones que manejan datos no estructurados [3]. Entre esos

modelos encontramos el modelo de *bases de datos de texto*, el *modelo de espacios métricos* y el *modelo métrico-temporal*.

En el caso de **bases de datos de texto**, la base de datos consiste de una gran colección de texto sobre la que se requiere resolver consultas de manera eficiente. Esta colección de texto se conceptualiza como un única y gran secuencia de caracteres T , que se encuentra dividida en varios archivos. Este texto T puede representar no sólo lenguaje natural, sino también música, códigos de programas, secuencias de ADN, secuencias de proteínas, etc. Una de las búsquedas más comunes en bases de datos de texto es la *búsqueda de un patrón*: el usuario ingresa un string P (*patrón de búsqueda*) y el sistema retorna todas las posiciones del texto T donde P ocurre. Resolver la búsqueda de un patrón de manera eficiente en grandes colecciones de texto requerirá la construcción de un índice.

El modelo de **espacios métricos** [4, 9, 12] permite modelar y manejar datos no estructurados sobre los que se quieren realizar búsquedas por similitud. El espacio está formado por un conjunto de objetos $\mathcal{X}$ y una función de distancia d definida entre ellos que mide cuan similares son. La base de datos es cualquier subconjunto finito $\mathcal{U} \subseteq \mathcal{X}$. Una de las consultas más comunes en este modelo de bases de datos es la *búsqueda por rango*. En esta búsqueda dado un elemento $q \in \mathcal{X}$, al que llamaremos *query*, y un radio de tolerancia r , la búsqueda por rango consiste en recuperar los objetos de la base de datos cuya distancia a q no sea mayor que r .

El modelo **métrico-temporal** [10, 2] fue pensado para manejar objetos no estructurados con tiempos de vigencia asociados, en los que existe necesidad de consultar por similitud y por tiempo en forma simultánea. Un *espacio métrico-temporal* es un par (U, d) , donde $U = O \times N \times N$, y la función d es de la forma $d : O \times O \rightarrow R^+$. Cada elemento $u \in U$ es una triupla (obj, t_i, t_f) , donde obj es un objeto (por ejemplo, una imagen, sonido,

cadena, etc) y $[t_i, t_f]$ es el intervalo de vigencia de obj . La función de distancia d , que mide la similitud entre dos objetos, cumple con las propiedades de una métrica (positividad, simetría y desigualdad triangular). Una *consulta métrico-temporal* implica buscar todos los objetos o de la base de datos que estén a una distancia a lo más r de q , y cuyo tiempo asociado t se solape con el tiempo de la consulta.

En todos estos modelos, para resolver eficientemente las consultas planteadas, se necesitan índices que sean específicos a cada tipo de dato y tipo de consulta planteada.

En este proyecto nos dedicamos al estudio, diseño y optimización de algoritmos de indexación y búsquedas sobre datos no estructurados y exploramos dominios de aplicación de las técnicas diseñadas. La aplicación de los métodos y técnicas diseñadas en casos reales de estudio nos permiten realizar transferencia tecnológica al medio y medir el desempeño de nuestras propuestas en problemas reales de la comunidad. En la actualidad nos encontramos trabajando sobre problemas relacionados a la medicina forense.

2 Líneas de Investigación

Describimos a continuación las principales líneas en las que nos encontramos trabajando actualmente.

2.1 Indexación en Memoria Secundaria

El proceso de paginación de un índice consiste en dividir el mismo en partes, cada una de las cuales se aloja en una página de disco. Luego el proceso de búsqueda consiste en ir cargando en memoria principal una parte, realizar la búsqueda en memoria principal sobre esa parte, para luego cargar la siguiente y proseguir la búsqueda.

Cuando un índice se maneja en disco, el costo de búsqueda queda determinado por la cantidad de accesos a disco realizadas [11]. Aun así, es importante no descuidar las operaciones que se hacen en memoria principal a fin de lograr un funcionamiento eficiente del índice.

En la actualidad nos encontramos trabajando en la implementación de dos índices para memoria secundaria: el *Trie de Sufijos (TS)* [8, 6] para bases de datos de texto y el *Historical-FHQT(H-FHQT)* [7] para bases de datos métrico-temporales.

En ambos casos hemos diseñado una técnica de paginado basándonos en el algoritmo de paginado para árboles binarios presentado en [5]. Esta técnica consiste en particionar el árbol en componentes conexas, denominadas *partes*, cada una de las cuales debe tener un tamaño que no supere la capacidad de una página de disco. El algoritmo procede en forma bottom-up tratando de condensar en una única parte un nodo con uno o más subárboles que dependen de él. En este proceso de particionado las decisiones se toman en base a la profundidad de cada nodo involucrado, donde la profundidad indica la cantidad de accesos a disco que deberá realizar el proceso de búsqueda para llegar desde esa parte a una hoja del árbol. En el caso del H-FHQT, se tiene en cuenta además de la profundidad, otros parámetros propios del diseño del índice para decidir qué partes condensar.

En este sentido, se ha avanzado en la implementación del TS y del H-FHQT en memoria secundaria, encontrándonos en la etapa de testing de las mismas.

2.2 Espacios Métricos y Medicina Forense

En medicina legal y forense existen varios temas de interés que necesitan manejar datos no tradicionales para resolver problemas de manera eficiente. Uno de esos problemas es la

identificación de cadáveres NN en el contexto de búsqueda de personas desaparecidas.

Dentro de los individuos que ingresan a los distintos Institutos de Medicina Forense del país, existen casos que no poseen las condiciones adecuadas para su identificación inmediata (indocumentados, en avanzado estado de descomposición, restos óseos, etc.) o sin posibilidad de identificación (fragmentos muy pequeños, restos carbonizados, etc.). Frente a esto, las instituciones deben investigar no sólo para determinar qué fue lo que sucedió (causa de la muerte) y cuándo sucedió (data de la muerte), sino también para poder dar con la identidad del cuerpo.

Claramente la identificación de cadáveres está directamente relacionada con la búsqueda de personas desaparecidas. En Argentina, no existe un sistema único de procesamiento para esta problemática. Cada provincia tiene su gobierno, su sistema forense y sus protocolos. Esto dificulta el proceso de identificación de cadáveres: si una persona desaparece en Chaco y aparece un cadáver similar en Chubut, no hay una forma rápida y correcta de relacionarlos. El Sistema Federal de Búsqueda de Personas Desaparecidas y Extraviadas (SIFEBU) es un intento de crear esta base de datos unificada pero sin tener automatizado el proceso de búsqueda.

En esta línea abordamos la aplicación de la teoría de Espacios Métricos para la identificación de cadáveres en el contexto de búsqueda de personas desaparecidas. El objetivo es desarrollar un sistema que permita mantener una base de datos, modelizada con un espacio métrico, con información sobre cadáveres no identificados para posteriormente realizar búsquedas de personas desaparecidas.

Para ello, por cada cadáver mantenemos un vector con los datos de las características físicas del mismo. Al momento de realizar una búsqueda, se deberá ingresar las características físicas de la persona buscada y con esa información generar el vector característico de la persona buscada. A partir de este vector se re-

aliza una búsqueda por similitud sobre la base de datos de cadáveres, obteniendo como resultado una lista de cadáveres con características similares a la persona buscada, rankeados según el grado de similitud. Hasta el momento, el trabajo desarrollado es el siguiente:

Generación de Vectores Característicos

Para la elaboración del sistema, como primer paso hubo que definir un core de datos que seas adecuado a la problemática de identificación. Usar pocos datos podría provocar que las búsquedas que se realicen sean de muy baja selectividad y usar demasiados puede provocar que se descarten elementos de la base de datos que sean de interés. En este sentido, ya hemos definido un primer core de datos sobre el cual trabajar: *color de ojos, color de pelo, color de piel, existencia de tatuajes y lugar de los mismos, cicatrices y/o marcas* (lunares por ejemplo), *si existen amputaciones y en qué lugar del cuerpo, la contextura física (atlético, atrófico, etc.)* y finalmente la existencia de *agenesias*. Para cada dato, se transforman los valores de su dominio en números que reflejen el grado de similitud entre los valores considerados y se genera el vector correspondiente. Claramente no todos los datos tienen el mismo grado de importancia, por ejemplo el color de pelo es menos importante que el color de ojos porque una persona puede cambiarse el color de pelo pero no el de ojos. Esto hay que tenerlo en cuenta para establecer para cada dato un peso que corresponda con el grado de importancia del mismo.

Función de Distancia Otro punto importante es la función de distancia a utilizar en las búsquedas. En una primera etapa usaremos la función coseno [1]. Si q es la query (vector de la persona buscada) y d_j es el j -ésimo vector de la base de datos, entonces el grado de similitud entre los vectores $\vec{d}_j$ y $\vec{q}$ se calcula como el coseno del ángulo formado entre ambos vectores.

Indexación y Búsquedas Con respecto al algoritmo de indexación comenzaremos usando algoritmos basados en pivotes. Cuando la base de datos se cargue con datos reales, se podrá analizar la dimensionalidad del espacio métrico sobre el cual se está trabajando y de ser necesario se cambiará el algoritmo de indexación. Con respecto a las búsquedas, utilizaremos las búsquedas de los k vecinos más cercanos, porque es la que más se adecúa a este problema. Esto permitirá al usuario del sistema decidir cuántos elementos desea recuperar en una primera instancia y luego, de ser necesario, podrá ampliar la búsqueda; por ejemplo: puede pedir los 10 cadáveres mas parecidos a la persona buscada y posteriormente puede ampliar la búsqueda pidiendo los 10 siguientes.

En función de los resultados que se obtengan con la primer versión del sistema se realizarán, de ser necesario, cambios que pueden implicar: aumentar o disminuir los datos del core, cambiar la función de distancia y/o cambiar el algoritmo de indexación.

Una extensión inmediata de esto, es mantener una bases de datos con los vectores característicos de las personas desaparecidas para que cada vez que ingresa un cadaver NN a un Instituto Forense del país, se pueda realizar la correspondiente búsqueda sobre la base de datos de personas desaparecida.

3 Resultados Esperados

Como resultados de los trabajos que estamos realizando, esperamos obtener índices eficientes en memoria secundaria para las bases de datos planteadas. Todo el trabajo realizado nos permite adquirir la experiencia suficiente como para abocarnos al diseño de otros índices que sean competitivos en este ámbito.

En cuanto a la transferencia tecnológica, el resultado esperado es un sistema web de

identificación de cadáver en el contexto de búsqueda de personas desaparecidas.

4 Formación Recursos Humanos

Dentro de esta línea se forman docentes y alumnos de la Universidad Nacional de San Luis y de la Universidad Tecnológica Nacional (FRCU). Actualmente hay en desarrollo 2 Tesis de Maestría y un Trabajo Final de la Licenciatura, todos de la Universidad Nacional de San Luis.

Bibliografía

- [1] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley, 1999.
- [2] A. De Battista, A. Pascal, N. Herrera, and G. Gutierrez. Metric-temporal access methods. *Journal of Computer Science & Technology*, 10(2):54–60, 2010.
- [3] H.M. Blanken, A.P. de Vries, H.E. Blok, and L. Feng. *Multimedia Retrieval. Data-Centric Systems and Applications*. Springer Berlin Heidelberg, 2007.
- [4] E. Chávez, G. Navarro, R. Baeza-Yates, and J.L. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [5] D. Clark and I. Munro. Efficient suffix tree on secondary storage. In *Proc. 7th ACM-SIAM Symposium on Discrete Algorithms*, pages 383–391, 1996.
- [6] J. Cornejo, D. Ruano, and N. Herrera. Una mejora en tiempo del trie de sufijos. In *Congreso Argentino de Ciencias de la Computación*, Rio Cuarto, Córdoba, Argentina, 2019.
- [7] A. De Battista, A. Pascal, G. Gutierrez, and N. Herrera. Un nuevo índice métrico-temporal: el historical-fhqt. In *Actas del XIII Congreso Argentino de Ciencias de la Computación*, Corrientes, Argentina, 2007.
- [8] G. H. Gonnet, R. Baeza-Yates, and T. Snider. *New indices for text: PAT trees and PAT arrays*. Prentice Hall, New Jersey, 1992.
- [9] Filip Nalepa, Michal Batko, and Pavel Zezula. Enhancing similarity search throughput by dynamic query reordering. In Sven Hartmann and Hui Ma, editors, *Database and Expert Systems Applications*, pages 185–200, Cham, 2016. Springer International Publishing.
- [10] A. Pascal, A. De Battista, G. Gutierrez, and N. Herrera. Métodos de acceso para bases de datos métrico - temporales. In *Actas del XV Congreso Argentino de Ciencias de la Computación*, pages 1061–1070, Jujuy, Argentina, 2009.
- [11] Jeffrey Scott Vitter. *Algorithms and Data Structures for External Memory*. Publishers Inc, 2006.
- [12] Pavel Zezula. Similarity management of data: The DISA experience. In Massimo Mecella, Giuseppe Amato, and Claudio Gennaro, editors, *Proceedings of the 27th Italian Symposium on Advanced Database Systems, Castiglione della Pescaia (Grosseto), Italy, June 16-19, 2019*, volume 2400 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2019.

Administración y Recuperación de Información en Bases de Datos Masivas

Luis Britos, Fernando Kasián, Verónica Ludueña, Diego Olivares,
Marcela Printista, Nora Reyes, Patricia Roggero, Gabriela Romano, Pablo Samat
LIDIC, Dpto. de Informática, Fac. de Cs. Físico Matemáticas y Naturales, Universidad Nacional de San Luis
{fkasian, vlud, mprinti, nreyes, proggero}@unsl.edu.ar,
{ldoorz, olivarestello.diego, gabrielasindaromano, samatpablo}@gmail.com

Karina Figueroa

Universidad Michoacana de San Nicolás de Hidalgo, México
karina@fisimat.umich.mx

Claudia Deco

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario
deco@fceia.unr.edu.ar

Resumen

Como se ha vuelto evidente durante los últimos tiempos, debido al uso masivo de dispositivos electrónicos, la cantidad de datos multimedia disponibles crece exponencialmente. Estos datos por su naturaleza ya son digitales, aunque sean de diversos tipos, y se acumulan en grandes repositorios de tipos de datos no estructurados. En particular, estos repositorios son difíciles de administrar bajo el modelo relacional de base de datos, por lo que para su administración y recuperación se hacen necesarias herramientas particulares, porque deben modelizarse sobre otros tipos de base de datos que no son las tradicionales.

Más aún, si consideramos el volumen de datos no estructurados que deben administrarse, se hace aún más evidente la necesidad de considerar nuevos modelos y herramientas para su administración y su indexación, considerando las operaciones de interés, porque es necesario resolverlas de manera eficiente, considerando la existencia de la jerarquía de memorias y la posibilidad de hacer uso de computación de alto desempeño.

Palabras Claves: bases de datos masivas, computación de alto desempeño, recuperación de información.

1. Contexto

La línea de investigación “Recuperación de Datos e Información” se enmarca dentro del Proyecto Consolidado 3-03-2018 de la Universidad Nacional de San Luis (UNSL - Res. CS 126/18) y en el Programa de Incentivos (Código 22/F834): “Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos”, se desarrolla en el Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) de la UNSL, finalizado en 2022. Actualmente se ha realizado una nueva presentación.

Esta línea cuenta con 4 docentes-investigadores y una estudiante de maestría y tiene como objetivo el desarrollo de herramientas eficientes que permitan no sólo administrar grandes bases de datos que almacenan datos no estructurados, sino también poder recuperar datos e información sobre ellas. Con este propósito se analizan nuevas técnicas que provean una buena interacción con el usuario y se desarrollan nuevas estructuras de datos capaces de indexar un gran volumen de datos no estructurados, que permitan resolver eficientemente los tipos de consultas de interés sobre estos tipos de datos (texto, secuencias de ADN, huellas digitales, audio, vídeo, etc.), que sean escalables y sean adecuadas para memorias jerárquicas. Para mayor eficiencia se pueden aplicar técnicas de computación de alto desempeño (HPC).

2. Introducción

Debido al uso masivo de internet, a la gran disponibilidad de dispositivos electrónicos y a la evolución de las tecnologías de información y comunicación, se están generando una cantidad masiva de datos digitales. Además, por provenir de fuentes muy diversas, los tipos de datos generados son también muy variados. En este contexto de problemas de “big data” (porque aparecen sus dos características importantes, que son: el volumen de datos y variedad de los mismos) se hace imposible restringir las consultas a búsquedas sobre datos estructurados tradicionales, porque obligaría a representar una visión parcial del problema. Entonces, para no perder información relevante, es necesario redefinir las

técnicas de procesamiento, análisis y obtención de información útil y formular nuevas metodologías.

Un enfoque significativo para sistemas de recuperación de información es la *búsqueda basada en contenidos*, donde se usa el dato no estructurado en sí mismo para describir lo que se está buscando. En este escenario los tipos de búsqueda más generales y aplicables son las búsquedas por similitud y para resolverlas eficientemente, es necesario utilizar *métodos de acceso* o *índices métricos* [3].

Por lo tanto, al considerar grandes cantidades de datos no estructurados y ser necesario responder consultas de recuperación de información, se pueden utilizar índices métricos para lograr eficiencia en las respuestas. Por otra parte, cuando se consideran aplicaciones reales, estos índices además de eficientes deben permitir actualizaciones (altas y bajas de elementos) y ser escalables; es decir, que su desempeño no se vea afectado muy negativamente a medida que crece el volumen de datos a indexar.

En este escenario, los *espacios métricos*, compuestos por un *universo* $\mathcal{U}$ de objetos y una *función de distancia* $d : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}^+$, resultan adecuados para modelizar las búsquedas por similitud. La función d mide la disimilitud entre dos objetos y cumple con las propiedades de reflexividad, positividad estricta, simetría y desigualdad triangular.

Entre las búsquedas por similitud, las más utilizadas son: la *búsqueda por rango* y la *búsqueda de los k vecinos más cercanos*. Una de las principales ventajas de este modelo es que, además de brindar un marco formal, es independiente del dominio de la aplicación. Al considerar una base de datos $S \subseteq \mathcal{U}$ tal que $|\mathcal{S}| = n$, cualquier consulta se resuelve de manera trivial calculando n evaluaciones de distancia. El cálculo de distancia sobre algunos tipos de datos no estructurados puede ser muy costoso (por ej.: comparación de huellas digitales). Por lo tanto, un objetivo importante es aprovechar los índices para ahorrar cálculos de distancia durante las búsquedas. Así, en la mayoría de las aplicaciones sobre datos masivos no es posible prescindir de un índice y resolver una consulta comparando el objeto de interés contra toda la base de datos (búsqueda exhaustiva). Para construir un índice se preprocesa la base de datos S , si se conocen de antemano sus elementos, sino se debe crear incrementalmente.

Sin embargo, en muchos casos es probable que la base de datos, el índice, o ambos, deban alojarse en memoria secundaria, porque no caben en memoria principal. Como se sabe, esto puede incrementar

mucho los costos de las búsquedas, porque las operaciones de E/S en memoria secundaria suelen ser muy costosas. Por lo tanto, para lograr eficiencia en estos casos, no basta con realizar pocos cálculos de d sino que también se debe minimizar el número de operaciones de E/S. Una solución adecuada en esta situación debe considerar el nivel de la jerarquía de memorias sobre la que se trabaja, utilizar técnicas de computación de alto desempeño (HPC) y, en algunos casos, admitir respuestas no exactas [5].

En esta línea, se busca desarrollar nuevas herramientas que soporten la interacción con el usuario, diseñar índices que permitan la manipulación eficiente de grandes volúmenes de datos no estructurados y faciliten la realización de diferentes tipos de consultas, para lograr recuperación de información eficiente sobre conjuntos masivos de datos. De esta manera, se espera contribuir al desarrollo de soluciones a problemas de big data para aplicaciones reales.

3. Líneas de Investigación

Dado que en esta línea se pretende contribuir a distintos aspectos de los sistemas de recuperación de información sobre grandes volúmenes de datos no estructurados, se ha considerado inicialmente optimizar índices existentes, diseñar nuevos índices, resolver distintas clases de consultas, y lograr eficiencia y escalabilidad para grandes volúmenes de datos; para luego incorporarlos en un sistema administrador para este tipo de bases de datos.

Métodos de Acceso Métricos

Como se ha mencionado, el modelo para estas bases de datos masivas que contienen datos no estructurados son las *bases de datos métricas*, en las cuales interesa responder a consultas por similitud. Para acelerar la respuesta a estas consultas, minimizando los costos involucrados en el proceso, se utilizan métodos de acceso métricos (MAMs) [3].

En general, los MAMs sacan partido de las distancias almacenadas en el índice y aprovechan que la distancia satisface la propiedad de desigualdad triangular, para ahorrar cálculos de distancia y también tiempo. Habitualmente, los MAMs almacenan algunas distancias precalculadas o referencias obtenidas al comparar los elementos de la base de datos respecto de un conjunto de objetos distinguidos llamados *centros*, *pivotes* o *permutantes*.

Durante las búsquedas es posible estimar la distancia entre cualquier objeto de consulta q y los elementos de la base de datos, aprovechando las distancias o referencias almacenadas y la desigualdad

triangular. Así, es posible determinar qué objetos no son relevantes y evitar cálculos de d .

Cuando se diseñan MAMs, se deben considerar varios aspectos: dinamismo (que soporten altas y bajas), en qué nivel de la jerarquía de memorias deben almacenarse, si es posible aplicar técnicas de HPC para mejorar los tiempos de respuesta, si son de respuesta exacta o de respuesta aproximada y la dimensionalidad del espacio métrico considerado.

Por el volumen de los datos con los que se trabaja, se debe considerar que posiblemente los índices no puedan almacenarse en memoria principal. En general, se considera directamente su almacenamiento en memoria secundaria y sólo se dispone de memoria principal para mantener información muy útil que permita determinar que partes del índice deben recuperarse desde el disco. Por ello, para lograr eficiencia en las búsquedas, se debe minimizar el número de cálculos de distancia y de operaciones de E/S.

Así, una parte de esta línea se dedica a diseñar MAMs, preferentemente dinámicos y adaptados para memoria secundaria, cuyo desempeño en las búsquedas sea adecuado para las aplicaciones de interés. Entre los índices que se han desarrollado se encuentran algunos completamente dinámicos y especialmente diseñados para trabajar sobre grandes volúmenes de datos, basados en la *Lista de Clusters*¹ por (LC) [2]. Uno de ellos, el *Conjunto Dinámico de Clusters* (DSC) [9], que tiene una buena ocupación de página y operaciones eficientes tanto en cálculos de distancia como en operaciones de E/S, logra un buen desempeño en espacios de alta dimensión. El DSC mantiene los clusters en memoria secundaria y organiza los centros de clusters en un DSAT [9] en memoria principal. La información en el DSAT permite bajar los costos en cálculos de distancia, y mantener bajos los costos de E/S.

Sin embargo, cada inserción en un DSC debe localizar el cluster donde se insertará el nuevo elemento. Luego, se lee el cluster desde el disco y se verifica si el nuevo objeto podría ser un mejor centro para el cluster; haciendo, si es posible, el cluster más compacto. Finalmente, se graba nuevamente el cluster en el disco y se actualiza el DSAT de centros si el nuevo objeto se transformó en el nuevo centro. Esta operación, además de los cálculos de distancia necesarios, realiza pocas operaciones de E/S en el disco. Para mejorar más respecto de la E/S se está analizando una variante que en lugar de insertar los ele-

mentos en el índice a medida que lleguen, demore la inserción de cada nuevo objeto a un cluster hasta tener varios elementos y determinar un mejor agrupamiento de los mismos. Así, se mejoran los costos de búsqueda, al lograr clusters aún más compactos y se asegura una total ocupación de la página del disco, achicando el tamaño del archivo y reduciendo los tiempos de acceso. Esta opción, podría reducir significativamente el costo de construcción, por amortizar el costo de la escritura de un cluster en disco, luego de varias inserciones.

Así, se ha diseñado el *Buffered On Line Dynamic Set of Clusters* (BOLDSC), una variante del DSC, que agrega un espacio en memoria para mantener los objetos que son insertados (bolsa), además de un índice auxiliar (de pivotes) sobre estos elementos. Como pivotes se eligen elementos de la bolsa, y cuando ésta se llena, se selecciona el pivote que necesita el menor radio de cobertura para encerrar la cantidad de elementos que caben en un cluster y todos éstos se sacan de la bolsa y se graban en un cluster completo. La grabación de clusters llenos mejora la cantidad de E/S, pero puede provocar que objetos cercanos, que se insertan en diferentes momentos, queden en clusters diferentes, ocasionando degradación en las búsquedas. Por ello, se está estudiando una variante del DSC que graba clusters con una cantidad de objetos que está en función de un factor de carga ρ (casi llenos). Como los clusters no estarán llenos, habrá espacio libre en cada cluster. Entonces, al insertar un nuevo objeto se verifica si hay clusters que lo puedan incorporar y si hay se elige al más cercano a él. Se verifica si el nuevo elemento sería un mejor centro para el cluster y si es así, se actualiza el DSAT de centros. Si no es posible incorporarlo a un cluster (todos llenos), va a la bolsa y se actualiza el índice de pivotes. Así, los clusters se adaptan para mejorar los costos de búsqueda. Esta variante está en etapa inicial de implementación.

El DSC basa su buen desempeño en la calidad de la partición generada por los centros de sus clusters. Así, otro aspecto a considerar es mejorar la poda en estos agrupamientos. Cada cluster en DSC está determinado por su centro y su radio. Sin embargo, puede haber “huecos” dentro de ellos. Por esta razón, es posible seleccionar un conjunto de “pivotes” globales que permitan caracterizar las zonas de los clusters dentro de las cuales existan efectivamente elementos [8], manteniendo para cada cluster información sobre la mínima y la máxima distancia a la que cada pivote encuentra elementos de

¹Usaremos la palabra *cluster* en inglés, en lugar de agrupamiento en español, ya que es de uso habitual en computación.

dicho cluster. Esta información identifica los “huecos” del cluster para que, durante las consultas, se pueda descartar un cluster si la bola de la consulta no posee intersección efectiva con la zona del cluster que realmente contiene elementos de interés. En las búsquedas se recupera desde memoria secundaria cada cluster que intersecta la bola de consulta con centro q y radio r . Al identificar las zonas del cluster que no contienen elementos, si la consulta intersecta al cluster en su zona “hueca” se lo puede descartar y ahorrar cálculos de distancia y lecturas a disco. Esta variante está en etapa avanzada de implementación.

Por otra parte, el *Árbol de Aproximación Espacial Distal (DiSAT)* [1], es un índice estático y muy competitivo respecto del estado del arte. Se ha demostrado que la razón principal de su buen desempeño en las búsquedas es por la partición que induce el árbol sobre la base de datos. Por ello, se está analizando utilizar la información de la partición que induce un *DiSAT* sobre la base de datos para obtener una *representación binaria compacta basada en esquemas* y realizar búsquedas por similitud aproximadas, en las que se admite intercambiar precisión en la respuesta por velocidad. Esta representación de secuencia de bits identifica la partición a la que pertenece el objeto y reduce el espacio de almacenamiento del índice. Durante las búsquedas, se puede obtener la representación binaria de la consulta q y compararla eficientemente con las representaciones de los objetos de la base de datos para determinar cuales parecen ser similares a q (sus representaciones deben ser similares entre sí). Luego, la consulta q se debe comparar usando d para determinar si esos elementos realmente son relevantes. Así, aunque podrían no recuperarse todos los objetos relevantes, por la calidad de la partición del *DiSAT*, se espera disminuir el costo de las búsquedas y obtener buena calidad de respuesta. Este nuevo índice está en etapa de diseño.

Otra manera de acelerar las búsquedas es paralelizar el índice usando técnicas de HPC. Se ha implementado un *DSC* paralelo, que no sólo baje la cantidad de cálculos de distancia y operaciones de E/S necesarias para responder a una consulta, sino resuelva más consultas simultáneas, aprovechando al máximo las operaciones de E/S que se realicen. Esta implementación considera un índice distribuido local. Cada nodo mantiene su propio índice local para los datos que almacena. La implementación se compone de: un broker, una cola de mensajería y los nodos del cluster conteniendo los índices locales. El broker se encarga de enviar los objetos al momento

de crear el árbol y al momento de hacer las búsquedas, dejando distintos mensajes en los tópicos correspondientes. Cada nodo lee los objetos del tópico de creación y agrega el objeto al índice local. También lee los objetos de búsqueda del tópico, busca en el índice local y deja los resultados en ese tópico.

Para la comunicación entre el broker y los nodos se utiliza Apache Kafka. Los datos se almacenan en Kafka en forma de tópicos. Al momento de la creación, el broker va a dejar un objeto en el tópico que corresponda, el nodo correspondiente leerá de ese tópico e insertará el elemento en su índice local. Así, se construyen los diferentes índices locales para cada nodo. En el caso de las consultas, el broker será el encargado de leer las respuestas del tópico, unir los resultados, y devolver la respuesta que corresponde a cada consulta. Para identificar qué respuesta corresponde a qué consulta, al realizarla se deja junto con el objeto un mensaje en el tópico con un identificador único. Así, la respuesta irá al tópico que corresponda y el broker será capaz de unir todos los resultados que correspondan al mismo identificador de tópico. Finalmente, para considerar aspectos de portabilidad, aislamiento, escalabilidad y disponibilidad y rendimiento de los nodos de cluster cada componente se ubica en un contenedor (“docker”) y corren en un cluster de Kubernetes. Actualmente, se está optimizando y evaluando la implementación.

DBMS para Bases de Datos Multimedia

Como ya se mencionó, las operaciones más comunes sobre bases de datos multimedia son las búsquedas por rango o de k -vecinos más cercanos. Sin embargo, no son las únicas operaciones a considerar en un sistema administrador para bases de datos multimedia [11]. En este caso, se debe incorporar también la operación de *join* por similitud [4].

Las variantes del *join* por similitud entre dos bases de datos A y B , con $A, B \subseteq \mathcal{U}$, dependen del criterio de similitud Φ utilizado. Cualquiera de ellas devuelve es un conjunto de pares (x, y) , $x \in A$ e $y \in B$, tales que se satisface Φ entre x e y . Algunas de las variantes más conocidas son: el *join* por rango, el *join* de k -vecinos más cercanos y el *join* de k pares de vecinos más cercanos [10]. Formalmente, dadas $A, B \subseteq \mathcal{U}$, el *join por similitud* entre A y B es $(A \bowtie_{\Phi} B) = \{(x, y) / x \in A \wedge y \in B \wedge \Phi(x, y)\}$. Si $A = B$, la operación se denomina *auto-join* [4]. Un DBMS para estas bases de datos, con datos métricos, debe resolver eficientemente las consultas.

PostgreSQL es un DBMS que permite realizar

consultas por similitud sobre algunos atributos. Por ejemplo, si se han creado índices *KNN-GiST* sobre texto, se podrían resolver búsquedas de *k*-vecinos más cercanos, pero no son aplicables sobre todo tipo de datos métricos. Además, estos índices brindan plantillas sólo para crear *árboles balanceados*, que no son adecuados para espacios métricos generales. Por ello, es importante incorporar en un DBMS la capacidad de manejar distintos tipos de datos no estructurados y las operaciones de interés.

Como las consultas de join son muy costosas, para acelerar los tiempos de respuesta se pueden disminuir los tamaños de los conjuntos obtenidos diversificando las respuestas [12]. El resultado es así un conjunto más pequeño de pares, con respuestas útiles y diversificadas, y en menor tiempo. Otra manera de acelerar la respuesta es dando una respuesta aproximada a la consulta. Por ello, se ha propuesto un algoritmo simple y eficiente para responder consultas de auto-join por similitud de *k* vecinos más cercanos aproximados, con una razonable precisión en la respuesta [4]. Por lo tanto, se espera lograr que *PostgreSQL* se transforme en un DBMS más aplicable sobre datos multimedia masivos.

4. Resultados

Recientemente, además de una estrategia para obtener buenas listas de candidatos para las búsquedas por similitud usando permutaciones en un índice invertido [5], se ha publicado una manera de acortar dichas listas de candidatos para acelerar las búsquedas [6], considerando una manera acortar las permutaciones y de evaluar las similitudes ellas [7].

Por otro lado, se está evaluando experimentalmente la versión paralela del índice *DSC* y se espera publicar estos resultados en el corto plazo. Además, se encuentran en etapa de publicación los resultados obtenidos con *BOLDSC* y en etapa de diseño e implementación las otras mejoras al *DSC*.

La extensión de *PostgreSQL* está en la etapa final para ponerla disponible y publicarla.

5. Formación de Recursos

Se están realizando tesis de Maestría: (1) “Estructuras Eficientes sobre Datos Masivos para Búsquedas en Espacios (UNSL)”, (2) “Sistema Administrador para Bases de Datos Métricas” (UNSL), (3) “Índices Métricos – Optimización del DSC usando Cortes de Regiones” (UNSJ), (4) “Optimización del BOLDSC por la Mejora de la Densidad y Solapamiento de los Clusters”(UNSJ) y (5) “Representa-

ción Basada en Esquemas para Búsquedas por Similitud Eficientes” (UNSL).

Referencias

- [1] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal sat. *Inf. Systems*, 59:15 – 47, 2016.
- [2] E. Chávez and G. Navarro. A compact space decomposition for effective metric indexing. *Pattern Recognition Letters*, 26(9):1363–1376, 2005.
- [3] L. Chen, Y. Gao, X. Song, Z. Li, X. Miao, and C. Jensen. Indexing metric spaces for exact similarity search. *ACM Comput. Surv.*, 55(6), 2022.
- [4] S. Ferrada, B. Bustos, and N. Reyes. An efficient algorithm for approximated self-similarity joins in metric spaces. *Inf. Systems*, 91:101510, 2020.
- [5] K. Figueroa, N. Reyes, and A. Camarena-Ibarrola. Candidate list obtained from metric inverted index for similarity searching. *Adv. in Comp. Intelligence*, 29–38, 2020. Springer.
- [6] K. Figueroa, A. Camarena-Ibarrola, and N. Reyes. Shortening the candidate list for similarity searching using inverted index. *Pattern Recognition*, 89–97, 2021. Springer.
- [7] K. Figueroa, A. Camarena-Ibarrola, N. Reyes, R. Paredes, and B. Hernández Martínez. A hybrid approach to boost the permutation index for similarity searching. In *Actas del XXVIII CACIC 2022*, 458–467, 2022.
- [8] J. Lokoč, J. Moško, P. Čech, and T. Skopal. On indexing metric spaces using cut-regions. *Inf. Systems*, 43:1–19, 2014.
- [9] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Inf. Systems*, 59:48 – 78, 2016.
- [10] R. Paredes and N. Reyes. Solving similarity joins and range queries in metric spaces with the list of twin clusters. *JDA*, 7:18–35, 2009.
- [11] C. Rong, C. Lin, Y. N. Silva, J. Wang, W. Lu, and X. Du. Fast and scalable distributed set similarity joins for big data analytics. In *2017 IEEE 33rd International Conference on Data Engineering*, pages 1059–1070, 2017.
- [12] L. Santos, L. Carvalho, W. Oliveira, A. Traina, and C. Jr. Traina. Diversity in similarity joins. *Similarity Search and Applications*, volume 9371 of *LNCS*, 42–53. Springer, 2015.

Diversas Aplicaciones de Bases de Datos Espaciales y Espacio Temporales

Edilma Olinda Gagliardi, María Gisela Dorzán,

Pablo Rafael Palmero y María Teresa Taranilla

Departamento de Informática

Facultad de Ciencias Físico Matemáticas y Naturales

Universidad Nacional de San Luis, Argentina

{oli, mgdorzan, prpalmero, tarani }@unsl.edu.ar

RESUMEN

En el proyecto *Tecnologías Avanzadas de Bases de Datos* de la Universidad Nacional de San Luis, se enmarca la línea de investigación *Bases de Datos Espaciales y Espacio Temporales*. Esta línea de investigación utiliza técnicas, métodos y herramientas tanto para la investigación de base como para el desarrollo de aplicaciones que contribuyen a la resolución de problemas en diversos campos de aplicación de las bases de datos espaciales y espacio-temporales.

Palabras clave: Bases de Datos, Bases de Datos Espaciales y Espacio Temporales, Geometría Computacional, AgroTIC, Gestión Integral de Incendios Forestales. Trazabilidad.

CONTEXTO

La línea de investigación *Bases de Datos Espaciales y Espacio Temporales* se encuadra en el Proyecto *Tecnologías Avanzadas de Bases de Datos* (PROICO N°03-2218) incluido en el Programa de Incentivos y en el contexto del Laboratorio de Investigación y Desarrollo en Base de Datos (LaBDa) de la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis.

El proyecto está integrado por tres líneas de investigación enfocadas en el estudio de bases de datos no convencionales y al desarrollo de nuevos modelos para gestionar y recuperar información alojada en repositorios de datos no estructurados.

Los modelos de bases de datos espaciales y espacio-temporales se emplean en situaciones donde se requiere almacenar y consultar información histórica y actual de ubicaciones referenciadas espacialmente. También se utilizan para observar los cambios en la forma de los objetos de estudio en distintos escenarios a lo largo del tiempo.

La línea de investigación se enfoca en el diseño y desarrollo de herramientas para gestionar y extraer información de sistemas de bases de datos. En este sentido, se plantea el diseño de aplicaciones que permitan modelar y administrar datos. Además, se busca obtener información para realizar análisis complejos y descubrir conocimiento que apoye en la toma de decisiones en un sistema.

El trabajo de investigación se lleva a cabo en colaboración con investigadores de proyectos afines en la Universidad Nacional de San Luis, y mediante acuerdos entre organizaciones nacionales y provinciales de San Luis.

1. INTRODUCCIÓN

Las Tecnologías de la Información y la Comunicación (TIC) han experimentado un importante avance y desarrollo en las últimas décadas. Estas tecnologías, que incluyen computadoras, internet, dispositivos móviles, software y redes, han tenido un gran impacto en la mayoría de las actividades humanas transformando la forma en que las personas trabajan, interactúan y se comunican en todo el mundo. Las TIC han propiciado el desarrollo de nuevas tecnologías y herramientas, las cuales han mejorado significativamente la eficiencia y productividad en diversos ámbitos. Estas mejoras han aportado beneficios concretos en situaciones del mundo real, optimizando aspectos como la prevención, gestión, calidad y economía, entre otros.

En ámbitos variados se observan experiencias que promueven la innovación en nuevas formas de comunicación y en prácticas que buscan optimizar tareas vinculadas a la gestión de diferentes temáticas.

En este contexto, los integrantes de la línea formaron parte del proyecto *Campo Conectado*, que surgió a partir de un convenio entre el Instituto Nacional de Tecnología Agropecuaria (INTA-Estación Experimental Agropecuaria San Luis) y la Universidad Nacional de San Luis. El objetivo principal del proyecto fue fomentar un espacio de investigación, intercambio y desarrollo, para estudiar el potencial y alcance de las TIC en el sector de la producción agropecuaria [1] [2].

Los Programas de Investigación Institucional de la Universidad Nacional de San Luis (Res. C.S 33/20-UNSL) tienen como objetivo promover el desarrollo de la investigación en áreas estratégicas de importancia para la universidad y la sociedad en general. Con este fin, se busca fomentar el trabajo colaborativo e interdisciplinario para abordar los desafíos actuales y futuros. En el marco de los programas mencionados, en el área estratégica de Medio Ambiente, se origina el proyecto "*Gestión Integral de Incendios Forestales*". Los miembros de la línea forman parte de este proyecto [11].

La ocurrencia frecuente de incendios forestales tiene gran impacto negativo en los ecosistemas naturales, al interrumpir los ciclos de vida de las especies vegetales y animales, lo que puede llevar a la desaparición de especies nativas, favorecer la aparición de especies invasoras y degradar el suelo en términos físico-químicos y biológicos. Es importante considerar el impacto económico que un incendio forestal puede tener en los recursos intangibles del bosque, como su valor para el ocio, el valor paisajístico y la fijación de carbono. Además, es necesario considerar que la recuperación del paisaje después del incendio puede ser difícil en varios ecosistemas únicos que tienen una menor capacidad de resiliencia. El proyecto *Gestión Integral de Incendios Forestales* tiene como objetivo abordar los problemas y desafíos asociados con los incendios forestales desde una perspectiva interdisciplinaria. Se propone un plan de estudio que involucra la investigación básica y aplicada, con el fin de diseñar herramientas y/o prototipos que permitan una toma de decisiones más informada y un análisis de impacto más preciso en la gestión de incendios forestales.

En junio de 2021, la niña Guadalupe Lucero desapareció en la ciudad de San Luis, su caso conmovió y movilizó a todo el país. A partir de este hecho, se planteó la idea de llevar a cabo el diseño de un sistema de información que logre aprovechar los recursos tecnológicos disponibles en la provincia para obtener la trazabilidad de los vehículos que se encontraban en una zona en un determinado lapso de tiempo, obteniendo así una herramienta utilizable en causas como desaparición de personas, tráfico, robos, entre otras.

En los dominios de aplicación mencionados (Campo Conectado, Incendios Forestales y Trazabilidad de Vehículos) se plantean diversos problemas donde surge la necesidad de representar y manipular tipos de datos tradicionales y tipos de datos complejos. Se requiere almacenar y consultar información actual e histórica, así como objetos de estudio referenciados espacialmente y su evolución en diferentes escenarios, a lo largo del tiempo.

En este contexto, los modelos de bases de datos espaciales permiten almacenar la información de la ubicación espacial de los objetos, sus características y las relaciones espaciales que existen entre ellos. Los modelos de bases de datos temporales incorporan el aspecto temporal para organizar la información y almacenan datos históricos junto con los datos actuales. Los modelos de bases de datos espacio-temporales combinan características de ambas, lo que permite analizar la evolución de los objetos y eventos en el espacio a lo largo del tiempo [4] [9] [12].

En estas áreas de trabajo se abordan problemas que surgen en diferentes campos de aplicación de las bases de datos espaciales y espacio-temporales. Se utilizan diversas técnicas y herramientas, incluyendo el diseño de índices espacio-temporales que sean apropiados en los diferentes escenarios y se plantea la aplicación de técnicas de geometría computacional para resolver los problemas involucrados [3].

2. LÍNEA DE INVESTIGACIÓN

En la línea se estudian diferentes dominios de aplicación de bases de datos espaciales y espacio-temporales, con uso de técnicas de indexación y diversas herramientas de apoyo en la resolución de problemas.

Entre los tópicos de estudio de la línea de investigación se destaca el diseño y aplicación de índices espaciales y espacio-temporales en diversos escenarios de movimiento; el desarrollo de herramientas y aplicaciones vinculadas con bases de datos espaciales y espacio-temporales; y la aplicación de la geometría computacional y su marco disciplinar en aspectos particulares de los problemas estudiados.

Como objetivos específicos se propone:

- Estudiar la indexación espacial y espacio temporal sobre objetos en movimiento en diversos escenarios. Implementar estructuras de almacenamiento, algoritmos de consulta y de evaluación experimental.
- Aplicar bases de datos espaciales y espacio-temporales, con apoyo de métodos

provistos por la geometría computacional en la resolución de problemas reales.

- Diseñar e implementar aplicaciones (app) y/o herramientas informáticas para diversos dispositivos.

3. RESULTADOS OBTENIDOS

Desarrollo de herramientas tecnológicas de apoyo para la gestión de la producción agropecuaria. En el contexto de proyecto Campo Conectado, se identificaron diversos requerimientos relacionados con la producción ganadera. Estos refieren a la necesidad de realizar un seguimiento de indicadores tanto productivos como ambientales, caracterizar y hacer seguimiento de cada individuo de un rodeo, garantizar la trazabilidad animal, controlar la alimentación del ganado, detectar el momento de celo, monitorear la condición corporal, entre otros aspectos. Todos ellos de interés en el ámbito de la ganadería de precisión.

Para satisfacer algunos de estos requerimientos, se propuso una plataforma que soporte eventos y sistemas de información. Esta plataforma está diseñada para seguir de forma espacial y temporal los rodeos, así como gestionar de modo tradicional y espacial los elementos que conforman un establecimiento agropecuario [5] [7].

En la plataforma integral se incluye una herramienta para el seguimiento espacial y temporal de individuos de rodeos ganaderos. Esta herramienta posibilita procesar tanto los datos tradicionales como los datos espaciales, permitiendo realizar consultas tradicionales, así como consultas espaciales y espacio-temporales sobre los objetos de estudio. Además, la herramienta ofrece una interfaz de comunicación con las bases de datos y posibilita la visualización del seguimiento geográfico del rodeo mediante el uso de dispositivos móviles [6] [10].

Trazabilidad de vehículos en estado de sospecha ante situaciones ilegales. Se propone la implementación de un sistema que sea capaz de analizar la circulación de

vehículos que pueden encontrarse en estado de sospecha o ilegalidad, con el fin de hallar posibles sospechosos de un potencial ilícito. Para lograr esto, se propone analizar los movimientos de los vehículos mediante la lectura de patentes en zonas barriales, rutas o caminos cercanos al lugar del hecho y fronteras provinciales [8].

Actualmente, se está trabajando en el desarrollo de un prototipo de la plataforma para simular casos de seguimiento y evaluar la instalación de los recursos técnicos necesarios para la implementación del sistema propuesto. El diseño de este prototipo será un trabajo final de Licenciatura en Ciencias de la Computación.

Gestión Integral de Incendios Forestales. En el marco del proyecto, el grupo aporta en el diseño y asistencia técnica para el desarrollo de bases de datos espaciales, procesamiento y optimización de consultas a bases de datos existentes asociadas a la problemática.

Además, se están estudiando herramientas especializadas para la gestión y análisis de información geográfica con el objetivo de presentar información relevante para el seguimiento y análisis de las transformaciones ambientales. El propósito es brindar apoyo en la toma de decisiones relacionadas con la prevención, gestión e impacto de los incendios forestales, así como en la recuperación del medio ambiente. Se trabaja en colaboración con integrantes de tres unidades académicas de la Universidad Nacional de San Luis, específicamente la Facultad de Ciencias Físicas, Matemáticas y Naturales (FCFMyN), la Facultad de Química, Bioquímica y Farmacia (FQBF) y la Facultad de Ingeniería y Ciencias Agropecuarias (FICA).

Como ejes de trabajo futuros, se propone continuar con actividades que involucran la pertinencia y transferencia de proyectos I+D+i en las áreas de intervención seleccionadas. Se proyecta maximizar la sinergia entre grupos de trabajo, afianzar el ámbito interdisciplinario de investigación y fortalecer las competencias de los investigadores.

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de la línea de investigación fortalecen su formación mediante actividades de cooperación e intercambio de información científica, tecnológica y la generación de nuevos conocimientos con investigadores locales y de otras universidades.

La formación de recursos humanos se refleja en la dirección de tesis de maestría y trabajos finales de Licenciatura en Ciencias de la Computación, así como en la dirección de becas y dictado de cursos de posgrado. Además, se llevan a cabo actividades de formación académica, tales como cursos de especialización, participación en actividades de divulgación científica, conferencias y otras actividades académicas y científicas relacionadas.

5. BIBLIOGRAFÍA

- [1] *AgroTic: Revolución Digital, Desarrollo y Producción*.
<https://inta.gob.ar/noticias/agrotic-revolucion-digital-desarrollo-y-produccion> (2016)
- [2] *Campo Conectado*, proyecto interinstitucional
<http://inta.gob.ar/noticias/campo-conectado-un-nuevo-proyecto-interinstitucional> (2017)
- [3] de Berg, M., Cheong O., van Kreveld, M., Overmars, M. : *Computational Geometry: Algorithms and Applications.*, Springer-Verlag, Heidelberg (2008)
- [4] Elmasri, R. Navathe, S.B.: *Fundamentals of Database Systems*. 7^{ma} Edition. Pearson (2015)
- [5] Gagliardi, E., Dorzán, M. G., Taranilla, M. T., Palmero, P., Casanova, C.: *Propuesta de plataforma para la integración de TIC orientadas al Agro* Congreso de AgroInformática CAI 2017, 46JAIIO (2017)
- [6] Gagliardi, E.; Dorzán, M.G.; Taranilla, M.T.; Palmero, P.; Casanova, C.: *GeoSeguimiento de Rodeos, hacia una plataforma integral para el Agro*. Congreso de AgroInformática CAI 2018, 47JAIIO (2018).

- [7] Gagliardi, E.; Dorzán, M.G.; Taranilla, M.T.; Palmero, P.; Casanova, C.: *Diseño de una herramienta de apoyo en la gestión de rodeos de cría*. Congreso de AgroInformática CAI 2018, 47JAIIO (2018).
- [8] Leiva Olguin G., Martinez R., Paez B., Gagliardi E. O. Diseño de un sistema de trazabilidad de vehículos en estado de sospecha. Congreso Nacional de Ingeniería Informática / Sistemas de Información- Universidad Tecnológica Nacional. (2022)
- [9] Manolopoulos, Y., Papadopoulos, A., Vassilakopoulos, M.: *Spatial Databases: Technologies, Techniques and Trends*. Idea Group (2005)
- [10] Palmero, P.; Gagliardi, E.; Dorzán, M.G.; *Seguimiento de rodeos en establecimientos agropecuarios*. CACIC: XXVI Congreso Argentino de Ciencias de la Computación (2020)
- [11] *Prevención y Gestión Integral de Incendios Forestales* - Programa de Investigación Institucional en el área estratégica Medio Ambiente. Universidad Nacional de San Luis. Res. N° CS 237-21 (otorgamiento)- Res. N° CS 1931-21 (integrantes) - UNSL.
- [12] Shekhar, S.; Chawla, S.: *Spatial databases: a tour*. Prentice Hall (2003)

Bases de Datos no Convencionales

M. D. Alba, J. Arroyuelo, M. E. Di Genaro, A. Grosso, M. Jofré, V. Ludueña, N. Reyes, D. Welch
Dpto. de Informática, Fac. de Cs. Físico-Matemáticas y Naturales, Universidad Nacional de San Luis
{bjarroyu, digeme, agrosso, vlud, nreyes, dwelch}@unsl.edu.ar, {mdaniela.alba, monicajofre}@gmail.com

Edgar Chávez

Centro de Investigación Científica y de Educación Superior de Ensenada, México
elchavez@cicese.mx

Karina Figueroa

Fac. de Cs. Físico-Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, México
karina@fismat.umich.mx

Rodrigo Paredes

Dpto. de Cs. de la Computación, Fac. de Ingeniería, Universidad de Talca, Chile
raparedede@utalca.cl

Resumen

El avance que han tenido las nuevas tecnologías en todos los ámbitos de la vida, ha ocasionado la producción de infinidad de datos digitales tan disímiles como lo son sus fuentes. Así, se generan datos provenientes de sensores de todo tipo, resultados de estudios médicos, datos científicos, secuencias biológicas, información obtenida de satélites, etc.; datos complejos que necesitan ser procesados para atender a los variados requerimientos del mundo actual. Este contexto motivó que las bases de datos debieran adaptarse, para ser capaces de administrar eficientemente grandes volúmenes de estos nuevos y diversos tipos de datos y así atender esos requerimientos.

El enfoque que se ha vuelto popular en los últimos años, para manejar las diversas bases de datos emergentes de objetos complejos y no estructurados, es el de *espacios métricos* y es el utilizado en las investigaciones presentadas en este artículo. Las mismas pretenden contribuir a la madurez de este modelo de bases de datos, *Bases de Datos Métricas*, considerando distintas perspectivas: administración del espacio disponible (crucial por la gran cantidad de datos); formas más sofisticadas de búsqueda sobre las mismas; optimización de estos depósitos, o desarrollo de nuevos, entre otros.

Palabras Claves: bases de datos métricas, índices, búsquedas por similitud.

Contexto

Las investigaciones mencionadas se realizan en el marco de la línea *Bases de Datos no Convencionales* perteneciente al Proyecto Consolidado *Tecnologías Avanzadas de Bases de Datos* que se desarrolla en la Universidad Nacional de San Luis (Código 03-

2218) y en el Programa de Incentivos (Código 22-F814) finalizado en diciembre de 2022. Actualmente se realizó una nueva presentación para su continuación. Dicho proyecto está incorporado al Laboratorio de Investigación y Desarrollo en Bases de Datos (LaBDa) en el cual colaboran investigadores de otros grupos de la región: Universidad de Talca (Chile), Universidad Michoacana de San Nicolás de Hidalgo y Centro de Investigación Científica y de Educación Superior de Ensenada (México).

El presente artículo expone el trabajo que se desarrolla en este ámbito, orientado a contribuir al modelo de Bases de Datos Métricas, dado que todavía no se ha alcanzado todo el potencial en las soluciones, el problema de las búsquedas sobre estos tipos de bases de datos es un problema abierto en varios aspectos. Para ello se trabaja, obteniendo nuevas estructuras de datos (índices) que resulten eficientes para memorias jerárquicas, que sean dinámicos, manejen grandes volúmenes de datos (escalables), con eficiencia en E/S y adecuados a la variedad de datos no estructurados existentes. De esta manera, se espera aportar a diferentes campos de aplicación: diseño asistido por computadora, sistemas de información geográfica, robótica, visión artificial, biología computacional, computación móvil, entre otros.

Introducción

La generación de grandes cantidades de datos digitales, provocada por el uso generalizado de dispositivos capaces de producirlos, aumenta constan-

temente. Así, tanto la cantidad como la variedad de datos provenientes de ámbitos tan diferentes como el de la salud, el educativo, productivo, laboral, recreativo, científico, etc. ha crecido de manera exponencial. En este contexto los repositorios especializados en datos *no estructurados* se vuelven indispensables. Como respuesta, las bases de datos han debido adaptarse rápidamente, tanto a la cantidad de datos que deben administrar, como a su variedad y disimilitud.

Más aún, son diferentes los requerimientos que las aplicaciones actuales plantean sobre estos datos; encontrar las huellas digitales más similares a una dada, u obtener composiciones semejantes a un trozo de una melodía, son algunos ejemplos. En estos casos las búsquedas tradicionales (exactas) carecen de sentido, en cambio las *búsquedas por similitud* resultan más adecuadas; en ellas, se suele dar un objeto como “ejemplo” de lo que se quiere recuperar y se busca en la base de datos los objetos que sean suficientemente similares a la muestra. Estas búsquedas son coincidas como consultas por contenido o consultas mediante un ejemplo (query by example).

Un modelo adecuado para la variedad de aplicaciones que utilizan búsquedas por similitud es el de *espacios métricos*, ya que, a pesar de las necesidades tan diversas de las mismas, todas comparten ciertas características que hacen a este modelo el más adecuado como su marco formal. Este modelo es determinado por un universo de objetos y una función de distancia definida entre ellos, que mide cuán diferentes son. Cualquier tipo de objetos no estructurados, que admita la definición de una medida que indique cuán diferentes son dos objetos, admite ser modelizado de esta manera. La única restricción es que esa medida cumpla con las propiedades que la hagan una métrica. En general, esas medidas son provistas por expertos (por ejemplo: distancias para comparar huellas dactilares).

Sin embargo, para responder eficientemente a los requerimientos tan dispares que se realizan sobre los objetos no estructurados almacenados en estas bases de datos, son necesarios los llamados *Métodos de Acceso Métricos* (MAMs), que permiten resolver este tipo de búsquedas, sin realizar una examinación secuencial del conjunto de datos [6, 9]. La diversidad de ámbitos en los que se aplica el modelo, vuelve esencial la actualización y optimización de los MAMs, el adaptarlos a cada caso particular y afrontar retos como soportar conjuntos masivos de datos, permitir actualizaciones (inserciones/eliminaciones) y resolver búsquedas más complejas.

Líneas de Investigación y Desarrollo

Bases de Datos Métricas

Como se expresó anteriormente, las *bases de datos no convencionales*, aquellas capaces de administrar vídeos, imágenes, texto libre, secuencias de proteínas o ADN, audio, etc., son modelizadas utilizando el *modelo de espacios métricos*, por lo que también se las conoce como *bases de datos métricas*. En este modelo, debido a lo costoso que suele resultar calcular la distancia entre dos objetos, el número de cálculos realizados durante alguna operación se usa habitualmente como medida de complejidad. Estos cálculos son necesarios tanto para crear un índice como para realizar búsquedas sobre él. Por lo tanto, para responder eficientemente a los distintos requerimientos sobre las mismas, evitando comparaciones exhaustivas sobre la base de datos, se requiere el uso de índices especializados.

Ésta es la razón por la cual optimizar los mismos se ha vuelto un objetivo prioritario, analizando aquellos que han mostrado buen desempeño, principalmente en las búsquedas, para reducir su complejidad. Para ello, cuando sea necesario, se considera el nivel de la memoria en la que se lo alojará, su capacidad de ser dinámico y de ser escalable.

Formalmente, un espacio métrico está definido por un universo de objetos $\mathbb{U}$ y una función de distancia entre ellos $d : \mathbb{U} \times \mathbb{U} \mapsto \mathbb{R}^+$, la cual permite medir la disimilitud entre los objetos del universo. En particular, d cumple con las propiedades de una métrica (reflexividad, positividad estricta, simetría y desigualdad triangular), lo que resulta muy útil al momento de resolver consultas por similitud. De sus propiedades, las que permiten ahorrar cálculos de distancia son la simetría y la desigualdad triangular. Conceptualmente, una base de datos X es un conjunto $X \subseteq \mathbb{U}$. Además, una consulta por similitud se expresa brindando un elemento $q \in \mathbb{U}$ y, en general, suele ser de dos tipos: *por rango* (q, r) o de *k-vecinos más cercanos* ($k\text{-NN}(q)$).

Métodos de Acceso Métricos

Como se comentó anteriormente, optimizar los índices métricos se ha vuelto un objetivo prioritario, y teniendo en cuenta el variado espectro de aplicación de los mismos, se deben considerar diferentes puntos de vistas para hacerlo. Hay que considerar su dinamismo; su escalabilidad; analizar si se adaptan adecuadamente al almacenamiento en memoria secundaria; entre otros. En este último caso, se de-

be considerar un cambio en la medida de complejidad, y que no sólo se debe tener en cuenta el número de cálculos de distancia, sino también el número de operaciones de E/S necesarias.

En relación al dinamismo, un buen representante en este sentido es el *Árbol de Aproximación Espacial Dinámico (DSAT)* [13], que es la versión dinámica de uno de los índices de mejor desempeño en espacios métricos de mediana a alta dimensión el *Árbol de Aproximación Espacial (SAT)*. El *DSAT* mantiene un buen desempeño en las búsquedas pero agrega un parámetro a sintonizar. Siguiendo la idea del *DSAT*, se emprendió la tarea de proveer de dinamismo al *Árbol de Aproximación Espacial Distal (DiSAT)* [5], que además de no necesitar ningún parámetro extra, lograba mejorar las búsquedas respecto de sus antecesores: *SAT* y *DSAT*, pero también era estático. Así, en una primera instancia se propuso el desarrollo de la *Foresta de Aproximación Espacial Distal (DiSAF)* [2], que es dinámica y para memoria principal. Este índice aplica la técnica de dinamización de Bentley y Saxe al *DiSAT* y aprovecha el profundo conocimiento que se tiene sobre la aproximación espacial para mejorar al máximo su desempeño. Sin embargo, los costos de construcción son altos debido a la necesidad de reconstruir subárboles luego de cada inserción.

Para solucionar este problema, se diseñó el *Árbol de Aproximación Espacial Distal Dinámico (DDiSAT)* [7]. Esta nueva versión utiliza *inserción perezosa* para amortizar los costos de reconstrucción, a la vez que mejora su desempeño en las búsquedas, sin usar espacio adicional. La propuesta es retrasar la inserción de nuevos elementos, asignándolos (en una bolsa) a su objeto más cercano en el árbol. Los objetos se insertan en las bolsas hasta que la cantidad de elementos pendientes iguala al número de nodos del árbol. En ese momento, se reconstruye el árbol completo. El (DDiSAT) reduce significativamente los costos de construcción con respecto a *DiSAF* y, sorprendentemente, obtiene un mejor rendimiento de búsqueda que sus parientes cercanos estáticos (*SAT* y *DiSAT*) y dinámico (*DSAT*). Ello se debe a la buena partición del espacio y a la buena distribución de elementos que hace el índice. Así, se sigue trabajando en posibles mejoras, por ejemplo, diseñando un algoritmo de carga masiva para crear el DDiSAT, si se conoce de antemano un subconjunto de elementos, evitando algunas reconstrucciones y combinándolo con la inserción diferida.

Al considerar el almacenamiento de un índice,

muchas veces sucede que los mismos no caben en memoria principal debido al tamaño de los objetos almacenados en él o la gran cantidad de objetos que almacena. Entonces surge la necesidad de diseñar MAMs que se puedan almacenar en memoria secundaria sin perder eficiencia. En este contexto, se está trabajando en lograr una versión dinámica del *DDiSAT* para memoria secundaria, que además de amortizar los costos de reconstrucción entre varias inserciones y mantener un buen desempeño en las búsquedas, se adapte a memoria secundaria realizando un buen uso de las páginas de disco para minimizar el número de operaciones de E/S. Para lograrlo, se debe considerar no sólo que logre un desempeño comparable al de la versión de memoria principal en cantidad de cálculos de distancia, sino que las páginas en disco mantengan una buena ocupación, que la cantidad de operaciones de E/S necesarias sea baja y que se mantenga la localidad en los accesos.

Otra faceta a considerar se relaciona con aplicaciones que priorizan la rapidez en las respuestas, a costa de perder algunos elementos de la misma. Este tipo de búsquedas, en las que se intercambia precisión (devolviendo sólo algunos objetos relevantes) por velocidad en la respuesta (esos objetos se devuelven más rápido), se conocen como *búsquedas aproximadas*. Cuando se trabaja con conjuntos de datos masivos, las búsquedas por similitud aproximadas permiten obtener un buen balance entre costo de las búsquedas y calidad de la respuesta obtenida. En este contexto, se propuso la *Lista Dinámica de Permutaciones Agrupadas (DLCP)* [11], que además de ser dinámica y para memoria secundaria, permitía realizar consultas aproximadas. Sin embargo, algunos aspectos de la estructura eran mejorables [12]. Así, se está trabajando en combinar el *Conjunto Dinámico de Clusters (DSC)* [13], un índice para memoria secundaria con muy buen desempeño; con uno de los mejores representantes para consultas aproximadas, el *Algoritmo Basado en Permutaciones (PBA)* [1], para lograr un nuevo índice para búsquedas aproximadas, con buena calidad de respuesta y eficiente en memoria secundaria.

Además, se espera poder incorporar estos nuevos índices, tipos de búsqueda y otras operaciones de interés en un DBMS que trabaja sobre bases de datos métricas, el cual se basa en PostgreSQL. Esto permitirá aumentar su capacidad de administración y procesamiento para grandes bases de datos.

Búsqueda de los k Vecinos

Entre las consultas por similitud en espacios métricos, una que resulta muy útil es la que obtiene los k -vecinos más cercanos de un elemento dado; este es el tipo de consulta que utilizamos cuando se quiere encontrar los restaurantes más cercanos a un lugar en particular. Una variación muy útil de la misma, es la *All- k -NN*, que relaciona cada elemento $u \in X$, con los k objetos en $X - \{u\}$ que tengan la menor distancia a él. La forma ingenua de resolverlo es comparar cada objeto en la base de datos con todos los demás y devolver los k más cercanos a él. Esta solución tiene una complejidad de n^2 cálculos de distancia ($|X| = n$). Una solución más eficiente es preprocesar la base de datos construyendo un índice y luego buscando en el mismo los k -NN de cada elemento del conjunto.

Sin embargo, existen situaciones en las cuales el costo de la construcción del índice, para luego realizar n consultas del tipo k -NN, puede resultar tan excesivo como la solución ingenua. Este es el caso de administrar una base de datos masiva, o cuando la función de distancia es demasiado costosa de calcular, o si se está trabajando con espacios métricos de alta dimensión. Estos casos pueden requerir revisar la base de datos completa, a pesar de la estrategia utilizada. Otro factor a considerar son los requerimientos de algunas aplicaciones particulares, que priorizan la velocidad de respuesta sobre la precisión de la misma [15, 6, 16, 10]. Para hacer frente a éstas circunstancias es que se han considerado las llamadas *búsquedas por similitud aproximadas*, las mismas mejoran su complejidad aceptando algunos “errores” en la respuesta.

Se sabe que al resolver el problema *All- k -NN* permite construir el *Grafo de los k -vecinos más cercanos* (k NNG)[14]. Dada una colección de objetos de un espacio métrico, el grafo de k vecinos más cercanos asocia cada nodo a sus k vecinos más cercanos. El k NNG resulta ser un índice eficiente, que admite mejoras y permite resolver búsquedas por similitud. Por ello hemos propuesto nuevas técnicas para resolver el problema de *All- k -NN*, que *no utiliza ningún índice* para buscar en él, y que permiten computar una aproximación del k NNG. Éstas conectan cada objeto u de la base de datos con k vecinos *cercanos*, relajando la condición que exige que no haya, en toda la base de datos, algún objeto más cercano a u que los k vecinos devueltos. Esto puede provocar que la respuesta pierda algún objeto muy cercano, devolviendo uno más lejano en su lugar, a cambio de una

respuesta más rápida. Este grafo se denominó *Grafo de vecinos cercanos* (k nNG) [4].

Una primera propuesta, utilizando un enfoque novedoso, surge aprovechando el profundo conocimiento que se tiene del *DiSAT* [5]. Aquí se consideró un caso particular del problema ($k = 1$) obteniendo el 1nNG. Esta propuesta utiliza la información obtenida durante la *construcción* del *DiSAT* para construir el 1nNG, conectando a cada elemento con un elemento cercano de la base de datos, que puede ser, o no, su vecino más cercano [4]. Con este algoritmo se consigue recuperar el 1nNG logrando un buen compromiso calidad/tiempo; los costos de las consultas resultan ser bajos, se alcanza muy buena precisión en la respuesta con un error bajo y todo esto *sin realizar ninguna búsqueda en el índice*.

Se realizaron otras tres propuestas que, a diferencia de la primera, se enfocan en el problema más general al responder a los *All- k -nN* y computar el k nNG. Estos desarrollos no utilizan el apoyo de ningún índice, no sólo no buscan en ellos, sino que ni siquiera recurren a la información provista por su construcción. La base de estos algoritmos es aprovechar de manera ingeniosa las propiedades que cumple la *función de distancia*. Para hacerlo, se proponen diferentes maneras de seleccionar muestras de la base de datos, a partir de las cuales se obtiene un conjunto de distancias que serán el punto de partida para resolver la consulta. Se analizan diferentes maneras de utilizar la información obtenida para calcular, en algunos casos, los vecinos exactos [3] y en otros los aproximados de todos los objetos de la base de datos, utilizando las propiedades de simetría o desigualdad triangular [8].

Resultados y Objetivos

Las investigaciones realizadas sobre el modelo de espacios métricos, permitieron mejorar el desempeño de los MAMs analizados, los resultados obtenidos conducen a estudiar su aplicación a otros métodos de acceso [7, 8, 3, 4, 12, 2].

Se espera brindar nuevas herramientas eficientes de administración para bases de datos métricas, que logren acercar su desarrollo al de los modelos tradicionales de base de datos. Así, se buscará profundizar en el estudio de nuevos diseños de estructuras de datos, buscando incrementar su eficiencia en espacio y en tiempo: que se adapten mejor al nivel de la jerarquía de memorias donde se almacenarán y a las características de los datos a ser indexados. Del mismo modo, se continuará indagando sobre técni-

cas innovadoras que, sin utilizar índices, permitan resolver consultas eficientemente.

Actividades de Formación

Dentro de esta línea de investigación se forman alumnos y docentes-investigadores participando en:

- **Maestría en Cs. de la Computación** (UNSL): tesis sobre una versión dinámica eficiente del *DiSAT*.
- **Maestría en Ing. de Software** (UNSL): tesis sobre una aplicación de estos índices a un DBMS métrico.
- **Maestría en Informática** (UNSJ): tesis sobre la evaluación del *knNG* para búsquedas por similitud.
- **Maestría en Informática** (UNSJ): tesis sobre una versión dinámica y para memoria secundaria del *DDiSAT*.

Referencias

- [1] E. Chávez, K. Figueroa, and G. Navarro. Effective proximity retrieval by ordering permutations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(9):1647–1658, Sept 2008.
- [2] E. Chávez, M. Di Genaro, N. Reyes, and P. Roggero. Decomposability of *disat* for index dynamization. *Computer Science & Technology*, pages 110–116, 2017.
- [3] E. Chávez, V. Ludueña, and N. Reyes. Solving all-*k*-nearest neighbor problem without an index. In *Procs. del XXV Congreso Argentino de Ciencias de la Computación*, pages 567–576. UniRío editora, 2019.
- [4] E. Chávez, V. Ludueña, N. Reyes, and F. Kasán. All near neighbor graph without searching. *Computer Science & Technology*, 18:61–67, 2018.
- [5] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal {SAT}. *Information Systems*, 59:15 – 47, 2016.
- [6] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [7] Edgar Chávez, María E. Di Genaro, and Nora Reyes. An efficient dynamic version of the distal spatial approximation trees. In *Actas del XXVIII Congreso Argentino de Ciencias de la Computación*, pages 468–477, Oct. 2022.
- [8] Edgar Chávez, Verónica Ludueña, and Nora Reyes. Heuristics for computing *k*-nearest neighbors graphs. In Patricia Pesado and Marcelo Arroyo, editors, *Computer Science – CA-CIC 2019*, pages 234–249, Cham, 2020. Springer.
- [9] Lu Chen, Yunjun Gao, Xuan Song, Zheng Li, Yifan Zhu, Xiaoye Miao, and Christian S. Jensen. Indexing metric spaces for exact similarity search. *ACM Comput. Surv.*, 55(6), dec 2022.
- [10] P. Ciaccia and M. Patella. Approximate and probabilistic methods. *SIGSPATIAL Special*, 2(2):16–19, 2010.
- [11] K. Figueroa, C. Martínez, R. Paredes, N. Reyes, and P. Roggero. Dynamic list of clustered permutations on disk. In *Computer Science and Technology Series: XXI Argentine Congress of Computer Science Selected Papers*, pages 201–211. EDULP, 2016.
- [12] K. Figueroa, N. Reyes, A. Camarena-Ibarrola, and L. Valero-Elizondo. Improving the list of clustered permutation on metric spaces for similarity searching on secondary memory. In *10th Mexican Conference on Pattern Recognition*, volume 10880, pages 82–92, 2018.
- [13] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.
- [14] R. Paredes, E. Chávez, K. Figueroa, and G. Navarro. Practical construction of *k*-nearest neighbor graphs in metric spaces. In *Proc. 5th Workshop on Efficient and Experimental Algorithms*, LNCS 4007, pages 85–97, 2006.
- [15] H. Samet. *Foundations of Multidimensional and Metric Data Structures (The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2006.
- [16] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach (Advances in Database Systems)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005. XVIII, 220 p., Hardcover ISBN: 0-387-29146-6.

Comparación de métodos y métricas para la predicción de series temporales aplicados a las ventas masivas de productos minoristas

Rodrigo Gonzalez, Juan Carlos Taffernaberr, Ana Laura Diedrichs Escudero, y Mario Sebastián Tobar

Universidad Tecnológica Nacional, Facultad Regional Mendoza, GridTICs,
Mendoza, Argentina.
{rodraz, cartaffe, mstobar, ana.diedrichs}@frm.utn.edu.ar

RESUMEN

La generación de pronósticos de ventas (sales forecasting) es una herramienta fundamental en la planificación de inventarios de una empresa. En el caso de grandes empresas de venta minorista (retail) como Amazon, Walmart o Mercado Libre el problema se vuelve muy complejo por la cantidad de productos cuyas ventas se deben predecir. Este proyecto se propone comparar los resultados de las competencias M5 del año 2020 y el Meli Data Challenge 2021, las cuales se basaron en predecir en forma masiva las ventas de productos minoristas. Si bien las mencionadas competencias parecen bastante similares, una notable diferencia entre ambas es que usaron distintas métricas para evaluar el desempeño de los modelos propuestos. Esto no permite que se pueda hacer una comparación directa entre los mejores modelos que obtuvieron los mejores puestos en ambos concursos, y por ende, no se puede llegar a una conclusión respecto a cuál modelo es mejor. Por tanto, este proyecto propone determinar si los algoritmos que mejor desempeño mostraron en una competencia también mostrarán un buen desempeño con los datos del otro certamen. Complementariamente, se tratará de entender qué influencia tienen diferentes métricas para determinar si un método de predicción es mejor que otro.

Palabras Claves: Series temporales, Ciencia de datos, Predicción masiva, ventas minoristas.

CONTEXTO

El presente proyecto se desarrolla en el marco del grupo GridTICs, Departamento de

Ingeniería Electrónica. El trabajo es parte de las líneas de investigación llevadas adelante por el Dr. Rodrigo Gonzalez.

1. INTRODUCCIÓN

La generación de pronósticos de ventas (sales forecasting) es un aspecto crucial en la planificación de inventarios de una empresa. Estas requieren predicciones de ventas para respaldar su desarrollo estratégico, tomar decisiones tácticas y administrar sus procesos de planificación de la oferta y la demanda, para de esta forma evitar problemas de servicios al cliente y altos costos de inventario (Fildes et al., 2019). Por lo tanto, es importante que estas predicciones sean lo más precisas posible, ya que almacenar demasiados productos incurre en costos adicionales, mientras que un almacenamiento insuficiente conduciría a la pérdida de ventas y bajas ganancias. En el caso de grandes empresas de venta minorista (retail) como Amazon, Walmart o Mercado Libre el problema se vuelve más complejo por la cantidad de productos cuyas ventas se deben predecir. Por otra parte, la demanda de muchos de estos productos es muy estacional o está impulsada por eventos externos o decisiones internas (por ejemplo, cambios de precios o promociones). Además, nuevos productos son continuamente añadidos al catálogo. Esto produce que muchos productos tengan demandas escasas e intermitentes, lo que agrega complejidad a la predicción de esta serie de tiempo en particular. Muchos de estos problemas son especialmente desafiantes para los métodos lineales de predicción tradicionales como ARIMA, ETS (Error, Trend, Seasonal) y algunas técnicas de

modelado aún más avanzadas (Seeger et al. 2016).

Las competencias Mx son un certamen internacional que propone un problema de predicción para que diferentes equipos muestren nuevos métodos para mejorar la precisión de predicción, tanto en valores absolutos como en distribuciones de probabilidad. Se han llevado a cabo durante casi 40 años. Los hallazgos obtenidos en estas competencias han influido significativamente en la teoría y la práctica de predicción en series temporales, ya que han proporcionado valiosa información sobre cómo se puede mejorar la precisión de los pronósticos (Hyndman, 2020, Semenoglou et al., 2021).

En particular, la competencia M5 amplió los objetivos de los cuatro certámenes anteriores centrándose en la predicción de ventas minoristas. Se proveyeron datos de ventas estructurados jerárquicamente de la cadena de supermercados minoristas Walmart, cuyas series de tiempo muestran características intermitentes y erráticas (Syntetos y Boylan, 2005, Syntetos et al., 2005). El conjunto de datos utilizado comprendió las ventas unitarias de 3.049 productos vendidos por Walmart en los EE. UU., los cuales se organizaron en forma de series temporales agrupadas en función de su tipo (categoría y departamento) y ubicación de venta (tiendas y estados), con un total de 42.840 series en 12 niveles transversales de agregación. La M5 se desarrolló durante el año 2020 y fue organizada en el marco del sitio Kaggle.com (Kaggle, 2022). Esto permitió conocer el enfoque aplicado por la mayoría de los participantes. Además, se cuenta con abundante bibliografía en la que se detallan varios aspectos de la competencia y qué algoritmos mostraron mejor desempeño de predicción (Makridakis, 2021) (Makridakis, 2022) (Lainder, 2022). Uno de los algoritmos de aprendizaje automático que mejor desempeño mostró y que fue usado en la práctica por los 50 principales competidores es LightGBM, un algoritmo basado en un árbol de decisión. Sin embargo, también se descubrió que los métodos fáciles de implementar y computacionalmente baratos,

como ETS, siguen siendo competitivos, especialmente cuando se usan para producir pronósticos a nivel de producto o tienda de productos.

Por otra parte, a mediados del año 2021 la empresa de ventas minorista Mercado Libre propuso una competencia similar a la M5 a la que llamó Meli Data Challenge 2021 (MDC) (Mercado Libre, 2021). En este caso, los datos suministrados incluían las ventas diarias por un periodo de 2 meses de alrededor de 400.000 productos en 3 países donde opera la empresa: Argentina, Brasil y México. El objetivo central era crear un modelo para obtener una distribución de probabilidad que indicara en cuántos días los artículos quedarían sin inventario, en función de los datos históricos de Mercado Libre y el stock existente. Al contrario de la competencia M5 en Kaggle, Meli Data Challenge 2021 no propició un foro para que los participantes pudieran intercambiar opiniones o publicar sus soluciones al desafío. Afortunadamente, los principales competidores han publicado en Github.com sus enfoques. Por ejemplo, el ganador de la competencia publicó su solución en código Python (Moreyra, 2021), la cual se basó en combinar una red neuronal recurrente (recurrent neural network, RNN) con un regresor XGBoost.

Si bien las mencionadas competencias parecen bastante similares, una notable diferencia entre ambas es que usaron métricas totalmente diferentes para evaluar el desempeño de los modelos propuestos. Esto no permite que se pueda hacer una comparación directa entre los mejores modelos propuestos en ambos concursos, y por ende, no se puede llegar a una conclusión respecto a cuál modelo es mejor. Por tanto, está pendiente determinar si los algoritmos que mejor desempeño mostraron en una competencia también mostrarán un buen desempeño con los datos del otro certamen. O por el contrario, si los mejores algoritmos de ambas competencias no pueden ser generalizados para cualquier serie de tiempo relacionada con la predicción de ventas, y son dependientes de los datos a los que fueron expuestos. Finalmente, está pendiente

entender qué influencia tienen diferentes métricas para determinar si un método de predicción es mejor que otro.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para el desarrollo del presente proyecto pueden diferenciarse 4 etapas principales.

1. Recabar información del estado del arte sobre algoritmos para la predicción de series de tiempo relacionadas con las ventas masivas de productos minoristas. En una primera etapa se realizará un estudio del estado del arte sobre la temática en referencia. Se analizarán los métodos ya propuestos en la literatura como ARIMA, ETS, LightGBM y XGBoost, sus alcances y limitaciones.

2. Seleccionar e implementar algoritmos para la predicción de series de tiempo relacionados con las ventas masivas de productos minoristas. En una segunda etapa se implementarán en R o Python (según bibliotecas disponibles) los algoritmos del estado del arte sobre la temática en referencia estudiados en la etapa 1. También se implementarán las soluciones que mostraron mejor desempeño en los desafíos M5 y Meli Data Challenge.

3. Considerar las métricas adecuadas para la comparación de predicciones de series de tiempo relacionadas con las ventas masivas de productos minoristas. Será esencial entender las diferentes métricas que existen en el área en referencia y qué métricas son las más adecuadas para comparar los algoritmos implementados en la etapa 2. Las mismas se implementarán en R o Python (según bibliotecas disponibles).

4. Evaluar el desempeño de los métodos del estado del arte para predicción de series de tiempo relacionadas con las ventas minoristas masivas con diferentes conjuntos de datos.

En una cuarta etapa se evaluará la hipótesis central del proyecto propuesto: si los algoritmos del estado del arte para la temática en referencia pueden generalizarse para diferentes conjuntos de datos, o en otras palabras, si estos pueden ser implementados en diferentes empresas de ventas minoristas masivas.

3. RESULTADOS ESPERADOS

Al término del tiempo de duración de este proyecto, se espera obtener los siguientes resultados:

1. Sólida comprensión de los beneficios y limitaciones de los métodos y métricas del estado del arte para la predicción de ventas minoristas masivas.

2. Creación de framework para el análisis y la visualización de series temporales, su modelado y predicción.

3. Contar con los conocimientos y herramientas necesarios para poder asistir a diferentes empresas a mejorar sus procesos de inventario y ventas, al poder predecir con una determinada exactitud la venta de los artículos que comercializan.

Finalmente, los resultados obtenidos serán publicados tanto en conferencias internacionales de alta calidad con referato, como en revistas especializadas de alto impacto, parte del catálogo de Scopus y Web of Science.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo propuesto estará integrado por 4 investigadores:

1. El Dr. Ing. Gonzalez, propuesto como director, es Profesor Adjunto de Ingeniería Electrónica e investigador del grupo GridTICs. Posee experiencia en investigación y dirección de proyectos desde el año 2010. Es especialista en procesamiento de señales, análisis de series de tiempo, álgebra lineal, estadística, procesos estocásticos, en el uso de técnicas de aprendizaje automático en series de tiempo (Gonzalez and Catania, 2019), programación en R y SQL. El Dr. Gonzalez participó del Meli Data Challenge, por lo que conoce la competencia en detalle. Obtuvo el puesto 84 de 162.

2. El Ing. Taffernaberry, propuesto como codirector, es Profesor Adjunto en Ingeniería Electrónica en Ingeniería Electrónica (UTN FRM) e investigador del grupo GridTICs. Cuenta con basta experiencia en investigación, dirección de proyectos de investigación, análisis de series temporales (Watteyne, 2018) y programación en Python.

3. La Ing. Diedrichs es docente de Informática II en Ingeniería Electrónica (UTN FRM) e investigadora del grupo GridTICs. Cuenta con experiencia en investigación desde 2014. Su tesis de doctorado está actualmente bajo revisión y es probable que antes de fin de año obtenga el título de Doctora en Ingeniería mención civil-ambiental (UTN FRM). Es especialista en el área de ciencia de datos. Cuenta con experiencia en la resolución de diferentes problemas aplicando algoritmos específicos del área de aprendizaje automático y ciencia de datos (Diedrichs, 2018, Watteyne, 2018). Es especialista en programación en R.

4. El Ing. Tobar, docente de Técnicas Digitales III en Ingeniería Electrónica (UTN FRM) e investigador del grupo GridTICs. Cuenta con experiencia en investigación en el área de redes y laboratorios remotos (Tobar, 2020). Está dando sus primeros pasos en el área de ciencia de datos y cuenta con experiencia en programación en Python.

En el campo de la formación de recursos humanos, se espera:

1. Iniciar en investigación a profesionales que estén interesados en postularse a diferentes becas doctorales, en el marco del proyecto en referencia.

2. Iniciar la investigación a alumnos de la UTN que deseen sumarse al proyecto como becarios.

3. Iniciar en investigación a alumnos de los últimos años de carrera que deseen desarrollar su proyecto final de carrera en el marco de este proyecto.

5. BIBLIOGRAFÍA

(Fildes et al., 2019) Fildes R., Ma S., Kolassa S. Retail forecasting: Research and practice. *International Journal of Forecasting* (2019).

(Kaggle, 2022) Kaggle Inc. <https://www.kaggle.com/>

(Hyndman, 2020) Hyndman R.J. A brief history of forecasting competitions *International Journal of Forecasting*, 36 (2020), pp. 7-14.

(Lainder, 2022) A. David Lainder, Russell D. Wolfinger, Forecasting with gradient boosted trees: augmentation, tuning, and cross-

validation strategies: Winning solution to the M5 Uncertainty competition, *International Journal of Forecasting*, 2022, ISSN 0169-2070,

<https://doi.org/10.1016/j.ijforecast.2021.12.003>.

(Makridakis, 2021) Spyros Makridakis, Evangelos Spiliotis, Vassilios Assimakopoulos, Zhi Chen, Anil Gaba, Ilia Tsetlin, Robert L. Winkler, The M5 uncertainty competition: Results, findings and conclusions, *International Journal of Forecasting*, 2021, ISSN 0169-2070, <https://doi.org/10.1016/j.ijforecast.2021.10.009>.

(Makridakis, 2022) Spyros Makridakis, Evangelos Spiliotis, Vassilios Assimakopoulos, M5 accuracy competition: Results, findings, and conclusions, *International Journal of Forecasting*, 2022, ISSN 0169-2070, <https://doi.org/10.1016/j.ijforecast.2021.11.013>.

(Mercado Libre, 2021) Mercado Libre. Meli Data Challenge 2021 <https://ml-challenge.mercadolibre.com/>

(Moreyra, 2021) Matías Moreyra. Meli Data Challenge 2021 - First Place Solution. <https://github.com/MatiMoreyra/meli-data-challenge-2021>

(Seeger et al. 2016) Seeger, M., Salinas, D., and Flunkert, V. (2016). Bayesian intermittent demand forecasting for large inventories. *NeurIPS Proceedings*.

(Semenoglou et al., 2021) Semenoglou A.-A., Spiliotis E., Makridakis S., Assimakopoulos V. Investigating the accuracy of cross-learning time series forecasting methods. *International Journal of Forecasting*, 37 (2021), pp. 1072-1084.

(Gonzalez and Catania, 2019) Rodrigo Gonzalez and Carlos A. Catania (2019). Time-delayed multiple linear regression for de-noising MEMS inertial sensors. *Computers & Electrical Engineering*, Volume 76, 2019, Pages 1-12, ISSN 0045-7906.

(Diedrichs, 2018) A. L. Diedrichs, F. Bromberg, D. Dujovne, K. Brun-Laguna and T. Watteyne, "Prediction of Frost Events Using Machine Learning and IoT Sensing

Devices," in IEEE Internet of Things Journal, vol. 5, no. 6, pp. 4589-4597, Dec. 2018. DOI: 10.1109/JIOT.2018.2867333

(Watteyne, 2018) Using SmartMesh IP in Smart Agriculture and Smart Building Deployments: it Just Works. Keoma Brun-Laguna, Co-Authors: Thomas Watteyne, Ana Laura Diedrichs, Diego Dujovne, Carlos Taffernaberry, Remy Leone, Xavier Vilajosana. Computer Communications journal, elsevier, 2018. DOI <https://doi.org/10.1016/j.comcom.2018.03.010>.

(Tobar, 2020) Sebastián Tobar, Carlos Taffernaberry, Cristian Pérez, Gustavo Mercado. Remote Embedded Systems Laboratory. V Congreso bienal de la sección Argentina del IEEE (IEEE ARGENCON 2020), 2020.

ÍNDICES DE DOMINIO PARA IDENTIFICACIÓN EFICIENTE MEDIANTE CODIGO DE IRIS

Juan J. Aguirre, Carlos E. Alvez, Graciela E. Etchart

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Av. Tavella 1424, Concordia, Entre Ríos - CP 3200

{juan.aguirre, carlos.alvez, graciela.etchart}@uner.edu.ar

RESUMEN

Las tecnologías biométricas representan un componente integral en sistemas de gestión de la identidad y de control de acceso; sin embargo, aún quedan muchos puntos a resolver respecto. Uno de estos puntos es la eficiencia en la búsqueda cuando se gestionan grandes cantidades de registros biométricos. Este problema se acrecienta cuando se busca identificar un individuo mediante un rasgo biométrico dentro de n individuos posibles, dado que no los sistemas de bases de datos actuales no cuentan con métodos de indexación para datos complejos, como lo son comúnmente las codificaciones de los rasgos biométricos. Existen varias alternativas para mejorar la performance en las búsquedas de códigos biométricos como ser, la implementación de diferentes tipos de índices o formas de segmentar la base de datos. En este trabajo se trabaja específicamente con el rasgo del iris humano, se busca mejorar el acceso a través de la adecuada representación del código del iris y estructuras de acceso eficientes. El principal objetivo de este trabajo es la definición de métodos de acceso eficientes para la identificación de iris utilizando índices de dominio.

Palabras clave: índices de dominio, biometría, iris, eficiencia.

CONTEXTO

Este trabajo se da en el marco del Proyecto PID 07/G058 “*Modelos de Machine Learning para la mejora de la precisión, seguridad y eficiencia en la gestión de datos biométricos*”.

El objetivo general del presente proyecto es mejorar la precisión, seguridad de datos y eficiencia en la gestión de datos biométricos. En este trabajo en particular mejorar la eficiencia del proceso de identificación biométrica con el código del iris, para lo cual se analizan diferentes métodos de indexación base de datos para mejorar el proceso de identificación mediante.

1. INTRODUCCIÓN

La creciente preocupación por la seguridad en las organizaciones, ha llevado al desarrollo de diferentes tecnologías para garantizar la autenticación y/o identificación de personas. Muchos de estos desarrollos están relacionados a los sistemas biométricos. La mayor importancia en el uso de esta tecnología radica en su aplicación como mecanismo de control que, a través de una serie de medidas de características específicas, permite el reconocimiento de personas. El funcionamiento general de los sistemas biométricos es el siguiente: por un lado, se

cuenta con una base de datos con los rasgos biométricos codificados (plantillas o template). Por otro lado, se debe contar con un sensor, que permite capturar el rasgo indicado a la persona que desea acceder (a un sistema o espacio físico). El sistema de control biométrico debe permitir comparar el rasgo capturado por el sensor, con los rasgos almacenados en la base de datos. Esta comparación puede ser: uno a uno en caso de que la persona se identifique (proceso de verificación) o, uno a n en caso de que se quiera identificar a la persona (proceso de identificación).

Los principales desafíos que aún se presentan en los sistemas biométricos son la precisión en la comparación, seguridad de datos y eficiencia. Este trabajo se enfoca en este último punto.

Respecto a los problemas relacionados a la eficiencia, se presentan principalmente por la necesidad de utilizar un método de identificación (comparación de 1 a n), con n muy grande. Para esto, las soluciones pueden consistir en métodos de indexación y/o segmentación de la base de datos. Este tema ha sido poco abordado en la literatura. Un trabajo relacionado a la indexación de datos biométricos se presenta en [1], donde se utiliza k-tree para indexación de plantillas de iris. Existen otras estructuras de índice basadas en métricas que podrían utilizarse para realizar índices de dominio en área de biometría, como M-tree [2] y Metric Index [3]. Una alternativa a la indexación para mejorar la eficiencia, es utilizar técnicas de ML (clustering) para la segmentación de la base de datos biométrica. De esta manera, se puede reducir el espacio de búsqueda mejorando los tiempos en el proceso de identificación. Un caso se puede ver en [4], donde se utiliza K-means para segmentar una base de datos de huellas dactilares. Básicamente, segmenta la base de datos por similitud. Para la búsqueda de una huella particular, comienza por el segmento con mayor similitud, luego por el siguiente, y así sucesivamente. Estas dos alternativas para mejorar la eficiencia no

son incompatibles. De hecho, se puede dividir la base de datos en segmentos e indexar cada segmento como se presenta en [5], donde se utiliza k-means para segmentar la base de datos y Hashing Multi-Index (MIH) para indexar los segmentos.

Otro punto relacionado con la eficiencia, aunque fuera del alcance de este proyecto, es cómo afecta a la eficiencia el tratamiento de la seguridad de los datos biométricos. En este sentido en [6], utilizan técnicas de extracción de características y de ML para mejorar el rendimiento, pero también reducen el esfuerzo del adversario para romper el sistema utilizando datos falsificados.

El propósito de esta investigación, es analizar cuales de estas alternativas se adecua mejor para mejorar la eficiencia en el proceso de identificación mediante el iris.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Una de las líneas más importantes, consideradas en el proyecto, consiste en la construcción de modelos para la representación de datos relacionados con los datos biométricos, de manera que los mismos puedan compararse y recuperarse de manera eficiente. En este trabajo, nos enfocamos en datos del iris humano.

Cuando se intenta verificar en un sistema biométrico si una persona es quien dice ser, no hay mayores problemas de eficiencia dado que sólo se debe comparar la plantilla almacenada de la persona que dice ser, con la plantilla generada a partir de los datos capturados por el sensor (búsqueda 1 a 1). En este caso, los índices tradicionales como hash tree o b-tree pueden utilizarse sin problemas, dado que se puede acceder directamente al registro por medio de un valor clave o identificación de usuario, y luego comparar el código del iris (iriscode) almacenado en la base de datos con el iriscode de la persona que se intenta autenticar.

Sin embargo, cuando se utilizan métodos de identificación (buscar quien es la persona que intenta acceder), esta comparación es uno a **n**. En este último caso es cuando se presenta el mayor problema de performance, dado que, sin una estructura adicional, se debe comparar el iriscode de la persona que se desea autenticar, respecto a todos los iriscodes almacenados en la base de datos. Además, en este caso, los índices tradicionales provistos por los sistemas de gestión de bases de datos (SGBD) no se adecuan para la búsqueda, dado que las estructuras del iriscode por lo general, son arreglos de valores.

En la propuesta presentada aquí, los iriscodes se almacenan en arreglos binarios utilizando el SGBD Postgresql, y como solución al problema se plantea implementar un índice de dominio basado en distancia.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Para lograr los objetivos del presente trabajo, se llevaron adelante las siguientes tareas:

1. Desarrollo de un sistema que permita generar código de iris en forma masiva, utilizando bases de datos de iris reconocidas, como: CASIA [7], BMBD [8] y IITD [9].
2. Búsqueda de representación de códigos de iris (iriscode), basado en curvas de Gabor en base de datos.
3. Implementación de métodos de comparación de iriscode.
4. Análisis de diferentes opciones de índices de dominio para iriscode.
5. Implementación de un índice de dominio para ser utilizado en la identificación mediante iriscode.
6. Evaluación de la performance del índice.

Las tres primeras tareas se han completado. El primer punto se logró con un desarrollo propio, utilizando librerías open source [10][11], el cual genera

códigos de iris binarios. Para el segundo punto, se utilizó una forma eficiente de almacenamiento en Postgresql mediante un el tipo *bit varying* [12] que permite trabajar con arreglos binarios y que cuenta con un conjunto de operadores bit a bit (AND, XOR, OR, NOT), lo cual es muy importante para el punto 3, dado que los métodos se pudieron implementar como funciones en PL/PgSQL (lenguaje de procedimiento de Postgresql). Esto permite que el cálculo de distancias entre iriscodes sea eficiente. Respecto a la actividad 4, para la selección de un método de indexación para el proceso de identificación, se tuvo en cuenta los siguientes criterios:

- Soporte de datos de alta dimensionalidad: condición fundamental para la estructura de código del iris propuesta.
- Eficiencia en búsqueda por rango: debe ser una estructura donde, preferentemente, evite que se superpongan los rangos.
- Capacidad de reorganización: para evitar reorganizar todo el índice ante nuevas inserciones.
- Orientado a la eficiencia en el almacenamiento secundario: para mejorar la eficiencia, es importante para minimizar el acceso a disco.

Luego del análisis de diferentes tipos índices de dominio bajo los criterios citados, se llegó a la conclusión de que el índice que mejor se adapta para el problema de identificación planteado es el M-Tree [13].

Actualmente se está trabajando en la implementación de un índice de dominio M-Tree adaptado la propuesta realizada en [14][15] para el código del iris. Para la implementación se está utilizando framework GiST (árbol de búsqueda generalizado) de Postgresql, por diferentes razones: Postgresql es una base de datos open source, por lo que no es necesario el pago de licencia; en laboratorio ya se cuenta con una extensión de Postgresql para gestionar datos biométricos [16] y, por

sobre todo, este framework brinda todo lo necesario para la implementación de índices métricos.

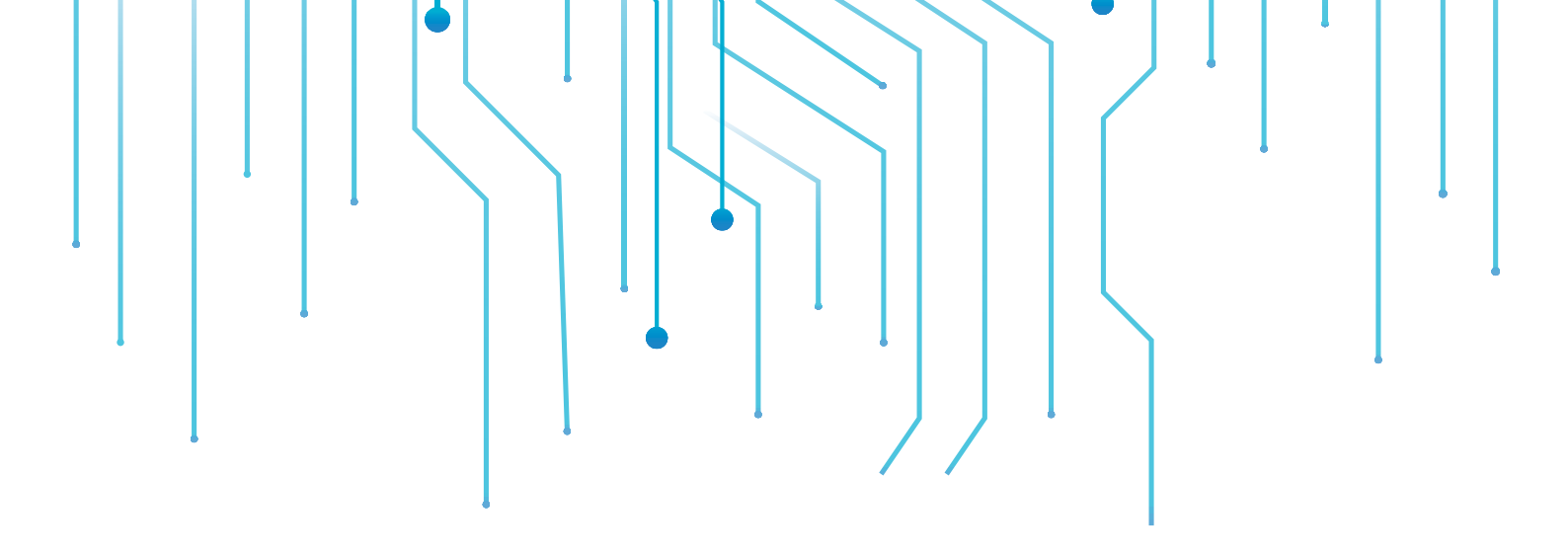
Se espera que, con esta implementación, las búsquedas por similitud se mantengan en un tiempo relativamente bajo aun cuando el número de iriscodes en la base de datos crezca. Para esto, una vez culminada la implementación del índice, se desarrollarán los ensayos para la evaluación de la performance del índice propuesto.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se finalizaron dos tesis de Maestría en Sistemas de Información y se están desarrollando otras tres tesis. Además, un integrante alumno se encuentra realizando el trabajo final de carrera de Licenciatura en Sistemas.

5. BIBLIOGRAFÍA

1. Alvez C., Miranda E., Etchart G. y Ruiz S. (2018). Efficient Iris Recognition Management in Object-Related Databases, JCS&T, vol. 18, no. 02, p. e12.
2. Zezula P., Amato G., Dohnal V., and Batko M. (2006). Similarity Search: The Metric Space Approach, Advances in Database Systems. vol. 32, 220 p. Springer, January 2006.
3. Novak D., Batko M., Zezula P. (2011). Metric Index: An efficient and scalable solution for precise and approximate similarity search, Information Systems, Volume 36, Issue 4, Pages 721-733.
4. Manhua Liu, Xudong Jiang, Alex Chichung Kot (2007). Efficient fingerprint search based on database clustering, Pattern Recognition, Volume 40, Issue 6, Pages 1793-1803.
5. Bai C., Wang W., Zhao T., Li M. (2018). Fast exact fingerprint indexing based on Compact Binary Minutia Cylinder Codes, Neurocomputing, Volume 275, Pages 1711-1724.
6. Sadeghi K., Banerjee A., Javad S., Gupta K.S. (2017). Performance and Security Strength Trade-Off in Machine Learning Based Biometric Authentication Systems. 1045-1048. 10.1109/ICMLA.2017.00-12.
7. CASIA: <http://english.ia.cas.cn/>
8. BMBD: <https://arxiv.org/abs/2111.08702>
9. IITD: <https://home.iitd.ac.in/>
10. Lee Y, Micheals RJ, Filliben JJ, Phillips PJ. VASIR: An Open-Source Research Platform for Advanced Iris Recognition Technologies. J Res Natl Inst Stand Technol. 2013 Apr 23;118:218-59.
11. OpenCV. <https://docs.opencv.org/>.
12. Bit String Types. <https://www.postgresql.org/docs/current/datatype-bit.html>
13. P. Ciaccia, M. Patella, F. Rabitti, P. Zezula. Indexing Metric Spaces with M-tree. Proc. Quinto Convegno Nazionale Seb. pp.67-86. 1997.
14. I. DONKÓ, J. M. SZALAI-GINDL, G. GOMBOS, and A. KISS. An implementation of the m-tree index structure for postgresql using gist. In 2019 IEEE 15th International Scientific Conference on Informatics, page 6, Poprad, Slovakia, nov 2019.
15. G. GOMBOS, J. M. SZALAI-GINDL, I. DONKÓ and A. KISS: "Towards on experimental comparison of the M-tree index structure with BK-tree and VP-tree", 2020, Acta Electrotechnica et Informatica. Vol. 20, No. 2, 2020.
16. C. Alvez, G. Etchart, S. Ruiz, E. Miranda and J. Aguirre. Extensión de una base de datos Objeto-Relacional para el soporte de datos de iris. XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). Universidad Nacional de La Plata-Argentina, 2017.



CGIV

Computación Gráfica, Imágenes y Visualización

TESTING DE VISUALIZACIONES

Martín Larrea, Matías Selzer, Dana K. Urribarri,

Clara Casalini, M. Luján Ganuza, Silvia M. Castro

Laboratorio de I+D en Visualización y Computación Gráfica

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)

Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

Laboratorio de I+D en Visualización y Computación Gráfica,

(UNS-CIC Prov. de Buenos Aires)

{mll, matias.selzer, dku, mcca, mlg, smc}@cs.uns.edu.ar

RESUMEN

El uso de visualizaciones de información ha crecido significativamente gracias a áreas como Big Data, Ciencia de Datos e Industria 4.0. Su aplicación a alcanzado sectores cada vez más críticos. En este contexto, las implementaciones de dichas visualizaciones deben cumplir con estándares de calidad más altos. Para garantizar dicha calidad, esta línea de investigación busca generar nuevas metodologías y herramientas desarrolladas desde una perspectiva de ingeniería de software, especialmente desde el área de verificación y validación de software. Todo nuestro desarrollo es libre y de código abierto, y por esto creemos que puede ser la base fundamental para futuros desarrollos que expandan su funcionalidad y dominios de aplicación.

Palabras Clave: Verificación y Validación, Visualización de Datos, Análisis Visual de Datos

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La toma de decisiones se ha transformado en una herramienta clave en toda organización e individuo, pasando de ser un proceso basado en la experiencia y la intuición a uno cada vez más establecido en el análisis de datos. Los procedimientos usados para tomar decisiones, en especial en base a grandes conjuntos de datos, son más sofisticados y requieren software especializado. Un tipo de software especializado es el de representaciones visuales. Las representaciones visuales son de suma importancia en la actualidad ya que permiten la exploración efectiva de un conjunto de datos y facilitan la tarea de identificar patrones y extraer conclusiones. Los sistemas de representaciones visuales no sólo ofrecen una vista de los datos, sino también proveen interacciones mediante las cuales el usuario puede manifestar sus necesidades y a partir de las cuales obtener la perspectiva deseada del conjunto de datos. Cada día más decisiones son tomadas en base al análisis visual mediante el uso de representaciones visuales. No sólo hay un aumento cuantitativo sino también cualitativo, las decisiones son más críticas y con mayor impacto en la sociedad, el medio ambiente y los individuos. En este contexto resulta fundamental desarrollar nuevas y mejores metodologías y herramientas que permitan asegurarnos el correcto funcionamiento de las

representaciones visuales y sus interacciones ([1]).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Basándonos en el estado del arte de las representaciones visuales de grandes conjuntos de datos ([2]), y teniendo en cuenta sus aplicaciones ([3]), los estados intermedios de los datos ([4]) y las posibles interacciones ([5]), es posible definir e implementar metodologías y herramientas que permitan detectar errores y/o problemas en los software de representaciones visuales de grandes conjuntos de datos. La principal línea de investigación y desarrollo consiste en el estudio y desarrollo de metodologías para la validación y verificación de representaciones visuales de grandes conjuntos de datos. De esta línea principal se desprenden sublíneas relativas al diseño y desarrollo de herramientas que asistan en el desarrollo de representaciones visuales de grandes conjuntos de datos para asegurar la calidad del producto final, así también como herramientas de testeo para software de propósito general ([13, 14, 15, 16]). Estas herramientas buscarán eliminar o reducir la incidencia de errores en los sistemas de representación visual. Esta línea de investigación expande nuestro trabajo sobre el área de Visualización ([6, 7, 8, 9, 10, 17, 18]).

3. RESULTADOS OBTENIDOS Y ESPERADOS

Sobre los ejes presentados se han obtenido resultados parciales. Desde el punto de vista de la teoría de base se ha comenzado a realizar un relevamiento de las diferentes metodologías disponibles en el ámbito de la Verificación y Validación del Software que sean aplicables a un dominio tan particular como el de la Visualización de Información. En lo referido al desarrollo de nuevas metodologías y herramientas se ha comenzado a publicar los

resultados obtenidos hasta el momento. Las siguientes publicaciones muestran los avances realizados:

- Schiaffino, M., Larrea, M. L., Ganuza, M. L., & Urribarri, D. K. (2022). A testing tool for information visualizations based on user interactions. *Journal of Computer Science & Technology*, 22. [19]
- Larrea, M. L., & Urribarri, D. K. (2021). Expanding the scope of a testing framework for Industry 4.0. In XXVII Congreso Argentino de Ciencias de la Computación (CACIC) [20]

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo y concluidas relaciones con las líneas de investigación presentadas, así también como un proyecto de grupo de investigación.

Tesis en Desarrollo: “Frontend Testing para Aplicaciones Móviles”, tesis de grado para la Ingeniería en Sistemas de Información. Alumno: Fausto Genovese Arrúe. Director: Martín Larrea. Codirector: Dana Urribarri.

Tesis Concluida: “Pruebas de front-end en aplicaciones Reactjs con Cypress”, tesis de grado para la Ingeniería en Computación. Alumno: Gianni Lucas Cerdá. Director: Martín Larrea.

Tesis Concluida: “Testing de Accesibilidad en Formularios Web. SIU-Kolla como Caso de Estudio.”, tesis de grado para la Ingeniería en Sistemas de Información. Alumna: Constanza Giorgetti. Director: Martín Larrea, Colaborador: Matías Selzer.

Tesis Concluida: “Metodologías de Análisis Estático en el Contexto de la Enseñanza de la Programación.”, proyecto final de carrera para la Ingeniería en Computación. Alumno: Tomas Miguel. Director: Martín Larrea.

Tesis Concluida: “Verificación y Validación de Software en el Entorno Symphony”, proyecto final de carrera para la Ingeniería en Sistemas de Computación”. Alumno: Tomas Miguel. Director: Martín Larrea

Tesis Concluida: “Alternativas a Selenium como Herramienta de Automatización para el Testing en la Web”, tesis de grado para la Licenciatura en Ciencias de la Computación. Alumno: Rosario Segovia Gonzalez. Director: Martín Larrea.

Tesis Concluida: “Testing Funcional en el Contexto de la Telemedicina”, proyecto final de carrera para la Ingeniería en Sistemas de Información. Alumno: Johanna Escudero. Director: Martín Larrea.

Proyecto: PGI 2019 “VERIFICACIÓN Y VALIDACIÓN DE REPRESENTACIONES VISUALES Y SUS INTERACCIONES”. Director: Martín Larrea, Miembros: Matías Selzer, Dana K. Urribarri, M. Clara Casalini. Proyecto acreditado para incentivos.

5. BIBLIOGRAFÍA

[1] Kirby, Robert M., and Cláudio T. Silva. “The need for verifiable visualization”. Computer Graphics and Applications, IEEE 28.5, 78-83, 2008.

[2] Rees, D., and R. S. Laramée. "A Survey of Information Visualization Books." Computer Graphics Forum. 2019.

[3] Börner, Katy, Andreas Bueckle, and Michael Ginda. "Data visualization literacy:

Definitions, conceptual frameworks, exercises, and assessments." Proceedings of the National Academy of Sciences 116.6 (2019): 1857-1864.

[4] Hagen, Hans. "DFU, Volume 2, Scientific Visualization: Interactions, Features, Metaphors, Complete Volume." Dagstuhl Follow-Ups. Vol. 2. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012.

[5] Ganuza, M. L. "Thesis Overview: Interactions in Visualization." Journal of Computer Science & Technology 18.2 (2018): 178-179.

[6] Castro, Silvia Mabel, et al. "Métricas, técnicas y semántica para la visualización de datos." XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).. 2018.

[7] Urribarri, Dana K., and Silvia M. Castro. "Prediction of data visibility in two-dimensional scatterplots." Information Visualization 16.2: 113-125. 2017

[8] Ganuza, María Luján, et al. "Visualización de Datos aplicada a resolver problemas en Ciencias Geológicas." XVI Workshop de Investigadores en Ciencias de la Computación. 2014.

[9] Larrea, Martín Leonardo, et al. "Ontologías y Semántica en el Proceso de Visualización." XVI Workshop de Investigadores en Ciencias de la Computación. 2014.

[10] Selzer, Matías Nicolás, et al. "Modelos de interacción y aplicaciones en realidad virtual mediante dispositivos móviles." XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires). 2017.

[11] Larrea, Martín Leonardo. "Black-Box Testing Technique for Information Visualization. Sequencing Constraints with Low-Level Interactions." *Journal of Computer Science & Technology* 17. 2017.

[12] Silva, J.I.R., Larrea, M.: White-box testing framework for object-oriented programming based on message sequence specification. In: XXIV Congreso Argentino de Ciencias de la Computación (Tandil, 2018). 2018.

[13] Larrea, M. & Urribarri, D.: Increasing Confidence in Industry 4.0 through new Software Verification and Validation Techniques. *International Conference on Production Research ICPR Americas*, 10. In press. 2020.

[14] Larrea, M. & Urribarri, D.: TAPIR: An Object-Oriented Programming Testing Framework based on Message Sequence Specification with Aspect-Oriented Programming. XXVI Congreso Argentino de Ciencias de la Computación. In press. 2020.

[15] Gonzalez, R. S. S., Urribarri, D. K., & Larrea, M. L. (2022). Automation Tools for Web Testing. Beyond Selenium. *Memorias de las JAIIO*, 8(3), 49-62.

[16] Larrea, M. "El análisis estático como herramienta de evaluación en cátedras con proyectos de programación." XVI Congreso de Tecnología en Educación & Educación en Tecnología-TE&ET 2021 (La Plata, 10 y 11 de junio de 2021). 2021.

[17] Wahler, S., Larrea, M., & Martinez, D. (2022). Una Herramienta de Análisis y Visualización de Redes Sociales para la

Identificación de Bandas Delictivas. *Memorias de las JAIIO*, 8(12), 23-36.

[18] Urribarri, D. K., & Larrea, M. L. (2022). A visualization technique to assist in the comparison of large meteorological datasets. *Computers & Graphics*, 104, 1-10.

[19] Schiaffino, M., Larrea, M. L., Ganuza, M. L., & Urribarri, D. K. (2022). A testing tool for information visualizations based on user interactions. *Journal of Computer Science & Technology*, 22.

[20] Larrea, M. L., & Urribarri, D. K. (2021). Expanding the scope of a testing framework for Industry 4.0. In XXVII Congreso Argentino de Ciencias de la Computación (CACIC)(Modalidad virtual, 4 al 8 de octubre de 2021).

PACS Cloud Open Source

Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito 9 de Julio 22, Chilecito, La Rioja, Argentina ¹

Lic. Ariel Quiroga Marin¹, Ing. Enrique N. Martinez¹, Ing. Elvio Sigampa¹, Dra. Laura Romero¹ {cquiroga, emartinez, esigampa, lromero}@undec.edu.ar, Palacios Cristian¹, Franco Padilla¹, Blanco Francisco¹ {cristianpalacios95, padilla.joni, franciscoblanco1907}@gmail.com

Hospital Eleazar Herrera Motta, Caveno 500, Chilecito, La Rioja²

Dr. Rivero José², Lic. Aguilar Norma², Dr. Omar Suarez², Leonel Alvarez², Ruben Quiroga² {riverotattoo, aguilarnorma, suarezomar, rubenhquiroga, leonelundec}@gmail.com

Resumen

Este trabajo se enfoca en un PACS CLOUD (Plataforma en línea para compartir y archivar imágenes médicas y documentos clínicos) que nos permite el almacenamiento, la búsqueda y la recuperación de imágenes médicas de una institución pública de salud “Hospital Eleazar Herrera Motta de Chilecito”, administrando las transferencias dentro del entorno de la nube y entre la nube y los dispositivos de acceso. Con PACS CLOUD se reduce la infrautilización del hardware, proporcionando una configuración flexible de máquinas virtuales, redes internas y externas eficientes y servicios adicionales como copia de seguridad y base de datos. Se estudia los problemas relacionados con la integridad de los datos y la seguridad. Los datos provienen de imágenes, en sus modalidades de tomografía computarizada (TC), resonancia magnética (MR) y radiografía computarizada (CR), los que se integran perfectamente en el entorno, a través del estándar de comunicación HL7, contiene los servicios DICOM (Imágenes digitales y comunicaciones en medicina), DICOM WEB, extensión de DICOM pensado en Internet y se puede utilizar como un visor WEB independiente o integrado en los sistemas dcm4chee Archive, Conquest PACS, ClearCanvas PACS,

adaptándose al sistema PACS del Hospital fácilmente en el flujo de trabajo en un entorno libre, compatible con código abierto, con una interfaz de integración flexible y abierta para conectarse a sistemas de almacenamiento en la nube (PACS en la nube).

PALABRAS CLAVE: Integrating Health Care Enterprise (IHE), Digital Imaging and Communications in Medicine (DICOM), PACS DICOM, Health Level7 (HL7), DCM4CHEE, Docker, arc-light, DICOMWeb (RestFul) y OHIF Viewer Oviyam, Cloud Server.

Líneas de I+D+I

- Computación gráfica, Imágenes y Visualización
- Arquitectura
- Redes y Sistemas Operativos
- Base de Datos
- Cloud Computing
- Minería de Datos

Contexto

DICOM es un estándar ampliamente adoptado para imágenes médicas, pero no es una solución integral para problemas de integración y transacciones en todos los componentes de

atención médica, para ello se necesita un sistema para captura y conversión y almacenamiento a formato DICOM de las imágenes de Diagnostico Medico de los distintos estudios que se realizan en el nosocomio “Hospital Eleazar Herrera Motta ”(tomografías, rayos x, ecografías, endoscopias etc.). El acceso al sistema será realizado mediante la intranet o internet, es decir, que se pueda realizar desde cualquier computador conectado en el hospital a la red interna o a su vez conectado a través de la red internet a la nube (PACS Cloud) accesible desde cualquier parte del mundo, estas imágenes quedan almacenadas en un servidor local. Los sistemas PACS basados en la nube nos permiten alojar fuera del sitio original (Hospital, clínica), y nos brinda seguridad y acceso a los datos de imágenes médicas desde cualquier lugar con conexión a Internet. Con un sistema PACS basado en la nube, la red informática del Hospital se amplía considerablemente, por lo que el alcance de los dispositivos que pueden actuar como estaciones de visualización es mucho más amplio. Cualquier dispositivo compatible con HTML5 puede acceder a un visor DICOM basado en la nube en línea.

Introducción

Hoy en día, los archivos DICOM incluyen también datos de imágenes y señales generados por áreas diversas como hematología, laringología, dermatología, cardiología, patología, oncología, urología y odontología. En la actualidad está en curso en el campo de las señales neurofisiológicas que tienen la capacidad para transferir conjuntos de datos de DICOM en 3D de alta resolución para aplicaciones neurológicas, además DICOM también se puede utilizar para almacenar formatos patentados, como los escáneres de microscopio. En las pruebas que

se realizaran, trataremos de mostrar las diferencias de comportamiento de escalado del servidor Dcm4chee open source implementado con sus componentes open source arc-light, las imágenes quedan totalmente vinculadas a la historia de Salud del Paciente, son accedidas netamente a través de DICOM Web (RestFul) y visualizadas a través de OHIF Viewer Oviyam, cumpliendo con los estándares DICOM server, HL7 server, Web-based interface. Se trabajará con los diferentes patrones de recuperación para evaluar el rendimiento en diferentes escenarios, sus conexiones paralelas y en el rendimiento general del servidor en la nube, además se prevé un entorno de producción que almacena más de 200 TB de datos de imágenes en un lapso de 4 años. El desafío para el PACS Cloud proviene de la cantidad de datos que se generan y archivan. Mantener y resguardar copias de imágenes u otros datos dentro del sistema es uno de los requisitos fundamentales para salvaguardar un gran volumen de datos médicos que están regulados por las leyes nacionales, esto requiere actualizaciones constantes de hardware y software, para garantizar una copia de seguridad a largo plazo es necesario proteger mediante sistemas de refuerzo en diferentes ubicaciones físicas o seleccionando almacenamiento externo. Solo el archivo local está disponible sin conexión a Internet, migrar un archivo en la nube es un desafío debido al gran volumen de datos. A pesar de que DICOM es un estándar ampliamente adoptado para imágenes médicas, no es fácil brindar una solución integral para problemas de integración y transacciones en la nube. La expansión de modalidades digitales provocó una explosión en la producción de imágenes médicas, aumentando la necesidad de tener repositorios más grandes para permitir toda esta cantidad de datos con alta disponibilidad y rendimiento. Las bases de datos NoSQL han ido

reemplazando a las bases de datos relacionales en algunos escenarios, debido a su escalabilidad horizontal y a su flexibilidad para adaptarse a requerimientos dinámicos.

El sistema PACS se integra perfectamente en el entorno hospitalario, a través del estándar de comunicación HL7, contiene los servicios DICOM y HL7 que se requieren para proporcionar almacenamiento, recuperación y flujo de trabajo a un entorno libre. El PACS DCM4CHE de código abierto con licencia LGPL permite la rápida adopción de imágenes de un archivo DICOM, su arquitectura y las características abarcan los estándares [DICOM, Health Level 7 (HL7), etc.], incluyendo un registro integral de transacciones DICOM, donde el archivo viene en una versión precompilada (binaria) diseñada para dos bases de datos de código abierto y fácilmente disponibles, MySQL y POSTGRES. Usaremos Software Open Source para automatizar el deploy de aplicaciones en contenedores, esto nos permitirá ahorrar el proceso de compilación y tener un modo de actualización simple (Docker). La plataforma en la nube es responsable del archivo de imágenes (a largo plazo), el flujo de datos y la administración de back-end. Funciona en modo multiusuario, con acceso DICOM de la interfaz local y sirve como puerta de enlace a los servicios en la nube (fig. 1).

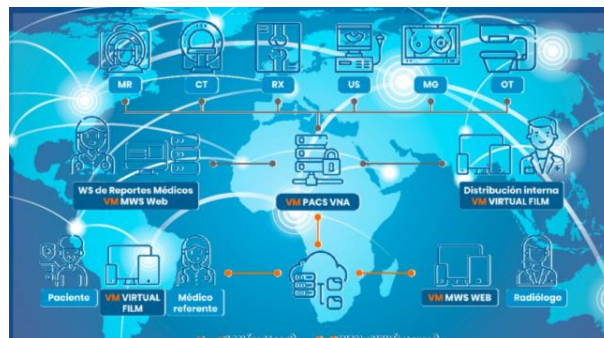


Figura 1. Diagrama Cloud PACS

Formación de Recursos Humanos

El equipo de trabajo está formado por docentes de las carreras Ingeniería en Sistemas, Licenciatura en Sistemas, Ing. Mecatrónica y Lic. En Enfermería de la UNdeC (acreditadas por CONEAU), dos de los cuales están realizando de la especialización de Big Data y Postgrado de Informática y Redes Mikrotik. También participan profesionales del Hospital Eleazar Herrera Motta de Chilecito y alumnos avanzados de grado, realizando su trabajo de tesina final en esta línea de I+D. Los integrantes son docentes de las asignaturas Sistemas Operativos I y II, Modelos y Simulación, Inteligencia Artificial, Taller Hardware e informática, Electrónica, Enfermería Hospitalaria. Estas asignaturas contemplan la aprobación mediante la participación en proyectos de investigación, por lo que pueden surgir nuevos trabajos en esta línea.

RESULTADOS ESPERADOS

Para que se pueda llevar a cabo sistemas de gestión y almacenamiento de imágenes médicas de manera eficiente en la nube (PACS Cloud), es necesario que los estudios cumplan con los estándares internacionales que garanticen su

reproducción y comunicación como son DICOM y HL7, Digital Imaging and Communications in Medicine y Health Level 7 respectivamente, estos formatos alimentan a los sistemas PACS para que finalmente haya una gestión y resguardo de estos estudios, lo que nos lleva a agregar mayor grado de seguridad (Redes y servidores) utilizando Fail2Ban, Firewall, OpenVPN con tunneling, encriptado de datos a nivel SO, además de incluir un esquema de backup de la información (archivos y DB), usar un dominio DNS Server y proxy reverso (NGINX).

Estos estándares de imágenes digitales de diversos campos, como la radiología, la cardiología, tomografía, endoscopia, etc., se combinan armoniosamente con la información de los sistemas de información clínica y otras bases de datos. Los usuarios pueden acceder fácilmente desde cualquier dispositivo y lugar. Esta solución PACS Cloud de código abierto nos dará la capacidad de integrarse con equipos de proveedores líderes en equipamiento de Imágenes DICOM. La solución Cloud Open Source propuesta resuelve el problema que tiene el Hospital de Chilecito y clínicas de Chilecito, que se interconectarán todos los servicios esenciales permitiendo un enlace práctico y dinámico para el flujo de información en la nube y brindar diagnósticos más precisos en menos tiempo, donde más de una persona puede acceder y ver registros médicos a la vez agilizando la colaboración entre los profesionales de la salud. Las características del PACS Cloud nos otorga una gran escalabilidad, fiabilidad de los datos y reducción de la infraestructura hardware. PACS Cloud permitira almacenar los datos

médicos del Hospital Eleazar Herrera Motta en una nube privada, cumpliendo con leyes de protección de datos. Gracias al desarrollo de software, las imágenes DICOM podrán ser procesadas y exportadas a archivos del tipo STL, PLY, OBJ entre otros, los cuales son archivos de intercambio para que puedan ser manipulados virtualmente en un software de diseño o llevados a una impresora 3D para la generación de un modelo físico. Al obtener un modelo 3D a partir de las imágenes DICOM se abren las posibilidades de análisis, diseño, diagnóstico, planeación e, inclusive, se puede explicar al paciente la anomalía y procedimiento a efectuar por medio del modelo virtual o el modelo físico impreso.

Referencias

- [1] Fennell, N.; Ralston, MD; Coleman, R. M. PACS y otros sistemas de gestión de imágenes. En *Informática de Imágenes Prácticas: Fundamentos y Aplicaciones para Imágenes Médicas*; Branstetter, BF, IV, Ed.; Springer: Nueva York, NY, EE. UU., 2021; págs. 131–142.
- [2] M. Lutz and D. Ascher, *Learning Python*, Second Edition. Sebastopol, CA: O'Reilly Media, Inc., 2003.
- [3] Huang, H. *PACS e informática de imágenes: principios básicos y aplicaciones*; John Wiley & Sons, Inc.: Hoboken, Nueva Jersey, EE. UU., 2010.
- [4] Linthicum, D.S. Cloud-Native Applications and Cloud Migration: The Good, the Bad, and the Points Between. *IEEE Cloud Comput.* 2017, 4, 12–14. [<https://ieeexplore.ieee.org/document/6864305>]
- [5] Warnock, M.J.; Toland, C.; Evans, D.; Wallace, B.; Nagy, P. Benefits of Using the DCM4CHE DICOM Archive. *J. Digit. Imaging* 2007, 20, 125–129. [<https://link.springer.com/article/10.1007/s10>]

278-007-9064-1]

- [6] Lebre, R.; Pinho, E.; Silva, J.M.; Costa, C. Dicoogle Framework for Medical Imaging Teaching and Research. In Proceedings of the 2020 IEEE Symposium on Computers and Communications (ISCC), Rennes, France, 7–10 July 2020; pp. 1–7. [<https://ieeexplore.ieee.org/document/9219545/>]
- [7] Costa, C.; Ferreira, C.; Bastiao, L.; Ribeiro, L.; Silva, A.; Oliveira, J.L. Dicoogle: un PACS punto a punto de código abierto. *J. Dígito. Imágenes* 2011, 24, 848–856. [<https://link.springer.com/article/10.1007/s10278-010-9347-9>]
- [8] Langer, S.G. Challenges for Data Storage in Medical Imaging Research. *J. Digit. Imaging* 2011, 24, 203–207.
- [9] Savaris, A.; Gimenes Marquez Filho, A.A.; Rodrigues Pires de Mello, R.; Colonetti, G.B.; Von Wangenheim, A.; Krechel, D. Integrating a PACS Network to a Statewide Telemedicine System: A Case Study of the Santa Catarina State Integrated Telemedicine and Telehealth System. In Proceedings of the 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), Thessaloniki, Greece, 22–24 June 2017; pp. 356–357. [<https://ieeexplore.ieee.org/document/8104218/>]
- [10] DICOM® Publications and DICOM Web™ Publications are published by and copyright owned by the National Electrical Manufacturers Association <https://www.dicomstandard.org/current>.

ANÁLISIS VISUAL DE MOVIMIENTOS OCULARES EN ENTORNOS DINÁMICOS

Leandro Luque^{1,2}, M. Luján Ganuza^{1,2,3}, Silvia M. Castro^{1,2,3} y Osvaldo E. Agamennoni³
{leandro.luque, mlg, smc}@cs.uns.edu.ar, oagamennoni@uns.edu.ar

¹ Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab)
(UNS-CIC Prov. de Buenos Aires)

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur
(DCIC-UNS)

² Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

³ Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC)

Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur
(DIEC-UNS)

RESUMEN

La utilización de herramientas basadas en eye tracking ha experimentado un importante crecimiento en los últimos años. La gran variedad de dispositivos que permiten registrar y almacenar como las personas se desempeñan en la realización de diversas tareas, permite contar con una gran cantidad de datos que son precisos de analizar.

Si bien los comienzos de los primeros experimentos estaban restringidos a entornos controlados donde las condiciones se mantenían inalterables, nuevos trabajos y experimentos proponen comprender cómo los participantes se desenvuelven en entornos dinámicos donde las condiciones de los estímulos presentados varían con el tiempo.

En la presente línea de investigación, se propone la investigación y el desarrollo de nuevas técnicas de visualización como complemento a los métodos tradicionales, que permitan dar soporte a los nuevos requerimientos de análisis dinámico, atendiendo no sólo al aspecto espacio-temporal de los datos, sino a las condiciones particulares del estímulo empleado durante el experimento.

Palabras claves: *Análisis Visual de Datos, Eye Tracking, Visualización de Datos*

Multidimensionales, Visualización de Datos Espacio-Temporales, Estímulos Dinámicos

CONTEXTO

El presente trabajo se realiza en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación (DCIC), de la Universidad Nacional del Sur (UNS) y en colaboración con el Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC, UNS-CIC). Los trabajos realizados en esta línea implican la participación de docentes investigadores y becarios doctorales.

1. INTRODUCCIÓN

El desarrollo del eye tracking como complemento de los modelos tradicionales para analizar comportamiento en diferentes actividades como *gaming*, *web browsing* y experimentos psicológicos, ha tenido un rol preponderante en el último tiempo. Particularmente para tareas como lectura o búsqueda de objetos en imágenes, el empleo de datos provenientes de un eye tracker permite comprender con mayor detalle el desempeño cognitivo de las personas. El uso de indicadores propios de la actividad ocular

permite cuantificar y analizar estados cognitivos de alto orden. Si bien estos tipos de estímulos han probado su utilidad en el desarrollo de experimentos, están limitados en cuanto a su naturaleza estática. En la actualidad y con la posibilidad de poder contar con aparatos registradores más precisos y de diferentes resoluciones, es posible incorporar estímulos dinámicos que permitan análisis más complejos y que se ajusten a la realidad cotidiana que experimentan las personas.

Lo expresado anteriormente presenta un nuevo reto en cuanto a cómo presentar los datos de manera visual para facilitar la exploración y análisis de los mismos. Al incorporar la naturaleza dinámica de un estímulo, como puede ser un video, una imagen que cambia sus condiciones iniciales en el tiempo o un experimento de lectura que modifica su contenido en función de lo que comprende el lector, es necesario el desarrollo de técnicas de visualización que permitan soportar estas variables. Es necesario no solo representar el aspecto espacio-temporal de los datos, sino que además se debe incorporar como estos se vinculan con este tipo de estímulo.

El desarrollo de nuevas técnicas de visualización supone el análisis de las distintas alternativas para explorar las propiedades dinámicas de los estímulos, y poder así detectar cuáles de los comportamientos de los participantes son focales (propios de una parte específica del estímulo que tiene sus propias características) y cuáles son globales.

Determinar la mejor manera visual de presentar estos datos es un tema bajo estudio dentro la comunidad de visualización, en la cual se han presentado opciones que tratan de abarcar todo el análisis en una sola técnica, en tanto que otros enfoques proponen las bondades del uso de vistas coordinadas, en

donde los usuarios tienen la posibilidad de interactuar con distintos aspectos de los datos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los nuevos paradigmas basados en datos de eye tracking introducen enfoques en donde la información más relevante se obtiene de la retroalimentación constante entre el sujeto y el estímulo al cual se lo somete.

Los diferentes tipos de indicadores oculares [11, 16] que se pueden recolectar de un experimento de eye tracking están relacionados con el movimiento (fijaciones y sacadas) y la dinámica pupilar. El primer tipo involucra a los dos movimientos más empleados en la literatura para el análisis de tareas que se desarrollan usando el registro ocular. Mientras que las fijaciones se refieren a las regiones puntuales en donde el sujeto focaliza su atención, las sacadas se definen a los distintos cambios de atención. La combinación de ambos movimientos da a lugar al concepto de *scanpath*, que resume cómo se compone la exploración visual del estímulo que el sujeto está analizando. En el caso particular de las fijaciones, estas son ampliamente estudiadas en la actualidad por tener propiedades relevantes que emergen de sus sub-movimientos como el drift, tremor y microsacadas [6, 13], que permiten un análisis mucho más detallado de cómo se moldea la atención de los sujetos. Los comportamientos como la exploración de detalles espaciales finos, implicaciones en la visión foveal y el tipo de estrategia de ocular empleada están fuertemente influenciados por este tipo de movimientos.

Por otro lado, el uso de los datos de la dinámica pupilar permite incorporar un aspecto que está fuertemente vinculado al desempeño cognitivo [10] de los sujetos bajo estudio. Históricamente, este tipo de indicador ha representado un gran desafío para los

expertos debido a que se encuentra afectado directamente por las condiciones ambientales sobre las cuales se lleva a cabo el experimento y por la fisiología del sujeto (si sufre algún tipo de desorden que no permite obtener datos útiles y/o representativos). Para ello han surgido una serie de nuevos indicadores [8, 9, 12, 16, 17] que están siendo empleados en trabajos recientes como soluciones parciales (o totales) a dichos problemas, pero que se encuentran acotados a ciertos tipos de tareas.

Si bien la rama estadística [19, 22] es la más ampliamente difundida en la comunidad de eye tracking, existe una necesidad creciente de emplear alternativas visuales que permitan explorar la gran cantidad de datos que generan estos experimentos como alternativa a los modelos más tradicionales. La naturaleza masiva, multivariada y espacio-temporal de los datos conlleva una gran cantidad de interrelaciones que no siempre son posibles de comprender de manera sencilla empleando modelos estadísticos.

Las nuevas técnicas visuales [3, 4, 5, 14, 15] permiten sintetizar una gran cantidad de información en poco espacio y ofrecen alternativas para aquellas situaciones donde se requiere mayor detalle de los datos. Sin embargo, estas técnicas están diseñadas para cubrir un amplio espectro de experimentos y no alcanzan a cubrir las necesidades especiales que son requeridas por el experto del dominio. Por otro lado, dichas representaciones visuales no incorporan ninguna información del estímulo dinámico que permita analizar cómo el mismo está afectando el desempeño del sujeto analizado y su utilidad en la comprensión de la atención humana.

Esta propuesta busca investigar cómo los movimientos oculares y la dinámica ocular extraídos de tareas que involucran estímulos dinámicos pueden ser empleados como

indicadores de actividad cognitiva. Adicionalmente, el desarrollo de técnicas de visualización permitirá conocer cómo es la relación de estos movimientos oculares con las propiedades dinámicas del estímulo. Mediante estas técnicas se buscará explotar el nuevo campo de *visual data science* como soporte adicional para los expertos relacionados con el eye-tracking.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Esta línea de investigación tendrá como finalidad el diseño y la implementación de técnicas de visualización que permita el análisis global de los datos obtenidos de experimentos desarrollados con estímulos dinámicos, como es el caso de videos o tareas que se desarrollan en el exterior que son similares a la realidad (conducción, interacción [1, 2, 20] entre personas, etc.). Las técnicas ad-hoc propuestas buscarán poner en contexto los movimientos oculares de los participantes con el estímulo bajo estudio, para comprender cómo se interrelacionan entre sí. Si bien los indicadores tradicionales, como el tiempo de fijación, cantidad de regresiones, y dilatación de la pupila resultan útiles para los análisis actuales, también es necesario el uso y desarrollo de nuevos indicadores que permitan captar nuevos patrones en los datos que resultan relevantes para los expertos que llevan a cabo estos experimentos de eye-tracking.

4. FORMACION DE RECURSOS HUMANOS

En cuanto a la formación de recursos humanos se detallan las tesis en desarrollo relacionadas con las líneas de investigación presentadas, así como también los proyectos de investigación y los becarios vinculados.

Tesis Completada: “Extracción y caracterización de microsacádicos”, Juan Arriola. Tesis de Doctorado en Matemática. Directora: Dra. Liliana Raquel Castro. Codirección: Dra. Marcela Patricia Álvarez.

Tesis en Desarrollo: “Análisis visual de datos provenientes de registradores oculares”, Leandro Luque. Tesis de Doctorado en Ciencias de la Computación. Directoras: Dra. Silvia Castro, Dra. M. Luján Ganuza.

Becario: Leandro Luque. Plan de trabajo: “Análisis Visual de Datos provenientes de Registradores de Movimientos Oculares”. Beca doctoral CONICET 2018, adjudicada a partir de abril de 2019.

Proyecto: PGI 24/N048, “Análisis Visual de Datos” (PGISeCyT-UNS, 2019-2023). Directora: Dra. Silvia Castro.

Proyecto:-PIBAA - CONICET (2872021010 0824CO) “Análisis Visual de Datos Multidimensionales sin Pérdida de Información”. Directora: Dra. M. Luján Ganuza.

Proyecto: PGI 24/K085, “Dinámica de los Sistemas Cognitivos” (PGISeCyT-UNS, 2019-2023). Director: Dr. Osvaldo Agamennoni, Co-Directora: Dra. Silvia Castro.

5. BIBLIOGRAFÍA

[1] Amati, M., McCarthy, C., Parmehr, E. G., & Sita, J. (2019). Combining eye-tracking data with an analysis of video content from free-viewing a video of a walk in an urban park environment. *JoVE (Journal of Visualized Experiments)*, (147), e58459.

[2] Aschwanden, D., Langer, N., & Allemand, M. (2019). Eye tracking in the wild: Piloting a

real-life assessment paradigm for older adults. *Journal of Eye Movement Research*, 12(1).

[3] Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2014, June). State-of-the-Art of Visualization for Eye Tracking Data. In *EuroVis (STARs)*.

[4] Blascheck, T., Raschke, M., & Ertl, T. (2013, August). Circular heat map transition diagram. In *Proceedings of the 2013 Conference on Eye Tracking South Africa* (pp. 58-61).

[5] Blascheck, T., Schweizer, M., Beck, F., & Ertl, T. (2017, June). Visual comparison of eye movement patterns. In *Computer Graphics Forum* (Vol. 36, No. 3, pp. 87-97).

[6] Bowers, N., Rucci, M., & Poletti, M. (2016). Microsaccades during reading. *Journal of Vision*, 16(12), 858-858

[7] Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, 10(10), 28-28.

[8] Duchowski, A. T., Krejtz, K., Krejtz, I., Biele, C., Niedzielska, A., Kiefer, P., ... & Giannopoulos, I. (2018, April). The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-13).

[9] Duchowski, A. T., Krejtz, K., Gehrer, N. A., Bafna, T., & Bækgaard, P. (2020, April). The low/high index of pupillary activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

- [10] Fernández, G., Buedo, P., Orozco, D., & Agamennoni, O. (2017). Eye Movement Behavior Analyses for Studying Cognitive Performance and Conversion to Pathologies. In *Psychiatry and Neuroscience Update-Vol. II* (pp. 281-292). Springer, Cham.
- [11] Holmqvist, K., & Andersson, R. (2017). Eye tracking: A comprehensive guide to methods. Paradigms and Measures.
- [12] Jayawardena, G. (2020, June). Raemap: Real-time advanced eye movements analysis pipeline. In *ACM Symposium on Eye Tracking Research and Applications* (pp. 1-4).
- [13] Krejtz, K., Żurawska, J., Duchowski, A. T., & Wichary, S. (2020). Pupillary and microsaccadic responses to cognitive effort and emotional arousal during complex decision making. *Journal of Eye Movement Research*, 13(5).
- [14] Kurzhals, K., Heimerl, F., & Weiskopf, D. (2014, March). ISeeCube: Visual analysis of gaze data for video. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 43-50).
- [15] Kurzhals, K., Hlawatsch, M., Heimerl, F., Burch, M., Ertl, T., & Weiskopf, D. (2015). Gaze stripes: Image-based visualization of eye tracking data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1), 1005-1014.
- [16] Mahanama, B., Jayawardana, Y., Rengarajan, S., Jayawardena, G., Chukoskie, L., Snider, J., & Jayarathna, S. (2022). Eye movement and pupil measures: A review. *Frontiers in Computer Science*, 3, 733531.
- [17] Marshall, S. P. (2002, September). The index of cognitive activity: Measuring cognitive workload. In *Proceedings of the IEEE 7th conference on Human Factors and Power Plants* (pp. 7-7). IEEE.
- [18] Noton, D., & Stark, L. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11(9), 929-938.
- [19] Onwuegbusi, T., Hermens, F., & Hogue, T. (2022). Data-driven group comparisons of eye fixations to dynamic stimuli. *Quarterly Journal of Experimental Psychology*, 75(6), 989-1003.
- [20] Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. *Scientific reports*, 8(1), 1-9.
- [21] Rucci, M., & Poletti, M. (2015). Control and functions of fixational eye movements. *Annual Review of Vision Science*, 1, 499-518.
- [22] Stadler, M., Doeblér, P., Mertins, B., & Delucchi, R. (2021). Statistical modeling of dynamic eye-tracking experiments: Relative importance of visual stimulus elements for gaze behavior in the multi-group case. *Behavior Research Methods*, 53, 2650-2667.

VISUALIZACIÓN COMPARATIVA: ESTUDIO Y DISEÑO DE VISUALIZACIONES ORIENTADAS A LA COMPARACIÓN DE CONJUNTO DE DATOS

Dana K. Urribarri, Martín Larrea, M. Luján Ganuza, Silvia M. Castro

Laboratorio de I+D en Visualización y Computación Gráfica

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)

Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

Laboratorio de I+D en Visualización y Computación Gráfica,

(UNS-CIC Prov. de Buenos Aires)

{dku, mll, mlg, smc}@cs.uns.edu.ar

RESUMEN

La comparación es una tarea importante en el análisis de datos, en la cual dos o más conjuntos de datos deben analizarse buscando similitudes o diferencias. La visualización, que cumple un rol importante en el proceso de entender, analizar y comunicar datos en general, también puede contribuir en la comparación de datos en particular. Estas comparaciones implican un desafío para la visualización, dado que se debe combinar tanto factores de los objetos individuales como de la relación entre ellos.

Mientras que la visualización de información tradicionalmente se ha centrado en examinar objetos o conjuntos de datos individuales, muchos sistemas han demostrado la importancia de desarrollar herramientas de visualización que soporten explícitamente la tarea de comparación. Si visualizar un único objeto o conjunto de datos complejo ya es una tarea desafiante, realizar una comparación suma desafíos nuevos.

La finalidad de este proyecto es proponer y desarrollar técnicas de visualización que faciliten el proceso de comparación de grandes conjuntos de datos y/o el proceso de comparación de múltiples conjuntos de datos.

Palabras Clave: *Visualización de información; Visualización comparativa; Grandes conjuntos de datos.*

CONTEXTO

Esta línea de investigación está inserta en el proyecto PGI 24/ZN39 financiado por la Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur, Argentina.

El proyecto se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) perteneciente al Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur y centro asociado CIC-PBA. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La comparación es una tarea importante en el análisis de datos, en la cual dos o más conjuntos de datos deben analizarse buscando similitudes o diferencias [1]. Aparece a lo largo de áreas de aplicación tan diversas como, por ejemplo, epidemiología [2], la oceanografía [3], la biología [4], la meteorología [5,6], y la ingeniería forestal [7] entre muchas otras, e involucra una amplia gama de tipos de datos, incluidos gráficos, tablas, superficies, series de tiempo, etc.

La visualización, que cumple un rol importante en el proceso de entender, analizar y

comunicar datos en general. En particular, la visualización de información puede dar soporte a la comparación de datos sin abordarla explícitamente. Sin embargo, las técnicas diseñadas explícitamente para facilitar la comparación son las que brindan mejor soporte al usuario [8].

La visualización comparativa plantea visualizar datos de dos o más fuentes diferentes para evidenciar similitudes y diferencias [9]. Pueden distinguirse tres tipos de objetivos en la comparación: comprender las similitudes y diferencias entre varios conjuntos de datos, comparar diferentes características de un conjunto de datos particular y comparar diferentes tipos de visualizaciones vinculadas. Además, pueden distinguirse tres grandes categorías de diseño de las visualizaciones comparativas [10]: la yuxtaposición que presenta los objetos a comparar por separado pero cercanos entre sí (en tiempo o en espacio), la superposición que presenta múltiples objetos en el mismo sistema de coordenadas (uno arriba del otro) y la representación explícita que codifica directamente las conexiones entre los objetos visuales. Finalmente, estas tres categorías pueden combinarse para formar categorías híbridas. Se han estudiado ventajas y desventajas de estas estrategias y se han esbozado guías de diseño [11].

Dentro de la visualización de datos se utilizan varias técnicas para la comparación, incluso las más tradicionales y simples [12]. La visualización *side-by-side* (una yuxtaposición espacial) es una de las estrategias más usadas. Si bien esta técnica es adecuada para la comparación, genera una alta carga cognitiva en el usuario y no escala bien ni con el tamaño ni con la cantidad de conjuntos de datos a comparar. El uso de esta estrategia para visualizar varios conjuntos de datos se denomina *small multiples* [13]. La superposición, a su vez, presenta problemas como la oclusión y la distorsión. Dentro de esta clasificación, los diagramas de espagueti son

muy utilizados aunque estos no escalan bien ni con la dimensionalidad ni con la cantidad de conjuntos a comparar. En este último caso, la acumulación de líneas en el área de visualización genera problemas de oclusión e interferencia en la información. Las visualizaciones comparativas 3D, considerando todas sus ventajas y desventajas, también son ampliamente utilizadas y necesarias [14].

El diseño de visualizaciones comparativas comparte los mismos desafíos que la visualización en general, añadidos aquellos inherentes a la propia tarea. Dado que la tarea de comparación puede cambiar drásticamente los objetivos de la visualización, las estrategias aprendidas de la visualización en general no siempre son apropiadas para este caso; por ejemplo, se ha estudiado el impacto de la tarea de comparación en la percepción [15,16,17]. Un buen diseño debe considerar el compromiso entre la complejidad del diseño y la eficiencia en la realización de la tarea [10], sin dejar de ser expresivo (i.e. reflejar los datos de manera óptima), efectivo (i.e. soportar la tarea de manera óptima) y apropiado (i.e. adaptarse al tipo de dato a visualizar). Los desafíos de diseño que la visualización comparativa comparte con la visualización en general son la escalabilidad visual, la incertidumbre de los datos y la percepción humana. Por otro lado, es importante contar con herramientas para medir y evaluar la calidad de las visualizaciones.

La escalabilidad visual [18] es la capacidad de las técnicas de visualización para mostrar eficazmente grandes conjuntos de datos, en términos de número o dimensión de los elementos individuales de estos. Y, dado que los problemas de comparación escalan no solo con la complejidad de los conjuntos de datos (cantidad de elemento y dimensionalidad) sino también con el número de conjuntos a comparar, la escalabilidad representa un desafío importante en la visualización comparativa [19]. Tres factores pueden afectar

a la escalabilidad visual: la cantidad de datos o su complejidad, la cantidad de conjuntos de datos a comparar, y la cantidad o la complejidad de las relaciones existentes entre los conjuntos.

La mayoría de los trabajos se centran en comparaciones 1 a 1, por lo que en estos casos la cantidad de datos o la complejidad de los datos de cada conjunto es el factor determinante en la escalabilidad visual. Por otro lado, tanto realizar comparaciones de varios conjuntos de datos a partir de estrategias 1 a 1 como encontrar representaciones visuales adecuadas para hacerlo no es sencillo, independientemente de la cantidad de datos en los conjuntos.

Otro gran desafío en visualización es la incertidumbre de los datos [20,21]. A pesar de la incertidumbre asociada a casi todos los conjuntos de datos, la visualización de información a menudo ha asumido que la información es en su mayoría precisa. Sin embargo, ignorar la incertidumbre asociada a los datos puede ser engañoso para los analistas y llevar a decisiones erróneas. La incertidumbre es un concepto complicado que tiene múltiples facetas y en la literatura existen múltiples definiciones. En general, hace referencia a que la información puede ser incompleta, inconsistente o errónea y, por lo tanto, el resultado final no es fiable; además, también pueden influir la falta de claridad, precisión o conocimiento de los usuarios. En el contexto de la visualización comparativa, la incertidumbre de los datos puede influenciar el proceso de comparación.

La medición es una parte integral de la ciencia moderna, ya que proporciona los medios fundamentales para la evaluación, la comparación y la predicción [22]. Las medidas pueden ser cuantitativas o cualitativas, y uno de sus principales objetivos es proporcionar medios objetivos para razonar sobre las visualizaciones y su eficacia [23,24,25].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación se centran en proponer y desarrollar técnicas de visualización que faciliten el proceso de comparación de grandes conjuntos de datos y/o el proceso de comparación de múltiples conjuntos de datos. En particular, se continuará trabajando en diseño de visualizaciones para la comparación de datos temporales multidimensionales [26,27,28,29]. Paralelamente, para las técnicas de visualización comparativa, se trabaja en el diseño de medidas objetivas para de evaluación de las técnicas, en el estudio de diferentes estrategias de visualización comparativa para diferentes tipos de datos, en el impacto de la cantidad de datos o la cantidad de conjuntos de datos a comparar en la escalabilidad visual de las técnicas y en la capacidad de las técnicas para representar la incertidumbre en los datos y en la comparación.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Sobre los ejes presentados se han obtenido resultados parciales. Se ha trabajado en una técnica de visualización comparativa de series de tiempo y basada en Dynamic Time Warping. Esta técnica se ha empleado tanto en la comparación de datos provenientes de movimientos de karate [26,27] como en la comparación de datos meteorológicos [28,29]. En la continuación del proyecto se espera contar con la implementación de una versión que genérica de la técnica que nos permita explorar la viabilidad de esta técnica en la comparación de series de tiempo con datos de otros tipos, por ejemplo, categóricos.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo que concierne a la formación de recursos humanos, se incentiva la incorporación de alumnos de grado o de posgrado que deseen trabajar en temas relacionados.

5. BIBLIOGRAFÍA

- [1] Schulz, H., Nocke, T., Heitzler, M., & Schumann, H. (2013). A Design Space of Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2366–2375. [10.1109/TVCG.2013.120](https://doi.org/10.1109/TVCG.2013.120)
- [2] Kreuz, T. (2021). *Comparative visualization of epidemiological data during various stages of a pandemic*. <https://arxiv.org/abs/2102.11315>
- [3] Yano, M., Itoh, T., Tanaka, Y., Matsuoka, D., & Araki, F. (2020). A comparative visualization tool for ocean data analysis based on mode water regions. *Journal of Visualization*, 23(2), 313–329. [10.1007/s12650-020-00629-y](https://doi.org/10.1007/s12650-020-00629-y)
- [4] Nagpal, S., Bakshi, K. D., Kuntal, B. K., & Mande, S. S. (2020). NetConfer: A web application for comparative analysis of multiple biological networks. *BMC Biology*, 18(1), Article 1. [10.1186/s12915-020-00781-9](https://doi.org/10.1186/s12915-020-00781-9)
- [5] Dasgupta, A., Poco, J., Wei, Y., Cook, R., Bertini, E., & Silva, C. T. (2015). *Bridging Theory with Practice: An Exploratory Study of Visualization Use and Design for Climate Model Comparison*. *IEEE Transactions on Visualization and Computer Graphics*, 21(9), 996–1014. [10.1109/tvcg.2015.2413774](https://doi.org/10.1109/tvcg.2015.2413774)
- [6] Rautenhaus, M., Böttinger, M., Siemen, S., Hoffman, R., Kirby, R. M., Mirzargar, M., Röber, N., & Westermann, R. (2018). *Visualization in Meteorology—A Survey of Techniques and Tools for Data Analysis Tasks*. *IEEE Transactions on Visualization and Computer Graphics*, 24, 3268–3296. [10.1109/TVCG.2017.2779501](https://doi.org/10.1109/TVCG.2017.2779501)
- [7] Nam, J. W., McCullough, K., Tveite, J., Espinosa, M. M., Perry, C. H., Wilson, B. T., & Keefe, D. F. (2019). Worlds-in-Wedges: Combining Worlds-in-Miniature and Portals to Support Comparative Immersive Visualization of Forestry Data. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 747–755. [10.1109/VR.2019.8797871](https://doi.org/10.1109/VR.2019.8797871)
- [8] Gleicher, M. (2018). Considerations for Visualizing Comparison. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), 413–423. [10.1109/TVCG.2017.2744199](https://doi.org/10.1109/TVCG.2017.2744199)
- [9] Pagendarm, H.-G., & Post, F. H. (1995). Comparative Visualization: Approaches and Examples. *Visualization in Scientific Computing (H. Göbel, H. Müller and B. Urban)*, February, 95–108. <https://elib.dlr.de/37368/>
- [10] Gleicher, M., Albers, D., Walker, R., Jusufi, I., Hansen, C. D., & Roberts, J. C. (2011). Visual comparison for information visualization. *Information Visualization*, 10(4), 289–309. [10.1177/1473871611416549](https://doi.org/10.1177/1473871611416549)
- [11] LYi, S., Jo, J., & Seo, J. (2021). Comparative Layouts Revisited: Design Space, Guidelines, and Future Directions. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 1525–1535. [10.1109/tvcg.2020.3030419](https://doi.org/10.1109/tvcg.2020.3030419)
- [12] Srinivasan, A., Brehmer, M., Lee, B., & Drucker, S. M. (2018, January). What's the Difference? *ACM Conference on Human Factors in Computing Systems (CHI)*. [10.1145/3173574.3173878](https://doi.org/10.1145/3173574.3173878)
- [13] Fuchs, J., Fischer, F., Mansmann, F., Bertini, E., & Isenberg, P. (2013, April). *Evaluation of alternative glyph designs for time series data in a small multiple setting*. [10.1145/2470654.2466443](https://doi.org/10.1145/2470654.2466443)
- [14] Kim, K., Carlis, J. V., & Keefe, D. F. (2017). Comparison techniques utilized in spatial 3D and 4D data visualizations: A survey and future directions. *Computers & Graphics*, 67, 138–147. [10.1016/j.cag.2017.05.005](https://doi.org/10.1016/j.cag.2017.05.005)
- [15] Jardine, N., Ondov, B. D., Elmqvist, N., & Franconeri, S. (2020). The Perceptual

- Proxies of Visual Comparison. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 1012–1021. [10.1109/TVCG.2019.2934786](https://doi.org/10.1109/TVCG.2019.2934786)
- [16] Ondov, B., Jardine, N., Elmqvist, N., & Franconeri, S. (2019). Face to Face: Evaluating Visual Comparison. *IEEE Transactions on Visualization and Computer Graphics*, 25(1), 861–871. [10.1109/TVCG.2018.2864884](https://doi.org/10.1109/TVCG.2018.2864884)
- [17] Ondov, B. D. (2021). *Revealing Perceptual Proxies in Comparative Data Visualization*. PhD Thesis. [Digital Repository at the University of Maryland]. [10.13016/RYUV-KZ1E](https://doi.org/10.13016/RYUV-KZ1E)
- [18] Eick, S. G., & Karr, A. F. (2002). Visual Scalability. *Journal of Computational and Graphical Statistics*, 11(1), 22–43. [10.1198/106186002317375604](https://doi.org/10.1198/106186002317375604)
- [19] Schmidt, J. (2016). *Scalable Comparative Visualization. Visual Analysis of Local Features in Different Dataset Ensembles*. PhD Thesis. [Faculty of Informatics (TU Wien)]. <https://diglib.eg.org/handle/10.2312/2631090>
- [20] Bonneau, G. P., Hege, H. C., Johnson, C. R., Oliveira, M. M., Potter, K., Rheingans, P., & Schultz, T. (2014). Overview and state-of-the-art of uncertainty visualization. In *Mathematics and Visualization* (Vol. 37, pp. 3–27). [10.1007/978-1-4471-6497-5_1](https://doi.org/10.1007/978-1-4471-6497-5_1)
- [21] Kamal, A., Dhakal, P., Javaid, A. Y., Devabhaktuni, V. K., Kaur, D., Zaiantz, J., & Marinier, R. (2021). Recent advances and challenges in uncertainty visualization: a survey. *Journal of Visualization*. [10.1007/s12650-021-00755-1](https://doi.org/10.1007/s12650-021-00755-1)
- [22] Bolte, F., & Bruckner, S. (2020). Measures in Visualization Space. In *Foundations of Data Visualization* (pp. 39–59). Springer International Publishing. [10.1007/978-3-030-34444-3_3](https://doi.org/10.1007/978-3-030-34444-3_3)
- [23] Urribarri, D. K., & Castro, S. M. (2017). Prediction of data visibility in two-dimensional scatterplots. *Information Visualization*, 16(2), 113–125. [10.1177/1473871616638892](https://doi.org/10.1177/1473871616638892)
- [24] Castro, S. M., Larrea, M. L., Urribarri, D. K., Ganuza, M. L., & Escarza, S. (2018). *Métricas, técnicas y semántica para la visualización de datos*. XX Workshop De Investigadores En Ciencias De La Computación (WICC 2018). <http://sedici.unlp.edu.ar/handle/10915/6745>
- [25] Urribarri, D., Larrea, M., & Castro, S. (2019). Automatic suggestions to improve the quality of scatterplots during its creation. *Journal of Computer Science and Technology*, 19(2), 100–109. [10.24215/16666038.19.e10](https://doi.org/10.24215/16666038.19.e10)
- [26] Urribarri, D. K., Larrea, M. L., Castro, S. M., & Puppo, E. (2019). Visualization to compare karate motion captures. *XXV Congreso Argentino De Ciencias De La Computación (CACIC 2019, Universidad Nacional De Río Cuarto)*. <http://sedici.unlp.edu.ar/handle/10915/90359>
- [27] Urribarri, D. K., Larrea, M. L., Castro, S. M., & Puppo, E. (2020). Overview+Detail Visual Comparison of Karate Motion Captures. In P. Pesado & M. Arroyo, *Communications in Computer and Information Science: Vol. 1184 CCIS* (pp. 139–154). Springer, Cham. [10.1007/978-3-030-48325-8_10](https://doi.org/10.1007/978-3-030-48325-8_10)
- [28] Larrea, M. L. & Urribarri, D. K. (2021). Visualization Technique for Comparison of Time-based Large Data Sets. *Conference on Cloud Computing Conference, Big Data & Emerging Topics (JCC-BD&ET), Springer International Publishing, 1444*, 179-187. [10.1007/978-3-030-84825-5_13](https://doi.org/10.1007/978-3-030-84825-5_13)
- [29] Urribarri, D. K. & Larrea, M. L. (2022). A visualization technique to assist in the comparison of large meteorological datasets *Computers & Graphics, Elsevier BV*. [10.1016/j.cag.2022.02.011](https://doi.org/10.1016/j.cag.2022.02.011)

TECNOLOGÍAS INMERSIVAS Y VISUALIZACIÓN SITUADA APLICADAS A GEOCIENCIAS

M. Luján Ganuza^{1,2}, Matías N. Selzer^{1,2}, Antonini Antonella S.^{1,2}, Luque, Leandro E.^{1,2}, Urribarri, Dana K.^{1,2}, Larrea, Martín L.^{1,2}, Ferracutti, Gabriela R.³, Asiain, Lucía³, Bjerg, Ernesto A.³, y Silvia M. Castro^{1,2}

¹Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (UNS-CIC)
Dpto. de Cs. e Ing. de la Computación, Universidad Nacional del Sur (DCIC-UNS)
{mlg, matias.selzer, antonella.antonini, leandro.luque, dku, mll, smc}@cs.uns.edu.ar

²Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

³INGEOSUR, Dpto. de Geología, Universidad Nacional del Sur
gferrac@uns.edu.ar; {lasiain, ebjerg}@ingeosur-conicet.gob.ar

RESUMEN

Las tecnologías inmersivas están avanzando a un ritmo vertiginoso, surgiendo cada día nuevos conceptos y aplicaciones. La visualización de datos, por otro lado, es una herramienta muy valiosa cuyo objetivo primario se basa en transmitir la información por medio de representaciones gráficas. Las tecnologías inmersivas pueden integrarse con la visualización de datos, ampliando su potencial, con el objetivo de transmitir información de manera que se potencie su análisis y entendimiento. Consideramos que la aplicación de tecnologías XR integradas en visualizaciones puede resultar de gran utilidad en diversas áreas, entre ellas las Geociencias. En este contexto se plantea como objetivo general el contribuir al diseño y desarrollo de técnicas inmersivas integradas en visualizaciones situadas interactivas en 3D que contribuyan a facilitar el trabajo de campo del geólogo.

Palabras claves: *Visualización Situada, Tecnologías Inmersivas, Tecnologías XR*

CONTEXTO

Este trabajo se realiza en estrecha colaboración con investigadores de centros de investigación de reconocido prestigio como el INGEOSUR-CONICET (Instituto Geológico

del Sur) y el Departamento de Geología de la Universidad Nacional del Sur, y el VyGLab (Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica) del Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS).

1. INTRODUCCIÓN

Las tecnologías inmersivas combinan la realidad con objetos virtuales (Realidad Aumentada y Realidad Mixta) o sumergen completamente al usuario en mundos totalmente virtuales (Realidad Virtual). La Realidad Extendida (XR, como se la conoce actualmente por sus siglas en inglés) es una tecnología novedosa que integra la Realidad Virtual (RV), la Realidad Aumentada (RA) y la Realidad Mixta (RM). A medida que las tecnologías digitales evolucionan hacia soluciones mejores, la RA, la RV y la RM se instalan lentamente en el centro de los negocios y de la vida cotidiana, provocando cambios sociales masivos. Aunque se suele pensar que estas tecnologías se utilizan principalmente para videojuegos, existen muchas aplicaciones en diversos campos que se benefician de ellas, como la educación [11, 12], el entrenamiento militar y médico [14, 15], el turismo, la salud mental y los deportes, entre otras [9,10].

La visualización de datos, por otro lado, es una herramienta muy valiosa siendo su objetivo primario el transmitir la información por medio de representaciones gráficas de datos para ayudar a los usuarios de la misma a generar insight, es decir, un entendimiento tanto de la información subyacente en éstos como de las posibles relaciones existentes entre éstos, para poder así llevar a cabo determinadas tareas de manera más eficiente. La XR puede integrarse con la visualización y al incorporar elementos virtuales a la realidad que percibimos (elementos sintéticos artificiales que se incorporan por un medio digital, como la computadora), o considerar una realidad totalmente virtual, brindará un abanico de nuevas formas de percibir e interactuar con estas realidades. Existen diversas alternativas tanto para integrar espacialmente los elementos virtuales a las escenas del mundo real o generar mundos totalmente virtuales, como para permitir la manipulación de los objetos virtuales en los ambientes de XR. Sin embargo, el éxito depende de cuán bien puedan comprenderse e integrarse las escenas generadas y cuán natural sea la interacción con ellas.

Consideramos que la aplicación de tecnologías XR integradas en visualizaciones puede resultar de gran utilidad en diversas áreas, entre ellas en las Geociencias. En particular, el trabajo de campo de los geólogos constituye un área de aplicación muy prometedora para la visualización de datos *in-situ* mediante la aplicación de tecnologías XR. Las actividades de campo resultan esenciales en lo concerniente a la búsqueda y explotación de recursos renovables y no renovables, emplazamientos de obras civiles e identificación y remediación de riesgos ambientales y uso de suelos entre muchas otras, involucrando cuidadosas observaciones y medidas de campo, recolección de muestras de rocas y almacenamiento de información

complementaria para su posterior análisis. Como soporte de estas tareas de campo, los geólogos utilizan simultáneamente diversos instrumentos de medición, además de imágenes de sensores remotos, mapas topográficos y mapas geológicos que deben referenciar y vincular al mundo real.

Adicionalmente, emerge el término de Visualización Situada para describir visualizaciones que son mostradas en el contexto en el cual son relevantes y el objetivo es que la combinación del mundo físico y la representación visual que se integra a éste tengan un significado integral [20]. En este caso, los datos visualizados deben estar claramente asociados con la ubicación física y/o los objetos reales y tener en cuenta al usuario y a las tareas a llevar a cabo por éste. La Visualización Situada también puede ser sumamente beneficiosa para el geólogo en el trabajo de campo. De este modo, podrá contar con herramientas informáticas que le brinden, en tiempo real e *in-situ*, no solo la documentación de soporte necesaria sino también los elementos que le permitan explorar nuevas alternativas para realizar el mapeo geológico incorporando los datos, tanto de laboratorio como los obtenidos con los instrumentos de medición propios e imprescindibles para el trabajo geológico (brújulas, clinómetros, magnetómetros, contadores Geiger, báculos, teodolitos, planchetas, GPS, etc.). La alternativa de integrar al ambiente real *in-situ* los mapas geológicos, los mapas topográficos, los datos recolectados durante una jornada y los obtenidos previamente, resultantes de los análisis de laboratorio, también permite reducir la cantidad de distintos tipos de soportes que contienen la información a utilizar (por ejemplo los esquemas en papel), facilitando el trabajo de campo y permitiendo una constante asociación de información 2D-3D.

2.LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación y desarrollo específicas planteadas son:

Diseño y desarrollo de herramientas de software para el campo de las Geociencias que integren técnicas inmersivas: Estas herramientas integrarán tecnologías de RA, RM y RV, que permitan interactuar con, y visualizar *in-situ* y de manera unificada, datos propios de la Geología.

La utilización de estas tecnologías de manera satisfactoria en exteriores presenta un gran desafío en cuanto a la forma de visualizar la información y a cómo interactuar con la misma. No es necesario contar con costosos dispositivos de hardware de alta gama para generar ambientes de XR en donde los usuarios puedan interactuar. Es posible desarrollar estas nuevas tecnologías a nivel nacional dados los recursos humanos y la tecnología actualmente disponibles.

Diseño y desarrollo de una visualización unificada de distintos estratos visuales que integre al ambiente real *in-situ* los datos multidimensionales que el geólogo utiliza (mapas geológicos, mapas topográficos, los datos recolectados durante una jornada, datos obtenidos de perforaciones y datos obtenidos previamente) en una aplicación de RA móvil. Esto permitirá reducir la cantidad de diferentes tipos de soportes que contienen la información a utilizar (por ejemplo los esquemas en papel), facilitando así la interpretación de los datos.

3.RESULTADOS OBTENIDOS Y ESPERADOS

Sobre los ejes presentados se han obtenido resultados parciales.

Con respecto al diseño y desarrollo de herramientas que integren tecnologías XR de bajo costo, se han desarrollado sistemas que

permiten a los usuarios explorar ambientes virtuales con una gran calidad visual, incluso en dispositivos móviles. Esas técnicas se complementaron con la detección de colisiones y la exploración de terrenos o ambientes con distintas alturas, lo que permite al usuario explorar e interactuar en ambientes más realistas [18, 19]. La exploración de mundos virtuales realistas y de gran tamaño presenta un desafío complejo en el área de tecnologías inmersivas de bajo costo; sin embargo, una gran ventaja es que esto puede utilizarse en diversas áreas, como por ejemplo la geología, en donde resulta muy útil la exploración grandes terrenos virtuales basados en terrenos reales.

Con respecto al diseño y desarrollo de una visualización unificada de distintos estratos visuales se ha trabajado de manera sostenida en el tiempo en temas relacionados con la Visualización de Datos aplicada a las Geociencias [4,5]. Se desarrollaron nuevas técnicas de visualización para datos geológicos aplicados a la mineralogía y a la prospección de recursos naturales [2, 6, 7, 8]. También se han realizado importantes avances en lo que respecta a visualización de datos multidimensionales en general [1, 3, 13, 16, 17], técnicas que son aplicables a datos provenientes del campo de las Geociencias.

4.FORMACION DE RECURSOS HUMANOS

A continuación, se detallan los proyectos de investigación, las tesis finalizadas y en desarrollo y las becas obtenidas vinculadas con las líneas de investigación presentadas.

Proyectos Vigentes:

-PGI 24/ZN38 “Tecnologías Inmersivas y Visualización Situada aplicadas a Geociencias”. Directora: Dra. M. Luján Ganuza.

-PIBAA - CONICET (2872021010 0824CO) “Análisis Visual de Datos Multidimensionales sin Pérdida de Información”. Directora: Dra. M. Luján Ganuza.

-PGI 24//N048 “Análisis Visual de Datos”. Directora: Silvia M. Castro.

-PICT-2017-1246 “Análisis Visual en Geociencias”. Directora: Silvia M. Castro.

Tesis Finalizada: “Métricas de Inmersión para Sistemas de Realidad Virtual”, tesis de Doctorado en Ciencias de la Computación. Alumno: Matías N. Selzer. Directores: Dra. Silvia M. Castro - Dr. Martín L. Larrea.

Tesis en Desarrollo:

- “Realidad Aumentada Móvil en Exteriores para Visualización de Datos Geológicos”, tesis de Doctorado en Ciencias de la Computación. Alumno: Juan Manuel Trippel Nagel. Directores: Dra. Silvia Castro - Dr. Ernesto Bjerg.

-“Análisis Visual de Datos Multidimensionales”, tesis de Doctorado en Ciencias de la Computación. Alumna: Antonella S. Antonini. Directora: Dra. Silvia Castro. Codirectora: Dra. M. Luján Ganuza.

Becas:

-Antonella S. Antonini. Título del plan propuesto: “Análisis Visual de Datos Multidimensionales”. Beca doctoral 2019 CONICET. Adjudicada a partir de abril de 2019 y por un término de 5 años. Directores: Dra. Silvia Castro - Dr. Ernesto Bjerg.

-Leandro Luque. Título del plan propuesto: “Análisis Visual de Datos provenientes de Registradores de Movimientos Oculares”. Beca doctoral 2019 CONICET. Adjudicada a partir de abril de 2019 y por un término de 5 años. Directora: Dra. Silvia Castro. Codirector: Dr. Osvaldo Agamennoni.

- Matías Selzer. Plan de trabajo: “Interacción y Navegación en Ambientes de Realidad Extendida Colaborativa”. Beca Posdoctoral CONICET 2020, adjudicada a partir de abril de 2021 y por un término de 2 años.

Proyecto Final de Carrera Concluido:

“Sistema Georeferenciado de Realidad Aumentada Colaborativo”, trabajo final de carrera Ingeniería en Sistemas de Computación. Alumno: Franco Raniolo. Directora: Dra. M. Luján Ganuza. Codirector: Dr. Matías N. Selzer.

5.BIBLIOGRAFÍA

- [1] Antonini, A. S., Ganuza, M. L. & Castro, S. M. (2022). VISUEL - A Web Dynamic Dashboard for Data Visualization. *Journal of Computer Science and Technology*, 22(1), e03.
- [2] Antonini, A. S., Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Matković, K., Gröller, E., Bjerg, E. A & Castro, S. M. (2021). Spinel web: an interactive web application for visualizing the chemical composition of spinel group minerals. *Earth Science Informatics*, 1-8.
- [3] Antonini, A. S., Luque, L., Ganuza, M. L. & Castro, S. M. (2022). Towards a Taxonomy for Non-Paired General Line Coordinates: A Comprehensive Survey. *International Journal of Data Science and Analytics*.
- [4] Ferracutti, G. R., Gargiulo, M. F., Ganuza, M. L., Bjerg, E. A., and Castro, S. M. (2015). Determination of the spinel group end members based on electron microprobe analyses. *Mineral. Petrol.* 109, 2 153-160.
- [5] Ganuza, M. L., Castro, S. M., Ferracutti, G., Bjerg, E. A., and Martig, S. (2012). Spinelviz: An interactive 3d application for visualizing spinel group minerals. *Comput. Geosci.* 48, 50–56.
- [6] Ganuza, M. L., Ferracutti, G., Gargiulo, M.F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. (2014). The spinel explorer - interactive visual analysis of spinel group minerals. *IEEE TVCG*. 20, 12, 1913–1922.
- [7] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E.,

- Matkovic, K. (2017). Interactive visual categorization of spinel-group minerals. Proc. of the Spring Conf. on Comput. Graph.
- [8] Ganuza, M. L., Gargiulo, M. F., Ferracutti, G., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. (2015). Interactive semi automatic categorization for spinel group minerals. In 2015 IEEE, VAST 2015, October 25-30, 2015, 197–198.
- [9] Grandi, J.G., Debarba, H.G., Nedel, L., and Maciel, A. (2017, May). Design and evaluation of a handheld-based 3d user interface for collaborative object manipulation. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 5881-5891).
- [10] Gutwin, C., Greenberg, S., (2002). A descriptive framework of workspace awareness for real-time groupware. Computer Supported Cooperative Work (CSCW) Vol. 11. n. 3-4 , pp. 411-446.
- [11] Huang, H.-M., Liaw, S.-S., Lai, C.-M., (2016). Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. Interactive Learning Environments, 24(1), 3-19.
- [12] Ibayashi, H., et al., (2015). Dollhouse VR: a multi-view, multi-user collaborative design workspace with VR technology. SIGGRAPH Asia 2015 Emerging Technologies. 1-2, 2015.
- [13] Kutak, D., Selzer, M. N., Byska, J., Ganuza, M. L., Barisic, I., Kozlikova, B. & Miao, H. (2021). Vivern A Virtual Environment for Multiscale Visualization and Modeling of DNA Nanostructures. IEEE Transactions on Visualization and Computer Graphics, 28(12), 4825-4838.
- [14] Lacoche, J., Pallamin, N., Boggini, T., and Royan, J. (2017). Collaborators awareness for user cohabitation in co-located collaborative virtual environments. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (pp. 1-9).
- [15] Le Chénéchal, M., Lacoche, J., Royan, J., Duval, T., Gouranton, V., and Arnaldi, B. (2016). When the giant meets the ant an asymmetric approach for collaborative and concurrent object manipulation in a multi-scale environment. In 2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE) (pp. 18-22).
- [16] Luque, L.E., Ganuza, M.L., Antonini, A.S. & Castro, S.M. (2021a). npGLC-Vis Library for Multidimensional Data Visualization. En Cloud Computing, Big Data & Emerging Topics: 9th Conference, JCC-BD&ET, La Plata, Argentina, June 22-25, 2021, Proceedings (Vol. 1444, p. 188).
- [17] Sabando, M. V., Ulbrich, P., Selzer, M., Byška, J., Mičan, J., Ponzoni, I., . . . Kozlíková, B. (2021). ChemVA: Interactive Visual Analysis of Chemical Compound Similarity in Virtual Screening. IEEE Transactions on Visualization and Computer Graphics, 27(2), 891-901.
- [18] Selzer, M. N., Ganuza, M. L., Urribarri, D. K., Larrea, M. L. & Castro, S. M. (2021). Stereoscopic Image-Based Rendering Technique for Low-Cost Virtual Reality. En P. Pesado & J. Eterovic (Eds.), Computer Science – CACIC 2020 (pp. 91-101). Cham: Springer International Publishing.
- [19] Selzer, M. N., Ganuza, M. L. & Castro, S. M. (2021). High Visual-Quality Scenes in Low-Cost Virtual Reality With Collisions and Irregular Surfaces. IEEE Computer Graphics and Applications, 41(4), 40-51.
- [20] White, S.; Feiner, S., Kopylec J., (2006). Virtual vouchers: Prototyping a mobile augmented reality user interface for botanical species identification. Proceedings of the 3D User Interfaces, p. 119–126.

VISUALIZACIÓN DE DATOS MULTIDIMENSIONALES PROCEDENTES DE LAS GEOCIENCIAS

Antonella S. Antonini¹, M. Luján Ganuza¹, Florencia Gargiulo², Gabriela Ferracutti², Ernesto A. Bjerg², Silvia M. Castro¹, Krešimir Matković³, Eduard Gröller⁴

⁽¹⁾ Laboratorio de I+D en Visualización y Computación Gráfica
Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica,
(UNS-CIC Prov. de Buenos Aires)

antonella.antonini@cs.uns.edu.ar; mlg@cs.uns.edu.ar; smc@cs.uns.edu.ar

⁽²⁾ INGEOSUR, Dpto. de Geología, Universidad Nacional del Sur
florenciagargiulo@gmail.com; gferrac@uns.edu.ar; ebjerg@ingeosur-conicet.gob.ar

⁽³⁾ VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH,
Donau-City-Strasse 11, 1220 Viena, Austria.
Matkovic@vrvis.at

⁽⁴⁾ Institute of Visual Computing & Human-Centered Technology
Favoritenstrasse 9-11 / E186, A-1040 Viena, Austria
groeller@cg.tuwien.ac.at

RESUMEN

En diversas disciplinas, incluidas las Ciencias Geológicas, a menudo se generan grandes volúmenes de datos multidimensionales.

Debido a la dimensionalidad y/o volumen de los datos, muchas técnicas tradicionales solo pueden producir una vista parcial de los mismos [13, 14, 15, 16].

Al aumentar considerablemente el número de variables e intentar representar dichos datos en un espacio muy limitado, la tarea del usuario de visualizar; interactuar y analizar posibles relaciones, patrones y tendencias, se vuelve cada vez más compleja.

Por ende, es valioso disponer de un conjunto de metáforas visuales y técnicas de visualización para analizar datos con estas características.

El objetivo general es promover el desarrollo de tecnologías y soluciones en torno al análisis visual de datos multidimensionales. En particular, nos enfocaremos en la

visualización de datos geológicos, con el fin de generar nuevos modelos, técnicas y herramientas para potenciar el análisis y la interpretación de los mismos.

Nos centraremos en la aplicación de estas técnicas al campo de la Geología, pero tenga en cuenta que estas técnicas se pueden utilizar en una variedad de disciplinas.

Palabras Clave: Visualización de Datos Multidimensionales; Análisis Visual; Visualización de Datos Geológicos

CONTEXTO

Esta línea de trabajo se lleva a cabo en el Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS). Trabajamos en estrecha colaboración con investigadores de reconocidos institutos de investigación nacionales e internacionales, incluidos el Institute of Visual Computing & Human-Centered Technology (TuWien,

Austria), VRVis (Center for Virtual Reality and Visualization, Austria), INGEOSUR-CONICET (Instituto Geológico del Sur) y el Departamento de Geología de la Universidad Nacional del Sur. Además, contamos con investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab).

1. INTRODUCCIÓN

La visualización de datos está siendo cada vez más utilizada en diversas áreas ante la necesidad de mostrar adecuadamente conjuntos de datos y la dificultad de extraer información relevante de ellos.

Dentro de la visualización de datos, existen múltiples técnicas que permiten un mejor entendimiento de la información. Sin embargo, la representación de datos multivariantes y multidimensionales requiere una visualización más compleja. Por ello, no es posible su representación de manera intuitiva utilizando una única gráfica. Estos datos se componen de multitud de variables que son importantes en conjunto y que necesitan una representación que ayude a analizarlos e inferir conclusiones sobre las posibles relaciones existentes entre ellos.

Nuestro objetivo y principal dificultad, es el estudio de técnicas que exploten al máximo las capacidades de percepción humana y permitan representar estos datos de una manera sencilla, evitando visualizaciones complejas y confusas.

En este contexto, exploramos nuevas técnicas multidimensionales y diversas visualizaciones correlacionadas que permiten al usuario el análisis visual de datos multidimensionales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En particular trabajamos con datos provenientes de las Ciencias Geológicas, pero

sin perder de vista el objetivo general de desarrollar técnicas de análisis de datos que sean útiles en diversas disciplinas.

Las diferentes líneas de investigación y desarrollo se orientan a la visualización de datos geoquímicos.

Un gran desafío para los geólogos consiste en caracterizar una región geológica en términos de su ambiente geotectónico. Para alcanzar este objetivo, se debe procesar y analizar una gran cantidad de información obtenida de rocas y minerales [10, 11, 12]. En particular, analizan los minerales que integran el grupo de los espinelos; un importante indicador petrogenético, que brinda información vital en lo referido al ambiente tectónico de las rocas presentes en un área, y proveen información valiosa acerca de los procesos que sufrieron las rocas que contienen a esos minerales.

Si bien existen técnicas de visualización que permiten analizar grandes conjuntos de datos multidimensionales provenientes de distintos contextos, en el análisis de datos geoquímicos sólo se trabaja con gráficos especiales que considera únicamente un conjunto reducido de estos atributos a la vez.

En este contexto, es importante diseñar y desarrollar métodos apropiados de visualización de datos multidimensionales que permitan a los geólogos considerar todos los datos obtenidos en el análisis.

Cabe destacar, que al trabajar juntamente con investigadores del Instituto Geológico del Sur (INGEOSUR-CONICET) y del Departamento de Geología de la Universidad Nacional del Sur, es posible validar los resultados obtenidos y verificar la utilidad de las nuevas técnicas propuestas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Desde hace tiempo trabajamos en diversas técnicas y herramientas para la visualización

de datos procedentes de las Geociencias [1, 2, 3, 4, 5].

En los últimos años, se implementó una herramienta web gratuita para la visualización de datos geológicos que integra la mayoría de los diagramas que utilizan los geólogos habitualmente para analizar las características químicas de los minerales e incorpora visualizaciones interactivas en 3D [6]. También proporciona vistas coordinadas e interacciones adecuadas para que los usuarios interactúen con sus conjuntos de datos.

En 2022 se desarrolló una aplicación de libre acceso [7] que integra un conjunto de técnicas de visualización coordinadas en un único dashboard. La herramienta puede ser utilizada en diversas disciplinas y brinda al usuario la posibilidad de visualizar un mismo conjunto de datos con diferentes técnicas de visualización y luego reutilizar la configuración de vistas con otros conjuntos de datos. VISUEL es totalmente interactivo y admite interacciones tradicionales como filtrado, selección, brushing, y zoom, entre otros. También permite al usuario configurar la representación visual de sus datos, eligiendo el color y la forma de las representaciones.

Además, en el último año llevamos a cabo una revisión sistemática de las técnicas de visualización presentes en la literatura y definimos una taxonomía para organizar las contribuciones recopiladas [8]. El trabajo realizado propone una clasificación completa de las técnicas y las interacciones asociadas; y ayuda a identificar áreas que no están suficientemente exploradas y son candidatas a ser investigadas.

También, desarrollamos una librería llamada np-GLCVis [9] que admite las Non-Paired General Line Coordinates (npGLC). Las Coordinadas de Líneas Generales (del inglés General Lines Coordinates o GLC) son representaciones visuales reversibles que

conservan información multidimensional para el descubrimiento del conocimiento. La librería presenta una colección de métodos de visualización e interacciones tradicionales, diseñada para experimentar con técnicas npGLC en el desarrollo de aplicaciones de visualización.

Actualmente, nos encontramos trabajando en el diseño e implementación de diversas herramientas de visualización que integran técnicas multidimensionales de análisis visual con modelos de Machine Learning. Nuestro objetivo es facilitar el trabajo de los geólogos en el proceso de categorización de una muestra de datos para determinar a qué ambiente tectónico corresponde.

4. FORMACIÓN DE RECURSOS HUMANOS

A continuación, se detallan los proyectos de investigación vinculados con las líneas de investigación presentadas y las becas adquiridas en centros de investigación científica.

Proyectos vigentes

- PGI SeCyT-UNS (24/N048) “*Análisis Visual de Datos*”. Directora: Silvia M. Castro.
- PGI SeCyT-UNS (24/ZN38) “*Tecnologías Inmersivas y Visualización Situada aplicadas a Geociencias*”. Directora: Dra. María Luján Ganuza.
- PIBAA - CONICET (2872021010 0824CO) “*Análisis Visual de Datos Multidimensionales sin Pérdida de Información*”. Directora: Dra. M. Luján Ganuza.
- PICT (2017-1246) “*Análisis Visual de Datos en Geociencias*”. Directora: Silvia M. Castro.

Proyectos finalizados

• PDTS, Proyecto de Desarrollo Tecnológico y Social: “*Desarrollo de un dashboard web integrado para el análisis visual de datos geoquímicos para la identificación de recursos naturales no renovables*”. Directora: Dra. Silvia Castro.

Beca Posdoctoral finalizada

M. Luján Ganuza. Título del plan propuesto: “*Análisis Visual de Datos Multidimensionales en Espacios Ad-Hoc*”. Beca Postdoctoral 2018 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET).

Beca Doctoral vigente

Antonella S. Antonini. Título del plan propuesto: “*Análisis Visual de Datos Multidimensionales*”. Beca doctoral 2019 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas (CONICET) de la República Argentina. Adjudicada a partir de abril de 2019 y por un término de 5 años.

5. BIBLIOGRAFÍA

- [1] Ferracutti, G. R., Gargiulo, M. F., Ganuza, M. L., Bjerg, E. A., and Castro, S. M. Determination of the spinel group end members based on electron microprobe analyses. *Mineral. Petrol.* 109, 2 (2015), 153-160.
- [2] Ganuza, M. L., Castro, S. M., Ferracutti, G., Bjerg, E. A., and Martig, S. Spinelviz: An interactive 3d application for visualizing spinel group minerals. *Comput. Geosci.* 48 (2012), 50–56.
- [3] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. The spinel explorer - interactive visual analysis of spinel group minerals. *IEEE TVCG.* 20, 12 (2014), 1913–1922.
- [4] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. Interactive visual categorization of spinel-group minerals. *Proc. of the Spring Conf. on Comput. Graph.* (2017).
- [5] Ganuza, M. L., Gargiulo, M. F., Ferracutti, G., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. Interactive semi automatic categorization for spinel group minerals. In *2015 IEEE Conference on Visual Analytics Science and Technology, VAST 2015, Chicago, IL, USA, October 25-30, 2015* (2015), 197–198.
- [6] Antonini, A. S., Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Matković, K., Gröller, E., Bjerg, E. A & Castro, S. M. (2021). Spinel web: an interactive web application for visualizing the chemical composition of spinel group minerals. *Earth Science Informatics*, 1-8.
- [7] Antonini, A. S., Ganuza, M. L., & Castro, S. M. (2022). VISUEL - A Web Dynamic Dashboard for Data Visualization. *Journal of Computer Science and Technology*, 22(1), e03.
- [8] Antonini, A.S., Luque, L., Ganuza, M.L. et al. Toward a taxonomy for 2D non-paired General Line Coordinates: a comprehensive survey. *Int J Data Sci Anal* (2022).
- [9] Luque, L.E., Ganuza, M.L., Antonini, A.S., Castro, S.M. (2021). npGLC-Vis Library for Multidimensional Data Visualization. In: Naiouf, M., Rucci, E., Chichizola, F., De Giusti, L. (eds) *Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2021. Communications in Computer and Information Science*, vol 1444. Springer, Cham.
- [10] Barnes, S. J., and Roeder, P. L. The range of spinel compositions in terrestrial mafic and ultramafic rocks. *J. Petrol.* 42, 12 (2001), 2279–2302.

[11] Lindsley, D.H. Oxide minerals: petrologic and magnetic significance. Review in Mineralogy 25 (1991).

[12] Roeder, P. L. Chromite; from the fiery rain of chondrules to the kilauea iki lava lake. The Canadian Mineralogist 32, 4 (1994), 729–746.

[13] Card, S. K., Mackinlay, J. D., and Shneiderman, B. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999.

[14] Spence, R. Information Visualization: Design for Interaction. 2007.

[15] Ware, C. Information Visualization: Perception for Design. Morgan Kaufmann, 2004.

[16] Ward, M. O., Grinstein, G., Keim, D. Interactive data visualization: foundations, techniques, and applications. CRC Press, 2010.

ANÁLISIS VISUAL GUIADO PARA EL ESTUDIO DE DATOS GEOQUÍMICOS

Antonini, Antonella S.^{1,2,3}, Luque, Leandro E.^{1,2,3}, Tanzola, Juan E.^{4,5}, Asiaín, Lucía^{4,5}, Ferracutti, Gabriela R.^{4,5}, Bjerg, Ernesto A.^{4,5}, Castro, Silvia M.^{1,2,3} y Ganuza, M. Luján^{1,2,3}

¹Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab) (UNS-CIC)

²Dpto. de Cs. e Ing. de la Computación, Universidad Nacional del Sur (DCIC-UNS)

{antonella.antonini, leandro.luque, smc, mlg}@cs.uns.edu.ar

³ICIC, Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)

⁴INGEOSUR, Instituto de Geología (UNS-CONICET)

⁵Dpto. de Geología, Universidad Nacional del Sur (DG-UNS)

gferrac@uns.edu.ar, {jetanzola,lasiain,ebjerg}@ingeosur-conicet.gob.ar

RESUMEN

Los datos geoquímicos, multidimensionales por naturaleza, son esenciales para comprender diversos procesos planetarios y medioambientales, entre los que se destacan el origen y la evolución de la Tierra y del Sistema Solar. Por otro lado, los avances tecnológicos en el instrumental analítico, han aumentado espectacularmente el volumen y la variedad de datos geoquímicos disponibles. La exploración de estos conjuntos de datos multidimensionales de naturaleza tan compleja exige técnicas de análisis visual (AV) adecuadas, un alto grado de interactividad y metáforas intuitivas. Dado que existe una correlación natural entre la complejidad de los datos y la complejidad de las herramientas para estudiarlos, los objetivos analíticos son más difíciles de alcanzar. Por lo tanto, tiene sentido considerar métodos que guíen o ayuden a los usuarios en el proceso de análisis visual. En este contexto, nuestro objetivo general es contribuir al diseño y desarrollo de técnicas de analítica visual que contribuyan efectivamente al análisis de datos geoquímicos tanto de rocas como de minerales.

Palabras claves: *Análisis Visual de Datos, Análisis Visual de Datos Geoquímicos,*

Visualización de Datos Multidimensionales, Explicabilidad, Guías de Exploración de Datos

CONTEXTO

Este trabajo se realiza en estrecha colaboración con investigadores del INGEOSUR-CONICET (Instituto Geológico del Sur), del Departamento de Geología de la Universidad Nacional del Sur, y el VyGLab (Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica) del Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS), instituciones de investigación de reconocido prestigio tanto nacional como internacional.

1. INTRODUCCIÓN

La capacidad actual de generación y recopilación de datos supera con creces la de su procesamiento de una forma adecuada y eficaz y más aún la de una asimilación inteligente de los mismos. En las últimas décadas, la capacidad de aprovechar esos datos no ha aumentado al mismo ritmo que la disponibilidad de los mismos, siendo ésta una constante en la mayoría de las disciplinas. En particular, se puede afirmar que en la actualidad las Geociencias han entrado en la era de los grandes conjuntos de datos.

En este campo, Reichstein *et al.* 2019 [Reichstein et al. 2019] plantean que, para sacar el máximo partido de este crecimiento explosivo de los datos del sistema terrestre y de su diversidad, en los próximos años debemos enfrentarnos a dos grandes tareas: extraer conocimiento de esta avalancha de datos y derivar modelos que aprendan mucho más de los datos de lo que pueden hacerlo los enfoques tradicionales para extraer información, colaborando con el usuario en su exploración y respetando al mismo tiempo la comprensión evolutiva de las leyes de la naturaleza por parte de éste.

En el área de la geoquímica aplicada, la existencia de grandes conjuntos de datos proporciona oportunidades y desafíos en el procesamiento de datos de estudios geoquímicos [Zuo et al. 2021] tanto de rocas como de minerales. Los datos geoquímicos son muy utilizados en la exploración minera, ya que desempeñan un rol destacado tanto en el descubrimiento de nuevos yacimientos minerales como en el estudio de los ya conocidos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Nuestras líneas de investigación investigación y desarrollo se basan en la utilización de uno de los métodos más ampliamente utilizados hoy en día para el preprocesamiento de datos geoquímicos de exploración, esto es, el análisis de datos composicionales.

Esto permite procesar los datos de exploración geoquímica con el fin, entre otros, de revelar la asociación espacial de patrones geoquímicos, identificar asociaciones elementales e identificar anomalías geoquímicas asociadas a los procesos mineralizantes.

Estas líneas específicas corresponden al diseño y desarrollo de métodos y herramientas

de analítica visual integrando técnicas de análisis visual y de machine learning aplicados a:

Clasificación y diferenciación automática de rocas máficas y ultramáficas y visualización de los métodos aplicados a las tomas de decisiones durante este proceso.

Según Ferracutti et al. [Ferracutti et al. 2017], afloramientos del cuerpo intrusivo Viorco junto con cuerpos intrusivos adyacentes, presentan interés económico [Ferracutti et al. 2005, Ferracutti et al. 2013]. Éstos son cuerpos intrusivos con estratificación magmática primaria. Esta característica se manifiesta como una alternancia de niveles de rocas máficas y ultramáficas. En lo que respecta a la actividad exploratoria de una empresa minera, es fundamental poder discriminar entre rocas máficas y ultramáficas durante la descripción de testigos de perforación.

La clasificación de rocas máficas y ultramáficas, está basada en los porcentajes de determinados minerales que integran estas rocas [Le Maitre et al. 2002]. La identificación de los mismos se lleva a cabo a partir de un análisis petrográfico de muestras de roca, utilizando microscopios de luz transmitida.

En lo que respecta al ámbito de la exploración minera, este proceso plantea una serie de problemas. En primer lugar, tanto la confección de secciones delgadas como el análisis petrográfico, conllevan un determinado tiempo, el cual no es compatible con los tiempos que maneja una empresa minera. En segundo lugar, un análisis petrográfico se realiza sobre muestras de roca puntuales, lo cual resultaría inviable de realizar sobre la totalidad de las rocas de perforaciones que abarcan miles de metros.

En este contexto, resulta de interés el estudio y desarrollo de técnicas de analítica visual que integren modelos de Machine Learning y

de Visualización. El uso de técnicas de Machine Learning nos permitirá automatizar el proceso de análisis de grandes cantidades de datos, considerar y evaluar las proporciones de los distintos minerales acelerando así la clasificación de las rocas máficas y ultramáficas. Por otro lado, en este proceso de automatización se debe explicar a los usuarios, geólogos en este caso, cómo y por qué se tomaron las decisiones que llevaron a dicha clasificación.

Análisis Visual Guiado de datos geoquímicos de espinelos (grupo de minerales). Las técnicas de visualización nos proporcionan herramientas útiles para comprender y explorar datos geoquímicos en minerales. El análisis visual proporciona información valiosa acerca de la información subyacente en estos datos. Los datos geoquímicos son complejos y multidimensionales. Además, contienen valores atípicos, el contenido de los elementos varía en miles de ppm (partes por millón) y los distintos elementos químicos pueden estar correlacionados. Si bien existen diversas herramientas que permiten visualizar datos multidimensionales, estas no proveen técnicas que se ajusten adecuadamente a datos geoquímicos en general y a espinelos en particular. Por otro lado, si bien hay algunas aplicaciones que permiten el análisis visual de estos datos [Ganuza et al. 2012, Ganuza et al. 2014, Ganuza et al. 2017], éstas se restringen a un máximo de 6 dimensiones. Esto motiva el desarrollo de métodos y herramientas adecuadas para analizar datos geoquímicos que sean poderosos. Un efecto adverso de las herramientas complicadas es que los objetivos analíticos son más difíciles de alcanzar; sin embargo, el análisis visual de datos se concibe como una colaboración entre el usuario y el sistema informático con el fin de completar una tarea determinada. Por lo tanto, tiene sentido considerar métodos que guíen o

ayuden a los usuarios en el proceso de análisis exploratorio de estos datos.

La búsqueda de una integración eficaz entre usuario y sistema (método/aplicación) en la que el sistema ayude activamente al usuario a alcanzar su objetivo de análisis, ha sido objeto de investigación durante muchos años [Ceneda et al. 2016, Ceneda et al. 2018, Escarza et al. 2011, May et al. 2012, Sacha et al. 2014, Streit et al. 2012]. Sin embargo, este problema sigue sin resolverse de forma sustancial. Como consecuencia, los usuarios pueden verse abrumados por sistemas de análisis visual poderosos pero complicados, lo que también limita su capacidad para producir resultados relevantes. Así, la guía es prometedora para lograr una colaboración eficaz usuario-sistema que favorezca el análisis visual.

3. RESULTADOS OBTENIDOS Y ESPERADOS

De las líneas de trabajo delineadas se han obtenido resultados parciales.

En lo que respecta a la clasificación automática de rocas máficas y ultramáficas y a la visualización de cómo se toman las decisiones durante este proceso, se han aplicado diversos modelos de aprendizaje automático supervisados y no supervisados para la clasificación de muestras de datos. Los resultados obtenidos de la clasificación automática fueron cotejados con los esperados por los geólogos, quienes afirmaron la utilidad y exactitud de estas nuevas técnicas.

Además, en la clasificación visual que realizan los geólogos se observan ciertos minerales en particular. Mediante el análisis de la importancia de ciertas características, pudimos determinar que al considerar adicionalmente otros minerales en la clasificación de las rocas los resultados podrían mejorarse significativamente.

Con respecto al Análisis Visual de datos geoquímicos de espinelos se ha desarrollado una herramienta de Visualización extensible, *SpinelWEB* [Antonini et al. 2021], para el Análisis Visual de datos geoquímicos. Si bien dicha herramienta está especialmente orientada a datos geoquímicos del grupo de minerales del espinelo, puede utilizarse para el análisis de datos geoquímicos en general. Actualmente se están desarrollando alternativas que permitan visualizar estos datos. Entre éstas, se está trabajando con diversas técnicas como *Paired y Non-Paired General Line Coordinates* [Luque et al. 2021, Antonini et al. 2022] y técnicas utilizadas en las visualización de clusters y conjuntos. Conjuntamente se está trabajando en guías de exploración con el objetivo de sacar ventaja de la sinergia entre los expertos del dominio y método/aplicación. Estas guías de orientación en la exploración, no se han explorado previamente en el contexto de aplicaciones geológicas.

4. FORMACION DE RECURSOS HUMANOS

A continuación, se detallan proyectos de investigación, tesis finalizadas y en desarrollo y becas obtenidas dedicadas a temáticas vinculadas con las líneas de investigación presentadas.

Proyectos Vigentes

-PIBAA - CONICET (2872021010 0824CO)
Análisis Visual de Datos Multidimensionales sin Pérdida de Información. Directora: Dra. M. Luján Ganuza.

-PGI 24//N048 *Análisis Visual de Datos.* Directora: Silvia M. Castro.

-PICT-2017-1246 *Análisis Visual en Geociencias.* Directora: Silvia M. Castro.

Tesis en Desarrollo

-Análisis Visual de Datos Multidimensionales, tesis de Doctorado en Ciencias de la Computación. Alumna: Antonella S. Antonini. Directora: Dra. Silvia Castro. Codirectora: Dra. M. Luján Ganuza.

Becas

-Antonella S. Antonini. Título del plan propuesto: *Análisis Visual de Datos Multidimensionales*. Beca doctoral 2019 CONICET. Adjudicada a partir de abril de 2019 y por un término de 5 años. Directores: Dra. Silvia Castro - Dr. Ernesto Bjerg.

5. BIBLIOGRAFÍA

[Ceneda et al. 2016] Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H.J., Streit, M. and Tominski, C., 2016. *Characterizing guidance in visual analytics*. IEEE TVCG, 23(1), pp.111-120.

[Ceneda et al. 2018] Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Streit, M. and Tominski, C., 2018, June. *Guidance or No Guidance? A Decision Tree Can Help*. In EuroVA@ EuroVis (pp. 19-23).

[Ceneda et al. 2019] Ceneda, D., Gschwandtner, T. and Miksch, S. 2019. *A review of guidance approaches in visual data analysis: A multifocal perspective*. In Computer Graphics Forum (Vol. 38, No. 3, pp. 861-879).

[Ferracutti et al. 2005] Ferracutti, G., Mogessie, A., and Bjerg, E. 2005. Chemical and mineralogical profile of the Las Águilas mafic-ultramafic drill core, San Luis Province, Argentina. *Mitteilungen der Österreichischen Mineralogischen Gesellschaft* 151: 40 (2005).

[Ferracutti et al. 2013] Ferracutti, G., Bjerg, E. and Mogessie, A. 2013. Petrology, geochemistry and mineralization of the Las Águilas and Virorco mafic-ultramafic bodies, San Luis Province, Argentina. *International Journal of Earth Sciences*, 102(3), 701-720.

[Ferracutti et al. 2017] Ferracutti, G., Bjerg, E., Hauzenberger, C., Mogessie, A., Cacace, F. and Asiain, L. 2017. Meso to Neoproterozoic layered mafic-ultramafic rocks from the Virorco back-arc

intrusion, Argentina. *Journal of South American Earth Sciences*, 79, 489-506.

[Le Maitre et al. 2002] Le Maitre, R. W., Streckeisen, A., Zanettin, B., Le Bas, M. J., Bonin, B., Bateman, P., Bellieni, G., Dudek, A., Efremova, S., Keller, J., Lameyre, J., Sabine, P. A., Schmid, R., Sørensen, H. and Wooley, A.R. 2002. *Igneous Rocks: A Classification and Glossary of Terms*. Cambridge 774 University Press, 236 p., Cambridge.

[Antonini et al. 2021] Antonini, A.S., Ganuza, M.L., Ferracutti, G.R., Gargiulo, M.F., Matković, K., Gröller, E., Bjerg, E.A., Castro, S.M. (2021). *SpinelWeb: an interactive web application for visualizing the chemical composition of spinel group minerals*. *Earth Science Informatics*, 1-8.

[Antonini et al. 2022] Antonini, A.S., Luque, L., Ganuza, M. L., Castro, S.M. (2022). *Towards a Taxonomy for Non-Paired General Line Coordinates: A Comprehensive Survey*. *International Journal of Data Science and Analytics*.

[Escarza et al. 2011] Escarza, S., Larrea, M., Urribarri, D., Martig, S., Castro, S., (2011) *Integrating Semantics in the Visualization Process*, Book "Scientific Visualization: Interactions, Features, Metaphors", pp. 92-102, Series: Dagstuhl FOLLOW-UPS Vol.2, Hans Hagen Ed, Publisher, SchlossDagstuhl--Leibniz-Zentrum fuer Informatik. Dagstuhl, Alemania. ISBN 978-3-939897-26-2.

[Ferracutti et al. 2015] Ferracutti, G.R., Gargiulo, M.F., Ganuza, M.L., Bjerg, E.A., Castro, S.M. (2015) *Determination of the spinel group end members based on electron microprobe analyses*. *Mineral. Petrol.* 109, 2, 153-160.

[Ganuza et al. 2012] Ganuza, M. L., Castro, S. M., Ferracutti, G.R., Bjerg, E. A., Martig, S. (2012) *Spinelviz: An interactive 3d application for visualizing spinel group minerals*. *Comput. Geosci.* 48, 50–56.

[Ganuza et al. 2014] Ganuza, M.L., Ferracutti, G., Gargiulo, M.F., Castro, S.M., Bjerg, E.A., Gröller, E., Matkovic, K. (2014) *The spinel explorer - interactive visual analysis of spinel group minerals*. *IEEE TVCG*. 20, 12, 1913–1922.

[Ganuza et al. 2017] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., Matkovic, K. (2017) *Interactive visual*

categorization of spinel-group minerals. *Proc. of the Spring Conf. on Comput. Graph.*

[Luque et al. 2021] Luque, L.E., Ganuza, M.L., Antonini, A.S., Castro, S.M. (2021). *npGLC-Vis Library for Multidimensional Data Visualization*. *En Cloud Computing, Big Data & Emerging Topics: 9th Conference, JCC-BD&ET, La Plata, Argentina, June 22-25, 2021, Proceedings* (Vol. 1444, p. 188).

[May et. al 2012] May, T., Steiger, M., Davey, J., Kohlhammer, J. *Using signposts for navigation in large graphs*. *Computer Graphics Forum*, 31(3pt2):985–994, 2012.

[Reichstein et al. 2019] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. (2019) *Deep learning and process understanding for data-driven Earth system science*. *Nature* 566:195–204.

[Sacha et al. 2014] Sacha, D., Stoffel, A., Stoffel, F., Kwon, B.C., Ellis, G. and Keim, D.A., 2014. *Knowledge generation model for visual analytics*. *IEEE TVCG*, 20(12), pp.1604-1613.

[Streit et. al 2012] Streit, M., Schulz, H.-J., Lex, A., Schmalstieg, D., Schumann, H. *Model driven design for the visual analysis of heterogeneous data*. *IEEE TVCG*, 18(6):998–1010.

[Zuo et al. 2021] Zuo R, Wang J, Xiong Y, Wang Z (2021) *The processing methods of geochemical exploration data: past, present, and future*. *Appl Geochem* 132:105072.

Industrias Inteligentes: Detección de uso de Elementos de Protección Personas en una Planta de Reciclaje

¹Diego Alberto Godoy, ^{1,2}Enrique Marcelo Alborno, ¹Lucas Gabriel Kucuk,
¹Hernán Bareiro, ¹Cesar Gallardo, ¹Ricardo Selva

¹Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.) Departamento de Ingeniería y Ciencias de la Producción-Universidad Gastón Dachary
Av. López y Planes 6519- Posadas, Misiones, Argentina.

²Instituto de investigación en Señales, Sistemas e Inteligencia Artificial - sinc(i), CONICET-UNL

diegodoy@citic.ugd.edu.ar, emalbornoz@sinc.unl.edu.ar, lukucuk@gmail.com,
hbareiro@citic.ugd.edu.ar, ricardosel@gmail.com, cesar.cgallardo@gmail.com

Resumen

En este trabajo se presenta un proyecto de investigación denominado “Tecnologías para Desarrollos Sostenibles de Ciudades Inteligentes apoyados por machine learning”. Particularmente en este artículo se presenta el desarrollo de un prototipo basado en visión computacional que permite detectar si operarios están utilizando correctamente elementos de protección personal. El mismo, implementa YOLOv5, una red neuronal convolucional para detección y clasificación de objetos en imágenes. Se realizaron las pruebas utilizando datasets obtenidos de la web y de autoría propia, que tiene que ver con los elementos de seguridad de los operarios. Se determina como resultado que se detectan adecuadamente los objetos incluso bajo dificultades propias de cada escenario de prueba como la luminosidad y variabilidad del entorno.

Palabras claves: visión computacional, Redes neuronales convolucionales, Elementos de protección personal.

Contexto

Este trabajo tiene como contexto al proyecto de investigación denominado “Tecnologías para Desarrollos Sostenibles de Ciudades Inteligentes apoyados por machine learning”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) con el número Código IP A12001/22 y radicado en el Centro de

Investigación en Tecnologías de la Información y Comunicaciones de dicha universidad.

El mismo fue incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°12 mediante la Resolución Rectoral (R.R). 37-A-22 y es una continuidad de los Proyectos Tecnologías Para Desarrollos Sostenibles De Ciudades Inteligentes. R.R. UGD N° 26/A/19. Diseño de Simuladores de Procesos de Desarrollo de Software Ágiles y Redes de Sensores Inalámbricos para la Industria y la Academia. R.R. UGD N° 07/A/17 y Simulación como herramienta para la mejora de los procesos de software desarrollados con metodologías ágiles utilizando dinámica de sistemas, R.R. UGD N° 18/A/14 y R.R. UGD N° 24/A/15.

Introducción

Las personas involucradas en una organización son el activo más valioso, ya que de ellos depende que se realice una determinada actividad productiva como una tarea operativa manual o el manejo de determinadas máquinas. Cada día, los empleados están expuestos a los peligros del ambiente del trabajo, donde un accidente puede resultar en heridas, lesiones, discapacidades e incluso el deceso de uno o más trabajadores [1]. Además, en el ambiente laboral están presentes distintos agentes de riesgo que, en el largo plazo, pueden generar diversas enfermedades profesionales [2]. La tarea de los EPP (Elementos de Protección

Personal) consiste en mitigar las consecuencias de estos accidentes y agentes de riesgo aumentando la seguridad del trabajador para evitar daños, por lo que la falta de uso o la utilización incorrecta de los mismos supone un riesgo aún mayor para los empleados, y para el empleador de ser sancionado por el ente controlador. Algunos de los EPP que pueden ser exigidos por los empleadores son cascos, protector facial, protector ocular, guantes, chalecos refractarios y zapatos.

Detectar a las personas, en términos computacionales, es una tarea compleja ya que se debe emular al cerebro humano para lograrlo: se requiere, por un lado, reconocer a cada trabajador y las partes de su cuerpo y, por otro lado, por cada parte del cuerpo, verificar que se esté utilizando el EPP correspondiente; por ejemplo, para la cabeza se debe verificar que se esté utilizando el casco o en las manos verificar que se esté utilizando guantes. Todo esto se realiza mientras los trabajadores cumplen con sus labores en movimiento, cambiando de posición y de posturas, alejándose y acercándose a la cámara. Es necesario tener en cuenta también la calidad de las filmaciones, las distintas condiciones de iluminación que se dan durante la jornada y los distintos colores y versiones disponibles para cada tipo de EPP.

Redes Neuronales Convolucionales

Una red neuronal convolucional (Convolutional Neural Network - CNN) es un tipo de red artificial donde las neuronas se corresponden con campos receptivos de una forma muy similar a las neuronas en la corteza visual de un cerebro biológico. Esta red es una variación del perceptrón multicapa, pero gracias a que utiliza matrices bidimensionales, es muy efectiva para tareas de visión artificial como la clasificación de imágenes. Funciona con aprendizaje de tipo supervisado, por lo cual dichas imágenes deben estar etiquetadas para ser utilizadas en el entrenamiento de la red, permitiendo a ésta aprender los patrones presentados [3].

La arquitectura de este tipo de red consiste en múltiples capas de filtros convolucionales. Las

primeras se encargan de la extracción de características, como líneas verticales u horizontales. Las últimas, en cambio, realizan la clasificación final sobre las características identificadas

Línea de Investigación

En esta línea de investigación se utilizó como caso de estudio la planta de reciclaje GIRSU (Gestión Integral de Resíduos Sólidos Urbanos) [4]. Esta institución municipal contribuye con el desarrollo sostenible, que se encarga de la reducción de residuos enviados a los centros de depósitos finales, promoviendo la reutilización y reciclaje de los mismos. Es un ambiente industrial con muchas maquinarias que presentan diversos riesgos para la salud. Los operarios deben usar en todo momento sus EPP personal, como cascos, guantes, zapatos y chaleco

Teniendo en cuenta la problemática planteada y el caso de estudio, se han planteado varios objetivos. El objetivo principal de la misma corresponde es:

Diseñar un prototipo para la identificación del uso de EPP en la planta GIRSU utilizando visión computacional.

Los objetivos específicos son: Analizar el estado del arte en la identificación de EPP. Caracterizar la utilización de EPP en la planta GIRSU en base a muestras de imágenes. Determinar los requisitos del prototipo para la detección de uso de EPP en GIRSU. Diseñar la arquitectura del prototipo y de la red neuronal convolucional. Definir escenarios de pruebas para verificar la utilidad del prototipo.

Resultados

Desde lo metodológico se optó por utilizar CRISP-DM (Cross industry standard process for data mining)[5], que es un enfoque iterativo, adaptable y flexible utilizado en los procesos de minería de datos. Ofrece flexibilidad de agregar o quitar etapas que no sean útiles para un proyecto, abriendo la posibilidad de iterar en caso que sea necesario.

La solución tecnológica consiste en el desarrollo de un prototipo basado en una red

neuronal convolucional capaz de detectar imágenes de objetos, en este caso EPP de los operarios de la planta, para este caso en concreto se utilizó YOLOv5 [6]. Este será capaz de estimar si los trabajadores visibles en dichas imágenes traen consigo los EPP. La salida de la red es procesada por un algoritmo que decide si la persona utilizando los elementos.

Para ello se procedió a recopilar las imágenes para conformar el dataset del proyecto, por un lado, se utilizó uno obtenido de la web del cual se pudo obtener 1280 imágenes y por otro lado se tomaron 136 fotos de los trabajadores de la planta realizando sus funciones (Figura 1), también se les solicitó tomar fotos de frente con el equipamiento (Figura 2).



Figura 1. Operarios de la planta



Figura 2. Operarios de la planta de frente

Una vez obtenido el dataset se procedió a la adaptación del mismo para que pueda ser utilizado por YOLOv5. Para agilizar la preparación, se procedió a realizar el etiquetado y formateo del dataset utilizando Roboflow que es un software que permite trabajar sobre la organización del mismo [7].

Antes de iniciar con el etiquetado, se analizaron las reglas del mismo considerando posibles falsos positivos. Entre las etiquetas de los elementos se determinaron las clases casco,

chalecos reflectarios, guantes, cabeza y el cuerpo entero de la persona, dando en total 6 etiquetas por persona (Figura 3).

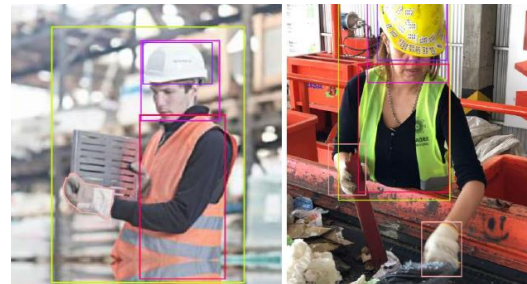


Figura 3. Etiquetado en las imágenes

En la siguiente tabla se describen los porcentajes y proporciones de cada una de las clases por los datasets utilizados para entrenar, validar y testear la red.

Tabla 1. Porcentajes de clases datasets

	Aprendizaje	Verificación	Testeo
Cabezas	32,08%	30,94%	27,86%
Cascos	28,68%	31,02%	19,20%
Chalecos	2,38%	2,57%	16,10%
Guantes	5,59%	4,95%	6,50%
Personas	31,27%	30,52%	30,03%
% totales	100%	100%	100%

La etapa de funcionamiento consiste en las pruebas del prototipo a través del ingreso de imágenes. La red se configura con el conjunto de pesos obtenidos de la etapa de entrenamiento y realiza la inferencia sobre las entradas, obteniendo así 4 coordenadas que encuadran el objeto en la imagen y un valor de confianza entre 0 y 1 que representa que tan segura está la red de la estimación (Figura 4).



Figura 4. Recuadros con valores de confianza

Finalmente, las detecciones sirven como entrada al algoritmo programado en python encargado de decidir si las personas están utilizando los EPP. Esta toma cada persona detectada y realiza una revisión en el orden: si

usa casco, luego chaleco reflectarios y luego los guantes. Resultando en un recuadro verde si está utilizando todos los EPP o sino un recuadro rojo (Figura 5).



Figura 5. Recuadros Rojo para faltante, verde no hay faltantes de EPP

Además del centro verde se utilizaron imágenes de distintos sitios, las fotografías representan a trabajadores de obras públicas de otros entornos, como ser trabajadores de la obra pública de la ciudad de Posadas, Vialidad Provincial y de Vialidad Nacional, el objetivo era probar la generalización del prototipo. En la imagen se puede observar que se detectaron correctamente las faltantes de EPP de varias personas, con excepción de dos personas del fondo donde a una no se le detectó correctamente el casco y a la otra el chaleco (Figura 6). También se trabajó sobre una imagen de trabajadores de Vialidad donde detectó correctamente la falta de casco (Figura 7).

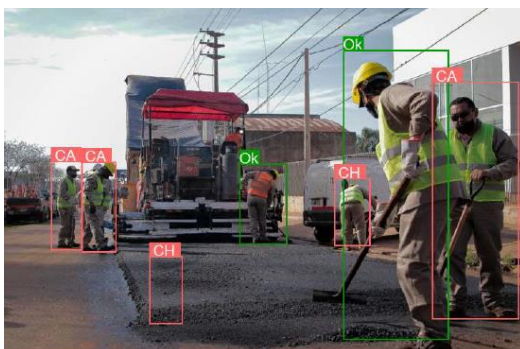


Figura 6. Operarios externos



Figura 7. Trabajadores de vialidad detección correcta de casco

Según los resultados obtenidos con las pruebas realizadas con el prototipo, es posible decir que resulta factible su implementación. Desde el punto de vista del modelo es destacable que con los parámetros del entrenamiento y el dataset se lograron grandes resultados en una primera iteración. Se propone como trabajo futuro mejorar el desempeño de la red generando un dataset con un mayor número de imágenes de guantes y manos, por lo que es deseable recolectar un mayor número de imágenes en el centro verde municipal.

Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por tres investigadores, Dos Doctores disciplinares en Tecnologías de la Información (uno por Parte de U.N.L. y uno por Parte de U.G.D.); dos Doctorandos en Informática y cinco estudiantes en período de realización de trabajos finales de carrera (TFC) de Ingeniería en Informática de la UGD. Actualmente, el número de TFC es de tres en proceso de desarrollo.

Bibliografía

- [1] C. Ray Asfahl, Seguridad industrial y administración de la salud. Pearson Educación, 2010.
- [2] L. E. S. Chávez, J. W. U. Vicuña, and L. B. C. Camacho, "Evaluación de la Exposición a agentes de riesgo físico en centros de salud," Polo del Conocimiento: Revista científico - profesional, vol. 5, no. 10, pp. 424–439, 2020.
- [3] M. Massiris, C. Delrieux, and J. Á.

Fernández Muñoz, “Detección de equipos de protección personal mediante red neuronal convolucional YOLO,” in XXXIX Jornadas de Automática, 2018, pp. 1022–1029.

- [4] Municipalidad de Posadas, “Centro Verde Municipal.”
<https://posadas.gov.ar/sustentable/centro-verde-municipal/> (accedido 12/01/2022).
- [5] “Conceptos básicos de ayuda de CRISP-DM.” <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview> (accedido 02/11/2022).
- [6] “ultralytics/yolov5,” GitHub.
<https://github.com/ultralytics/yolov5> (accedido 26/09/2022).
- [7] “Roboflow: Give your software the power to see objects in images and video.” <https://roboflow.com/> (accedido 02/11/2022).

LA IMPORTANCIA DE LA REALIDAD EXTENDIDA EN EL METAVERSO

G. Rodríguez, N. Jofré, Y. Alvarado, J. Fernandez, R. Guerrero
Laboratorio de Computación Gráfica / Dpto. Informática / FCFMyN
Universidad Nacional de San Luis
Ejército de los Andes 950, Tel: 0266 4420823, San Luis, Argentina
{gbrodriguez, npasinetti, ymalvarado, jmfer, rag}@unsl.edu.ar

RESUMEN

El crecimiento exponencial de las nuevas tecnologías prácticamente ha impactado en todos los ámbitos de nuestras vidas. En consecuencia, esto ha generado Realidades Alternativas que permiten alterar la percepción de un mundo canónico. En este contexto ha surgido una innovación comunicacional: el Metaverso.

El objetivo del Metaverso es construir un mundo virtual paralelo al mundo real con una sociedad y un sistema económico estables, que permita a cada usuario editar el mundo y producir contenidos. Es una tecnología emergente en evolución que experimenta ciclos de sobreexpectación y, al tiempo que es potencialmente disruptivo, todavía no ha sido completamente comprendido ni tampoco suficientemente investigado. Al día de hoy, el desafío consiste en cómo llevarlo a un próximo nivel.

Por su parte, la Realidad Extendida (concepto que abarca la Realidad Virtual, la Realidad Aumentada y la Realidad Mixta) da soporte de base al Metaverso. En los últimos años, ésta ha experimentado un notable progreso. Las investigaciones sobre la Realidad Extendida han aportado beneficios a partir de dos conceptos clave: el grado de inmersión y la sensación de presencia. En la presente propuesta de trabajo se propone analizar las aportaciones de la Realidad Extendida para el desarrollo del Metaverso.

Palabras clave: Metaverso, Realidad Extendida, Interacción Natural.

CONTEXTO

La propuesta de trabajo se lleva a cabo dentro del proyecto “*Realidades Alternativas como lenguaje generativo aplicado a la solución de problemas reales*”. Este proyecto es desarrollado en el ámbito del Laboratorio de Computación Gráfica de la Universidad Nacional de San Luis.

1. INTRODUCCIÓN

El constante desarrollo tecnológico ha ocasionado la aparición de nuevas tecnologías emergentes. En los últimos años, el crecimiento exponencial en su utilización ha impactado en todos los ámbitos de nuestras vidas, y en muchos casos, ha mejorado nuestra calidad de vida. Las nuevas tecnologías generan un gran volumen de información y conocimiento, lo cual ha modificado la manera en que el ser humano se comunica con sus semejantes; convirtiendo a la información y al conocimiento en primordiales para el avance social. Más aún, éstas han obligado a desarrollar nuevas estrategias para la transmisión y adquisición del conocimiento [1].

En particular, la comunicación juega un rol clave en situaciones donde se requiere una visión generalizada del problema, tales como la toma de decisiones y la resolución de problemas. En este contexto de innovación comunicacional, la Realidad Extendida (Realidad Virtual - RV, Realidad Aumentada - RA y Realidad Mixta - RM) se ha convertido en una alternativa válida que provee de recursos a los sistemas computacionales para

la transmisión de información al usuario en forma rápida, eficiente, natural e intuitiva [2].

Como corolario de la pandemia mundial reciente, ha surgido un nuevo concepto que evoca de manera subliminal la existencia de mundos virtuales paralelos: el Metaverso. Esta es una tecnología emergente compuesta de un conjunto de herramientas, conceptos, innovaciones y avances utilizados en diversos contextos al servicio de diversos propósitos. Pretende combinar todos los aspectos de la vida real en un solo lugar, permitiendo a los usuarios trabajar, reunirse, jugar y socializar en diversos espacios virtuales. El metaverso tiene el potencial de extender el mundo físico utilizando como base a las tecnologías de RV, RA y RM. Más aún, aunque no es obligatorio, para ofrecer una experiencia virtual inmersiva perfecta en el metaverso, las empresas de tecnología están incorporando tecnologías de vanguardia para impulsar el desarrollo del mundo 3D. Dichas tecnologías incluyen Interacción Humano-Computadora, Blockchain, Inteligencia Artificial, Visión por Computadora, Internet de las Cosas, Robótica, Computación en la Nube y Computación Frontera [3-4].

Paralelamente, el contexto reciente de pandemia ha puesto el foco en las tecnologías base del metaverso (RV, RA, RM), esperando de ellas un impulso para la industria y los sectores relacionados. Estas tecnologías son lenguajes de comunicación visual con probada importancia y eficacia en brindar soluciones, estando cada vez más adentradas en diferentes ámbitos de la sociedad. Sin embargo, su problema se centra en la robustez de las aplicaciones que se han logrado con dichas tecnologías. Según su definición, este tipo de tecnología es comunicación pura (directa e indirecta) que relaciona al usuario con el hardware asociado; y considerando como punto de vista a la transmisión de información, uno de los debates actuales más significativos en el área es el referido a cómo robustecer la comunicación provista por este tipo de aplicaciones [5-6].

El disponer de información y conocimiento es trascendental para la resolución de problemas y toma de decisiones en diferentes

ámbitos. La capacidad de resolución de problemas de una persona está íntimamente ligada con su voluntad, capacidad y agilidad para aprender de la experiencia y luego aplicar estos conocimientos en la resolución exitosa de nuevas situaciones. En los últimos años existe una creciente preocupación en mejorar las habilidades relacionadas con los procesos que optimizan la comunicación de información. Bajo esta consideración, el núcleo de construcción del metaverso, a través de la composición de numerosos espacios virtuales compartidos, debe considerar, y en consecuencia fusionar, las acciones simultáneas entre todos los objetos, avatares que representan a sus usuarios, y sus interacciones. Es importante señalar que diferentes usuarios de un mundo virtual, es decir, un subconjunto del metaverso, deben recibir información idéntica y consistente al mismo tiempo que los usuarios puedan interactuar entre sí de manera coherente y en tiempo real [2].

Actualmente, los investigadores, desarrolladores, empresarios y emprendedores concuerdan en percibir al metaverso como el nuevo paradigma para crear y optimizar procesos de intercambio y comunicación de información. En consecuencia, el metaverso se ha convertido en eje de investigación para los grandes centros de investigación, desarrollo e innovación, quienes buscan descubrir su potencial y extender sus desafíos [7].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La industria tecnológica trabaja, desde hace varias décadas, en el desarrollo de las tecnologías de RV y RA. En sus orígenes, estas tecnologías se hicieron fuertes, en forma aislada y totalmente independiente entre ellas, en un nicho clave y muy relacionado a la innovación: el gaming. Los videojuegos fueron terreno de pruebas fértil y accesible, con usuarios jóvenes y dispuestos a dar sus primeros pasos en los mundos virtuales. Hoy en día, los dispositivos antes pensados como

futuristas e inaccesibles se acercan cada vez más a los consumidores del mundo, achicando la brecha entre lo virtual y lo real. Con este antecedente, el metaverso propone la inmersión completa para trabajar y relacionarse con otros, a través de diversas tecnologías pertenecientes a las ciencias de la computación y la comunicación. Las áreas que hasta el momento se han visto claramente beneficiadas son la salud, medicina y negocios/comercio.

Dado que el metaverso toma como base a la Realidad Extendida, el punto de partida consiste en definir como líneas de investigación básicas a las tres tecnologías involucradas (RV, RA, RM), intentando potenciar las áreas mencionadas al mismo tiempo que se exploran nuevas áreas de aplicación.

- **RV para el Metaverso.**

La RV tiene como objetivo lograr una experiencia completamente inmersiva y recreada totalmente gracias a las herramientas que brindan las tecnologías de la información. Por esta razón, las tecnologías de RV están en camino a convertirse en futuros periféricos de los usuarios del metaverso.

En educación, las salas de RV permiten a los alumnos un acceso grupal. Los profesores usan estos entornos virtuales para transportar a los alumnos, de forma rápida y segura, no sólo a otro lugar físico sino incluso ubicado en una época diferente, viajando en el tiempo, y posibilitando el aprendizaje de los alumnos de manera práctica y emocionante [8].

En medicina, la experiencia inmersiva y la interoperabilidad del metaverso abren nuevas vías para la formación médica. Por ejemplo, los estudiantes pueden interactuar entre sí al mismo tiempo que se sumergen en un cuerpo humano virtual para obtener una visión completa de los órganos humanos o simular una intervención quirúrgica real [9].

En economía, la RV habilita una nueva forma de comprar, muy parecida a

la vida real, pero totalmente virtual, en dónde los clientes pueden interactuar de manera inmersiva con el entorno. Una experiencia para destacar es el uso activos digitales (Criptomonedas o NFT - *Non-Fungible Token*) que han demostrado tener la capacidad de reemplazar a objetos reales [10].

- **RA para el Metaverso.**

La RA permite la superposición de una capa de conocimiento virtual sobre elementos reales para mejorar la comprensión de la realidad por parte del usuario [11]. Es decir, la RA brinda una vista mejorada de una imagen real.

En el ámbito educativo, las aplicaciones de RA brindan a los estudiantes la capacidad de interactuar entre sí a través de un entorno de aprendizaje virtual dentro del mundo virtual-físico [12].

Con la reciente pandemia, la RA ha sido utilizada por la Organización Mundial de la Salud (OMS) para formar a los expertos en salud mental sobre cómo utilizar la RV para tratar a personas con enfermedades mentales y dificultades emocionales. También se ha utilizado en facultades de medicina con fines educativos [13].

En el ámbito empresarial, los consumidores pueden conocer a fondo la calidad y el estilo de los productos antes de comprarlos, lo que puede reducir la tasa de devoluciones de productos y ofrecer comodidad tanto a comerciantes como a consumidores [14]. Por otro lado, la pandemia COVID-19 ha acelerado la transformación digital del turismo. Los museos digitales rompen con el formato tradicional de galería presencial y ofrecen a los visitantes experiencias sociales y culturales novedosas en un espacio virtual [15].

Las tecnologías de RA en conjunto con redes de comunicación y el internet de las cosas permiten que se puedan recopilar diversos datos de una ciudad y analizarlos para obtener soluciones inteligentes que mejoren la vida pública

(ciudades inteligentes) [16].

- **RM para el Metaverso.**

La Realidad Mixta es, sin lugar a duda, la más desconocida e innovadora de las tecnologías contenidas dentro de la Realidad Extendida [17]. En este caso, lo que se busca es crear una nueva forma de interactuar con los entornos físicos/reales y los tecnológicos/virtuales al mismo tiempo. La RM combina lo mejor de los elementos de la RV y la RA de una forma diferente y novedosa. Mezcla en tiempo real componentes del mundo digital con el mundo real al punto en que es posible interactuar con los elementos digitales como si fueran objetos reales [2].

Los usos de la Realidad Mixta abarcan desde la educación [18], pasando por el entrenamiento [19] hasta el entretenimiento [19]. Algunos ejemplos son las experiencias educativas que permiten a los niños ir a un safari en las llanuras de África desde su aula escolar, o incluso la reparación de cápsulas espaciales que los astronautas usan para su entrenamiento sin salir de sus instalaciones en tierra. Otras aplicaciones se relacionan con Medicina [20], Museos [21], Turismo [22], Comida [23], entre otras.

3. RESULTADOS OBTENIDOS/ESPERADOS

Las actividades realizadas hasta el momento por este grupo se han enmarcado dentro de 1 proyecto de Investigación de la UNSL, 1 proyecto de la Comunidad Europea, 4 proyectos de Desarrollo Tecnológico de la Secretaría de Políticas Universitarias y 3 proyectos de Extensión de Interés Social. Como consecuencia del trabajo elaborado, se ha logrado desarrollar varios sistemas relacionados a RV con interacción multimodal involucrando aspectos verbales y gestuales, así como también se han elaborado aplicaciones de innovación y desarrollo asociadas a la RV y RA.

Actualmente las acciones se encuentran focalizadas en la incorporación de nuevas estrategias que permitan alcanzar una mejor percepción e interacción, al mismo tiempo que se logra incentivar al usuario tanto cognitiva como físicamente.

Además, se analizará el impacto logrado por el uso de la RV, RA y RM en la adquisición y mejora de habilidades funcionales en procesos físicos y cognitivos asociados a diversas áreas de aplicación para el metaverso. De esta manera, se pretende detectar y evaluar la evidencia científica resultante para determinar la envergadura de dichas intervenciones.

4. FORMACIÓN DE RECURSOS HUMANOS

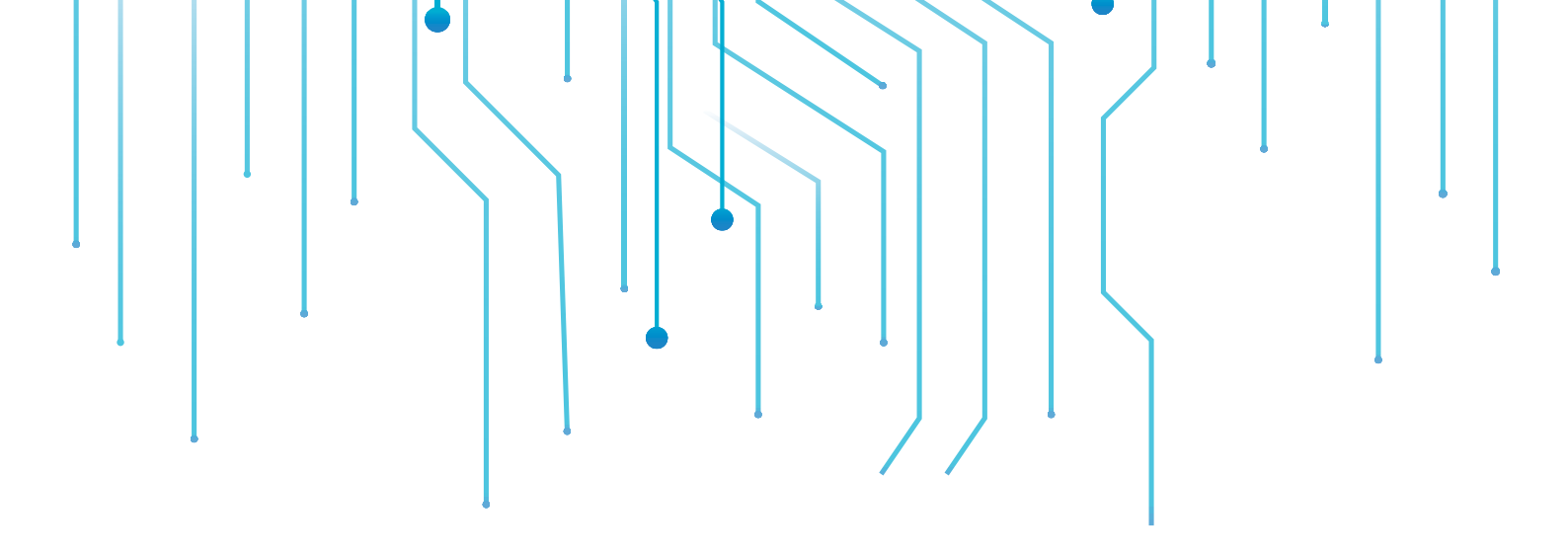
Los trabajos realizados han permitido la definición de trabajos finales de carrera de la Licenciatura en Cs. de la Computación (4 finalizados), tesis de Especialización en Educación Superior (1 finalizada), tesis de Maestría en Cs. de la Computación (1 finalizada y 2 en ejecución) y tesis de Doctorado en Ciencias de la Computación (1 en ejecución).

Adicionalmente se ha obtenido una beca de iniciación a la investigación y una beca de perfeccionamiento en investigaciones otorgadas por la Secretaría de Ciencia y Técnica de la UNSL; y una beca doctoral de CONICET.

5. BIBLIOGRAFÍA

1. Almenara, J. et al. Ecosistema de tecnologías emergentes: realidad aumentada, virtual y mixta. Revista Tecnología, Ciencia y Educación, 2022, no 23, p. 7-22.
2. Morimoto, T. et al. XR (Extended Reality: Virtual Reality, Augmented Reality, Mixed Reality) Technology in Spine Medicine: Status Quo and Quo Vadis. J. Clin. Med. 2022, 11, 470.
3. Sparkes, Matthew. What is a metaverse. New Scientist, 2021, 251(3348), 18.

4. Ortega-Rodríguez, P. From extended reality to the metaverse: A critical reflection on contributions to education. *Teoría de la Educación. Revista Interuniversitaria*, 2022, vol. 34, no 2.
5. Faraboschi, P. et al. Virtual worlds (Metaverse): From skepticism, to fear, to immersive opportunities. *Computer*, 2022, vol. 55, no 10, p. 100-106.
6. Chung, J.. Metaverse XR Components. En *Emerging Metaverse XR and Video Multimedia Technologies: Modern Streaming and Multimedia Systems and Applications*. Berkeley, CA: Apress, 2022. p. 61-98.
7. Park, S. et al. A metaverse: taxonomy, components, applications, and open challenges. *IEEE access*, 2022, vol. 10, p. 4209-4251.
8. Contreras, G. S. et al. (2022). The importance of the application of the metaverse in education. *Modern Applied Science*, 16(3), 1-34.
9. Manrique Morante, H. H. (2021). Simulador de realidad virtual como soporte al desarrollo de prácticas académicas en cirugía laparoscópica de los estudiantes de medicina.
10. Ordoñez Velecela, K. J., & Jácome Parra, E. R. (2018). Simuladores inmersivos de negocios basados en tecnologías de realidad virtual y ambientes 360° para la gestión empresarial (Bachelor's thesis, Universidad de Guayaquil Facultad de Ciencias Administrativas).
11. Azuma, R. (1997). *A Survey of Augmented Reality*. Presence: Teleoperators & Virtual Environments.
12. H. Lin et al , "Metaverse in education: Opportunities and challenges," in *IEEE International Conference on Big Data*. IEEE, 2022, pp. 1–10.
13. Ahmadi M. et al. A revolution in health: Opportunities and challenges of the Metaverse. *EXCLI J.* 2022 May 31;21:791-792.
14. Periyasami, S.; Periyasamy, A.P. Metaverse as Future Promising Platform Business Model: Case Study on Fashion Value Chain. *Businesses* 2022, 2, 527–545
15. Buhalis, D.; Leung, D.; Lin, M. Metaverse as a disruptive technology revolutionising tourism management and marketing. *Tour. Manag.* 2023, 97, 104724.
16. Zaheer A. et al. The metaverse as a virtual form of smart cities: Opportunities and challenges for environmental, economic, and social sustainability in urban futures. *Smart Cities*, 5(3):771–801, 2022.
17. Rokhsaritalemi, S. et al. A review on mixed reality: Current trends, challenges and prospects. *Applied Sciences*, 2020, vol. 10, no 2, p. 636.
18. Donally, Jaime. Learning transported: Augmented, virtual and mixed reality for all classrooms. *International Society for Technology in Education*, 2022.
19. Hughes, Charles E., et al. Mixed reality in education, entertainment, and training. *IEEE computer graphics and applications*, 2005, vol. 25, no 6, p. 24-30.
20. Sahija, Dhaval. Critical review of mixed reality integration with medical devices for patientcare. 2022, 2022, vol. 10.
21. Trunfio, Mariapina; Campana,, Salvatore. A visitors' experience model for mixed reality in the museum. *Current Issues in Tourism*, 2020, vol. 23, no 9, p. 1053-1058.
22. Hammady, Ramy, et al. Design and development of a spatial mixed reality touring guide to the Egyptian museum. *Multimedia Tools and Applications*, 2020, vol. 79, p. 3465-3494.
23. Chai, Jackey JK, et al. Augmented/mixed reality technologies for food: A review. *Trends in Food Science & Technology*, 2022.



IEI

Innovación en Educación en Informática

ADOPCIÓN DE VERILOG EN MATERIAS DE DISEÑO DIGITAL

Gho, Edgardo^{1,2} Carlos Maidana¹ Jair Hnatiuk¹

¹ Grupo de investigación en lógica programable, Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza, Buenos Aires, Argentina.

{egho,cmaidana,jhnatiuk} @unlam.edu.ar

² Universidad Abierta Interamericana. Facultad de Tecnología informática. Centro de Altos Estudios en Tecnología Informática. Buenos Aires, Argentina.

RESUMEN

El presente trabajo describe la migración realizada por el grupo de investigación en lógica programable (GILP-UNLaM) del lenguaje descriptivo de hardware VHDL hacia Verilog. Esto se realizó en el marco de un proyecto de investigación sobre la arquitectura RISC-V, cuyo objetivo consistió en migrar el núcleo de un procesador RISC-V previamente elaborado por dicho grupo en VHDL a Verilog. El desarrollo del procesador RISC-Vp escrito en VHDL fue limitado en ciertas pruebas debido a que las herramientas de desarrollo y depuración ofrecen algunos análisis solo a proyectos diseñados usando Verilog. Surgió entonces la necesidad de migrar el desarrollo a Verilog y adoptar el mismo como lenguaje predeterminado para nuevos diseños. Como resultado intermedio de esta migración se elaboró un curso de aprendizaje de diseño digital utilizando Verilog como lenguaje descriptivo de hardware. Se analizan la necesidad de dicho curso y los objetivos que busca alcanzar el mismo. Actualmente se encuentra incorporado dentro de la currícula de Ingeniería en Electrónica en la materia de programación de hardware.

Palabras Clave: Verilog, HDL, VHDL, RISC-V, FPGA

CONTEXTO

La presente investigación surge de los proyectos I/D del Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM C2-ING-082 “Implementación de la expansión de un procesador RISC-Vp en un entorno de desarrollo de lógica programable” y de su proyecto antecesor del programa PROINCE C-219 “Desarrollo e implementación de una arquitectura basado en el conjunto de instrucciones RISC-V”. Los proyectos fueron llevados a cabo por el Grupo de Investigación en Lógica Programable (GILP) y se vinculan con la materia de la carrera de Ingeniería en Electrónica “Programación de Hardware” y la materia “Arquitectura de Computadoras” de la carrera de Ingeniería en informática del mencionado departamento.

1. INTRODUCCIÓN

Los lenguajes descriptivos de hardware son una herramienta muy útil a la hora de describir, documentar y simular un circuito lógico. Desde su introducción producto de una necesidad en la década de los 80s[1] han reemplazado otros métodos tradicionales de diseño, como la captura de esquemáticos [2] o métodos fotográficos. Estos métodos tradicionales suelen ser engorrosos a la hora de

documentar el funcionamiento de circuitos. En el caso de esquemáticos gráficos, si bien utilizan una representación visual del funcionamiento de un circuito modelado con compuertas que son entendidas por cualquier diseñador, cuando crece la complejidad del diseño por la gran cantidad de compuertas que representa, se hace difícil plasmar la funcionalidad completa en un tamaño reducido y manejable. Algo similar ocurre cuando se representa un algoritmo complejo utilizando diagramas de flujo de datos, que visualmente explican el funcionamiento del algoritmo, pero si los mismos se extienden por varias hojas se pierde fácilmente el hilo de ejecución del algoritmo. En estos casos describir el algoritmo en pseudocódigo demuestra ser una alternativa superadora, ya que brinda al lector una descripción a alto nivel del funcionamiento en un tamaño reducido. De forma similar los lenguajes descriptivos de hardware, que utilizan un lenguaje para describir, implementar y simular circuitos lógicos, demuestran ser más prácticos que las representaciones visuales. Más aún ciertos lenguajes descriptivos permiten hacer verificaciones sobre el funcionamiento con aserciones ([3], [4]) que validan casos de prueba unitarios sobre módulos o conjuntos de módulos permitiendo con una sola herramienta hacer cobertura de código.

Incluir un circuito lógico programable como un FPGA en un diseño permite ciertas modificaciones que serían imposibles de no utilizar estos circuitos [5]. Aunque en la última década se vió una tendencia en el mercado hacia abandonar los circuitos lógicos programables como componente principal en favor de integrar diversos componentes en un único chip (SoC), el uso de aceleradores basados en lógica programable se vio incrementado. Un claro ejemplo es el procesamiento de algoritmos de hashing -

como el SHA3 utilizado en criptomonedas- que puede ser implementado de forma eficiente utilizando circuitos lógicos programables obteniendo una performance muy superior a su contraparte software [6], [7] . Por los motivos expuestos cualquier carrera de ingeniería orientada al diseño de sistemas, como electrónica o informática, se verá beneficiada al incluir en sus contenidos el diseño de circuitos utilizando herramientas de lógica programable. Si bien existen muchas opciones a la hora de buscar un lenguaje descriptivo de hardware, existen dos que predominan por su nivel de adopción en el mercado [8], a saber, VHDL y SystemVerilog. Su origen es distinto, ya que VHDL nace como lenguaje formal auspiciado por la necesidad de estandarizar circuitos provistos al Departamento de Defensa de Estados Unidos y Verilog (luego expandido a SystemVerilog) surge como herramienta de-facto del mercado para la simulación y luego síntesis de circuitos. No obstante, ambos se encuentran estandarizados por la IEEE y cuentan con soporte por parte de la gran mayoría de fabricantes de circuitos lógicos programables. En el caso particular de Verilog, este estándar fue reemplazado por SystemVerilog en 2009 [9] , el que incluye el Verilog tradicional agregando nuevas funciones modernas que extienden el uso de este lenguaje. Cualquiera de las estructuras de Verilog se encuentra soportada en SystemVerilog.

2. LÍNEA DE INVESTIGACIÓN / DESARROLLO

A lo largo del desarrollo del núcleo RISC-Vp [10] escrito en VHDL surgieron problemas fruto de la utilización del software Vivado, ya que los modelos de los chips necesarios para hacer simulaciones post-síntesis y post-implementación no se encuentran disponibles

para VHDL [11]. Esto es una limitación de la biblioteca SIMPRIM [12] que solo posee modelos de retardos de tiempos en Verilog. Dado que al diseñar una unidad central de procesos para correr dentro de un circuito lógico programable es imperativo verificar el cumplimiento de las restricciones de tiempos tal que todo funcione de forma sincrónica, no disponer de esta herramienta en VHDL generó serias complicaciones. Es por ello que se planteó un proyecto de investigación con el fin de migrar el diseño del núcleo RISC-Vp a Verilog. Dentro de este proyecto se logró:

- Conocer la sintaxis de Verilog
- Discernir las diferencias entre Verilog y VHDL
- Implementar circuitos lógicos utilizando Verilog en las herramientas de desarrollo de Xilinx.

Considerando que el grupo de investigación posee larga experiencia utilizando VHDL en aplicaciones prácticas pero también como herramienta en el dictado de clases y cursos en lógica programable, se procedió a diseñar un curso introductorio al lenguaje Verilog que pudiese ser dictado a alumnos de la carrera de Ingeniería en Electrónica [13]–[16]. Se decidió utilizar Verilog en lugar de SystemVerilog ya que es el conjunto mínimo soportado por la mayoría de las herramientas. Es digno de mención que ciertos kits de desarrollo más antiguos solo funcionan con herramientas de programación discontinuadas, las cuales soportan Verilog pero no necesariamente SystemVerilog. Por cuanto los mencionados kits discontinuados se siguen empleando en ciertos ámbitos, debido principalmente por el alto costo de reemplazarlos, se consideró conveniente preparar un curso que soportara un lenguaje compatible con todos los kits.

Teniendo como base la experiencia de haber dictado cursos similares utilizando VHDL se propusieron los siguientes objetivos:

- Introducir al alumno en la historia del diseño digital.
- Plantear la necesidad de los lenguajes descriptivos de hardware como alternativa superadora a la captura digital de esquemáticos.
- Describir el funcionamiento de los circuitos lógicos programables.
- Presentar de forma concreta y resumida los elementos mínimos para realizar un diseño elemental y de baja complejidad como introducción a la sintaxis y las herramientas de desarrollo en Verilog.
- Incrementar el nivel de conocimiento agregando herramientas tales como simulaciones y funciones de ayuda provistas por el lenguaje.
- Realizar prácticas de circuitos combinacionales.
- Introducir circuitos secuenciales y los distintos tipos de asignaciones provistas por el lenguaje.
- Describir los elementos más avanzados del lenguaje, como la instanciación de elementos físicos, secuencias encadenadas y parametrización de módulos.
- Introducir al diseño e implementación de módulos de propiedad intelectual (IP Cores).
- Estudiar circuitos complejos como máquinas de estado.

3. RESULTADOS OBTENIDOS

Los resultados del proyecto de investigación han sido satisfactorios[17]. Se diseñó el curso y el mismo fue utilizado por primera vez durante el primer cuatrimestre de 2022 en la materia Programación de Hardware de la carrera Ingeniería en Electrónica. Sobre el mismo se hicieron ciertas mejoras, particularmente al abordar las máquinas de estado finito, que fueron utilizadas en el segundo cuatrimestre de 2022. El material de apoyo visual se encuentra publicado. El curso se dicta en laboratorio con acceso al software de Xilinx Vivado, versión 2019 en adelante, y kits de desarrollo de la familia FPGA Xilinx series 7, como el Artix 7 o Spartan 7. A los alumnos se les exige realizar un proyecto integrador donde demuestren sus conocimientos del tema y utilicen algunos de los componentes más avanzados de los kits de desarrollo. El objetivo es cubrir la brecha que existe entre proyectos realizables con un microcontrolador de bajo porte versus proyectos involucrando un sistema de computación completo basado en microprocesadores con costos elevados donde se depende de un sistema operativo. Considerando que la aplicación más factible de circuitos lógicos programables es en aceleradores de procesos, se busca que los alumnos implementen interfaces a medida o algoritmos traducidos de software a hardware.

La meta para la siguiente etapa consiste en agregar nuevo material focalizándose en el diseño sincrónico, considerando las líneas de retardo de tiempos para mantener la sincronía dentro de los FPGA. Las ventajas de utilizar SystemVerilog quedan planteadas para una versión avanzada del curso en un futuro proyecto de investigación.

Dada la familiaridad del grupo de investigación tanto con VHDL como con

Verilog, y el hecho de haber realizados similares diseños en ambos lenguajes así como haber dictado cursos en los mismos, creemos importante destacar que Verilog es un típico lenguaje simple de escribir, es decir, se logra escribir mucho código en un tiempo reducido ya que la sintaxis no es tan estructurada como VHDL. Sin embargo, a la hora de leer el código VHDL resulta más simple de comprender en una primera mirada, teniendo características de ser un lenguaje difícil de escribir pero sencillo de leer. Esta impresión está en línea con lo planteado por [18] sobre la inspiración de lenguajes de programación tradicionales como lo es el lenguaje C que se observa en Verilog, mientras que VHDL sigue una estructura en la línea de otros lenguajes fuertemente tipados como ADA.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto trabajan docentes investigadores del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza. La mayoría son docentes en carreras tanto de Informática como de Electrónica. Uno de los miembros se encuentra realizando un Doctorado en Informática en la Universidad Abierta Interamericana sobre Arquitecturas de Computadora utilizando como herramienta de prototipado y pruebas Verilog para el modelado de arquitecturas basadas en el set de instrucciones RISC-V.

5. BIBLIOGRAFÍA

- [1] M. Meredith y S. Svoboda, «The Next IC Design Methodology Transition Is Long Overdue», *Open System C Initiative*, 1 de febrero de 2010. : https://www.accellera.org/images/resources/articles/icdesigntrans/ic_design_transition_feb_2010.pdf
- [2] R. Lin *et al.*, «Beyond schematic capture: Meaningful abstractions for better electronics design tools», en *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.
- [3] O. Amin, Y. Ramzy, O. Ibrahim, A. Fouad, K. Mohamed, y M. Abdelsalam, «System verilog assertions synthesis based compiler», en *2016 17th International Workshop on Microprocessor and SOC Test and Verification (MTV)*, 2016.
- [4] Y. Li, W. Wu, L. Hou, y H. Cheng, «A Study on the Assertion-Based Verification of Digital IC», en *2009 Second International Conference on Information and Computing Science*, may 2009, vol. 2, doi: 10.1109/ICIC.2009.114.
- [5] F. Fallahlalehzari, «How does the Mars Perseverance rover benefit from FPGAs as the main processing units?» <https://www.aldec.com/en/company/blog/188--how-does-the-mars-perseverance-rover-benefit-from-fpgas-as-the-main-processing-units#:~:text=FPGAs%20are%20used%20in%20newer,of%20the%20main%20processing%20units>.
- [6] Y. Jararweh, L. Tawalbeh, H. Tawalbeh, y A. Moh'd, «Hardware Performance Evaluation of SHA-3 Candidate Algorithms», vol. 2012, abr. 2012, doi: 10.4236/jis.2012.32008.
- [7] C. Bittware, «CVP-13 FPGA Powered Cryptocurrency Mining Card with Xilinx Virtex UltraScale+ VU13P FPGA», *BittWare*, <https://www.bittware.com/cvp-13/>.
- [8] H. Foster, «2020 Wilson Research Group Functional Verification Study», octubre de 2020. <https://uobdv.github.io/Design-Verification/WilsonResearchGroupFunctionalVerificationStudy/2020-WRGFV-Study/2020-WrG-FV-Study-Webinar-Oct13.pdf>
- [9] «IEEE Standard for SystemVerilog—Unified Hardware Design, Specification, and Verification Language», *IEEE Std 1800-2017 Revis. IEEE Std 1800-2012*, feb. 2018, doi: 10.1109/IEEESTD.2018.8299595.
- [10] E. Gho, «RiscVP». 1 de marzo de 2022. <https://github.com/edgardogho/RiscVP>
- [11] Xilinx, «Vivado Design Suite User Guide», junio de 2016. <https://docs.xilinx.com/r/en-US/ug900-vivado-logic-simulation>.
- [12] I. Tuzov, J.-C. Ruiz, y D. de Andrés, «Accurately Simulating the Effects of Faults in VHDL Models Described at the Implementation-Level», en *2017 13th European Dependable Computing Conference (EDCC)*, sep. 2017, doi: 10.1109/EDCC.2017.26.
- [13] E. A. De Maria, E. Gho, C. E. Maidana, C. A. Rodriguez, F. I. Szklanny, y H. R. Tantignone, «Real Time FPGA based Thresholding Segmentation in a Multi Touch System», en *2008 4th Southern Conference on Programmable Logic*, 2008.
- [14] E. A. De Maria, E. Gho, C. E. Maidana, F. I. Szklanny, y H. R. Tantignone, «A low cost FPGA based USB device core», en *2008 4th Southern Conference on Programmable Logic*, 2008.
- [15] F. I. Szklanny *et al.*, «Utilización de dispositivos y sistemas de lógica programable en sistemas de control numérico para aplicaciones industriales.»
- [16] F. I. Szklanny *et al.*, «Introducción al diseño con VHDL», en *II Congreso de Microelectrónica Aplicada (μEA 2011)(La Plata, 7 al 9 de septiembre de 2011)*, 2011.
- [17] E. Gho, «CursoVerilog», *GitHub*, <https://github.com/edgardogho/CursoVerilog>
- [18] S. Bailey, «Comparison of vhdl, verilog and systemverilog», *Available Download Www Model Com*, 2003.

APLICACIÓN CON REALIDAD AUMENTADA PARA EL APRENDIZAJE DE LOS CONCEPTOS DE PROGRAMACIÓN

Acosta, Denis L. y Álvarez, Margarita M.

Instituto de Investigaciones en Informática y Sistemas de Información
Departamento de Informática - Facultad de Ciencias Exactas y Tecnologías
Universidad Nacional de Santiago del Estero (UNSE)
e-mail: denislionelacosta@gmail.com; alvarez@unse.edu.ar

RESUMEN

Las actividades de aprendizaje asociadas a los conceptos básicos de programación son consideradas a menudo una tarea compleja debido al alto grado de dificultad involucradas en ellas. Algunas investigaciones indican que las razones para no lograr los niveles de aprendizaje deseados, pueden estar relacionadas con determinadas características que suceden dentro del aula y con ciertas habilidades cognitivas que son relevantes al momento del aprendizaje de dichos conceptos. Entre ellas, está la capacidad de abstracción, una buena aptitud lógica-matemática y la facilidad para la resolución de problemas de orden algorítmico. Ante esta problemática, se propone el desarrollo de una aplicación de aprendizaje ubicuo (AU) con realidad aumentada (RA) que apoye en el aprendizaje de los conceptos básicos de programación. La misma serviría para iniciar a los estudiantes en el aprendizaje de los conceptos básicos, mostrándole un objeto, empleando técnicas de RA y el algoritmo que realiza cuando se desplaza en un determinado ambiente en búsqueda del mismo. Para reforzar la experiencia, la aplicación le permitirá interactuar con un objeto de aprendizaje (OA) para evaluar su conocimiento.

Palabras clave: Aprendizaje ubicuo, Realidad aumentada, Estructuras de Control, Geolocalización.

CONTEXTO

Esta línea de investigación se desarrollará en el marco del proyecto de investigación “Modelos basados en Inteligencia Artificial y

Computación Ubicua para la resolución de Problemas en Educación y otros dominios”, correspondiente a la convocatoria 2021 de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SICYT - UNSE).

1. INTRODUCCIÓN

El aprendizaje de los conceptos básicos de programación, como las estructuras de control, son consideradas difíciles debido a su complejidad y al nivel de abstracción requerido por el estudiante (Insuasti, 2016). La enseñanza de los fundamentos de programación, es un punto clave en la formación de profesionales en Ciencias Computacionales. Pero existen evidencias que sugieren que el aprender a programar no es una tarea fácil (Soloway & Spohrer, 1989). Estos autores afirman que la creación y el control de ambientes y soluciones computacionales a través de la programación, son cosas que para un individuo pueden ser difíciles de realizar. La prueba de esta afirmación puede ser constatada a través de las altas tasas de deserción y cancelación de cursos de fundamentos de programación que numerosos estudios han demostrado (Carter & Jenkins, 2002), (Azad & Mensch, 2008) y (Kelleher & Pausch, 2005), todos teniendo un común denominador en materia de aprendizaje: la forma de abordar los saberes.

En ocasiones, se debe ver a la programación como un ejercicio educativo que es mejor aprendido en un ambiente de juego, diseñado para ilustrar los conceptos seleccionados con configuraciones sencillas. En el que, mediante acciones limitadas, sea posible aprender rutinas simples como condicionales, ciclos e

iteraciones (Hartmann, Nievergelt, & Reichert, 2001). Así mismo, la articulación del mundo real con los fundamentos de la programación, es un elemento crucial para mejorar el aprendizaje de los objetos del saber y, una de las tecnologías que permite integrar el mundo físico y el virtual es la RA.

La RA es una tecnología que enriquece la percepción y la interacción del usuario con el mundo físico, al combinar elementos virtuales en 3D que parecen coexistir en el espacio real. Esto permite que los objetos virtuales se superpongan al mundo físico, mejorando la experiencia de los usuarios y proporcionando una mayor comprensión del entorno en el que se encuentran (Azuma, y otros, 2001). Por otro lado, el AU puede entenderse como aquel en el que la adquisición de nuevos conocimientos puede producirse en cualquier situación en la que se encuentre el sujeto, independientemente del momento y lugar (Burbules, 2012). El AU permite a los estudiantes acceder a todo tipo de información, a partir de la interacción con los objetos circundantes, sean físico o virtuales (Villalustre Martínez & del Moral Pérez, 2018). Resitúa la ubicación del aprendizaje, tanto dentro como fuera del aula, y se adapta a las características de los estudiantes y su contexto, ofreciendo contenidos, presentaciones y actividades personalizadas. La RA, se usa generalmente con las tecnologías móviles y ubicuas y han encontrado en el campo de la educación un área muy rica de aplicación. La combinación de los dispositivos móviles y portátiles, la RA y el AU proporcionan un aprendizaje inmersivo, experiencias enriquecidas, situadas y fluidas (Bozkurt, 2017).

Además, la utilización de la tecnología en el campo educativo aparece como una de las principales alternativas para mejorar la calidad de la enseñanza. En este contexto, la generación de un espacio virtual, y la utilización de material de apoyo, compuesto por textos, imágenes y videos, es una forma de aprovechar las nuevas tecnologías y fomentar la exploración activa y una construcción personal del conocimiento (Salazar & Durán, 2014). Una propuesta interesante para esto,

son los OA, que son: un conjunto de recursos reunidos con un propósito educativo, autocontenible e independiente, diseñado y creado en pequeñas unidades digitales que pueden ser adaptadas para maximizar el número de situaciones en que puede ser reutilizado (López de Munain & Rozaingo, 2013).

Un enfoque común en la enseñanza de los fundamentos de la programación es abordar primero lo básico del lenguaje de programación (o los conceptos ligados a estos), para luego guiar a los estudiantes a las estrategias donde se contemple la totalidad del proceso de programación (Insuasti, 2016). Es por ello que, ante la problemática planteada y considerando las ventajas que proporcionan el AU y la RA, en esta investigación se propone el desarrollo de una aplicación de aprendizaje ubicuo con realidad aumentada que apoye en el aprendizaje de los conceptos básicos de programación. Mediante técnicas de RA, la aplicación mostraría al estudiante un objeto que debe localizar; el estudiante caminaría hasta el objeto y, como resultado, la aplicación generaría un programa con las acciones realizadas. Para ello, será necesario un módulo de compilación que tomará como entrada los datos de los sensores del dispositivo móvil del estudiante (como las coordenadas del GPS, el giroscopio y el acelerómetro), que luego de ser procesados y manipulados, los traducirá a un programa escrito en un lenguaje diseñado a tal fin. También, contará de un módulo gestor de 3D que permitirá insertar y manipular los objetos en RA, un módulo de interfaz de usuario y un módulo principal. Finalmente, para reforzar la experiencia, la aplicación contará de un módulo de OA, que permitirá al estudiante interactuar con un OA, el cual deberá pasar satisfactoriamente la evaluación para desbloquear nuevos objetos y/o niveles.

La interacción con la aplicación le proporcionará al estudiante una respuesta inmediata, favoreciendo al alumno la experimentación en el mundo real y, permitiéndole formar rápidamente un modelo mental gracias a las respuestas que arrojará la herramienta.

La aplicación informática resultante será probada con estudiantes ingresantes a la carrera Licenciatura en Sistemas de Información (LSI) y Programador Universitario en Informática (PUI) de la Facultad de Ciencias Exactas y Tecnologías de la Universidad Nacional de Santiago del Estero.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación tiene como objetivo el análisis, diseño y desarrollo de una aplicación, con los módulos que conformen la arquitectura de la misma y, en particular, de un módulo capaz de traducir coordenadas geográficas en un lenguaje de instrucciones atómicas a usarse para tal fin que llamaremos módulo del compilador. Como así también, la exploración y aplicación de herramientas que permiten implementar RA en dispositivos móviles. En consecuencia, el objetivo general es:

1. Contribuir en el proceso de enseñanza-aprendizaje de los conceptos básicos de la programación en los estudiantes que se inician en su estudio.

Además, se ha definido el siguiente objetivo específico:

1. Diseñar y desarrollar una aplicación de AU, que implemente RA, que apoye a la enseñanza-aprendizaje de los conceptos básicos de fundamentos de la programación a estudiantes principiantes en carreras de informática y que permita:
 - Mostrar objetos georreferenciados en función de la ubicación geográfica del estudiante.
 - Trazar el camino recorrido por el estudiante hasta el objeto georreferenciado mediante la traducción de las coordenadas geográficas a instrucciones atómicas.
 - Mostrar los OA y evaluar el conocimiento adquirido por los estudiantes.

Con el propósito de dar cumplimiento a los objetivos planteados, se realizarán las siguientes actividades:

1. *Estudio exploratorio y recopilación del dominio del problema:* a través de la investigación y revisión bibliográfica, se realizará una búsqueda en publicaciones, artículos, libros u otro contenido científico disponible en la web sobre: AU: técnicas y metodologías; RA: herramientas, ventajas y desventajas; OA estructura y características, Compiladores: estructura; Geolocalización: nociones básicas y Desarrollo Móvil: lenguajes de programación, sensores, herramientas y frameworks.
2. *Definición de la arquitectura de la aplicación:* se diseñará la arquitectura de la aplicación y se determinará cuales serán cada uno de los módulos que conformarán la misma; cuáles serán sus componentes y cómo interactuarán entre ellos.
3. *Definición, análisis y desarrollo del compilador:* para ello se realizará la definición de los lenguajes fuente y objeto; diseño de las expresiones regulares y de la gramática libre de contexto; definición de las frases del compilador que se implementaran; definición del mecanismo de traducción apropiado entre los lenguajes; desarrollo de las fases del compilador y prueba del compilador.
4. *Definición del módulo de gestión 3D y gestión de OA:* se deberá definir un formato de archivos compatible y un gestor de almacenamiento que actúe como un repositorio de los mismos. En cuanto a los OA, al igual que los objetos 3D, se requiere de un flujo de datos que permita persistir a los mismos, desarrollar una herramienta que soporte su visualización en un dispositivo móvil y permita la interacción con el mismo.
5. *Definición del módulo de sensores:* se definirán las interfaces que se comunicarán con los distintos sensores (GPS, giroscopio, acelerómetro), la compatibilidad con los dispositivos y el formato de los datos que se obtendrán de ellos.
6. *Definición del módulo de interfaz de usuario:* en esta etapa, se realizarán diagramas de alto nivel (o wireframes) utilizando herramientas de diseño de interfaces de usuario que reflejen los

aspectos esenciales de la aplicación. Se deberá decidir qué componentes visuales son los adecuados y cuales son compatibles con la mayoría de los dispositivos Android.

7. *Implementación de la aplicación*: se implementará, en la aplicación, cada una de las etapas definidas previamente y cada uno de los módulos. Haciendo énfasis en los módulos del compilador, gestión 3D y gestión de OA.
8. *Testeo/Validación de la aplicación*: mediante el uso de casos de pruebas, se realizarán la validación de cada uno de los componentes de la aplicación, pruebas unitarias para cada uno de los módulos y pruebas de integración para la aplicación en general.
9. *Prueba de la aplicación en estudiantes*: mediante la evaluación de diferentes indicadores, se busca validar que la herramienta diseñada cumpla o no con los objetivos propuestos. Para ello, se realizará una serie de pruebas a través de alumnos ingresantes de la carrera LSI y PUI de la UNSE que permitan medir la usabilidad de la aplicación y el grado satisfacción con el contenido presentado.
10. Conclusiones finales: se realizará un análisis de los resultados para concluir si la aplicación cumplió con los objetivos establecidos.

3. RESULTADOS ESPERADOS

A partir de la ejecución de esta propuesta, el resultado que se espera obtener, conforme a los objetivos, es una herramienta que permita complementar el aprendizaje de los conceptos básicos de la programación e integre las nociones de AU, RA y OA.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación se integra por un estudiante y la Co-Directora del Proyecto de Investigación indicado en el contexto.

Además, con este trabajo, el estudiante podrá optar por el título de Licenciado en Sistemas

de Información de la Universidad Nacional de Santiago del Estero.

5. BIBLIOGRAFÍA

- Azad, M., & Mensch, S. (2008). Issues and challenges for selecting a programming language in a technology update course. *Proceedings of the Information Systems Education Conference*,. Phoenix, AZ.
- Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., & Macintyre, B. (2001). Recent advances in augmented reality. *Comput Graphics Appl.* 21, 34-47. doi:10.1109/38.963459
- Bozkurt, A. (2017). Augmented Reality with Mobile and Ubiquitous Learning: Immersive, Enriched, Situated, and Seamless Learning Experiences. En *Digital Tools for Seamless Learning* (Vol. 2, págs. 27-41). IGI Global. doi:10.4018/978-1-5225-1692-7.ch002
- Burbules, N. C. (2012). Ubiquitous Learning and the Future of Teaching. Board. doi:https://doi.org/10.24908/eoe-ese-rse.v13i0.4472
- Carter, J., & Jenkins, T. (2002). Gender differences in programming? *Proceedings of the 7th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, (págs. 188-192). Aarhus, Denmark. doi:10.1145/377435.377681
- Hartmann, W., Nievergelt, J., & Reichert, R. (2001). Kara, finite state machines, and the case for programming as part of general education. *Proceedings IEEE Symposia on Human-Centric Computing Languages and Environments*, (págs. 135-141). doi:10.1109/HCC.2001.995251
- Insuasti, J. (2016). Problemas de enseñanza y aprendizaje de los fundamentos de programación. *Revista Educación Y Desarrollo Social*, 2(10), 234-246. doi:10.18359/reds.1966
- Kelleher, C., & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM*

- Computing Surveys*, 37, págs. 83-137.
doi:10.1145/1089733.1089734
- López de Munain, C., & Rozaingo, Z. (2013).
Objetos de Aprendizaje y Simulación en el
proceso de enseñanza-aprendizaje.
*Conferencia Latinoamericana de Objetos y
Tecnologías de Aprendizaje*. 4. Chile:
LACLO 2013.
- Salazar, N., & Durán, E. (2014). Objeto de
aprendizaje para la enseñanza de la
simulación. *IX Congreso sobre Tecnología
en Educación & Educación en Tecnología*,
(págs. 315-324). La Rioja. doi:978-987-
24611-1-9
- Salinas, J., & Marín, V. (2014). Pasado,
presente y futuro del microlearning como
estrategia para el desarrollo profesional.
III(2), 46-61.
- Soloway, E., & Spohrer, J. (1989). *Studying
the Novice Programmer* (1st Edition ed.).
(P. Press, Ed.) New York.
doi:10.4324/9781315808321
- Villaluestre Martínez, L., & del Moral Pérez,
E. (2018). Geolocalización y realidad
aumentada para un aprendizaje ubicuo en la
formación inicial del profesorado. *@tic.
revista d'innovació educativa*(21), 40-48.
doi:DOI:
<https://doi.org/10.7203/attic.21.12633>

Innovación Tecnológica y Metodológica en Educación

Mónica Sarobe¹, Tamara Ahmad¹, Claudia Russo¹, Mariana Adó¹, Nicolás Alonso¹, María Mercedes Guasch¹, Paula Lencina¹, Fernando López Gil¹, María Rosa Piergallini¹, Gustavo Gnazzo¹, Marina Rodríguez¹, Natalia Sinde¹, Guillermo Casanova¹, Matías Dimase¹, Ludmila Falcioni¹, Maricel Molina¹, Santiago Pérez¹, Abril Carini¹, Carla Decoud².

¹ Instituto de Investigación y Transferencia en Tecnología (ITT), UNNOBA-CIC.
Sarmiento N° 1169, 2^{do} Piso, Junín (Bs. As.), Argentina – Te: +54 236 4477-050 INT 11610.

² Facultad Politécnica (FP), Universidad de Asunción (UNA).
Acceso peatonal de Mcal. Estigarribia a la UNA, San Lorenzo, Paraguay – Te: +595 21 588 7000.

monica.sarobe, tamara.ahmad, claudia.russo, mariana.ado, nicolas.alonso, mercedes.guasch, paula.lencina, fernando.lopezgil, rosana.piergallini, gustavo.gnazzo, marina.rodriguez, natalia.sinde, guillermo.casanova, matias.dimase, ludmila.falcioni, maricel.molina, santiago.perez, [@itt.unnoba.edu.ar](mailto:abril.carini), carladecoud@gmail.com

Resumen

Este proyecto de investigación, desarrollo e innovación (I+D+i) se centra tanto en las tecnologías innovadoras vinculadas a la educación digital como en los procesos de evaluación de dichas tecnologías aplicadas a la educación. En el marco del presente y considerando que las tecnologías actuales implican un impacto en los modelos educativos, se investiga sobre realidad aumentada (RA), realidad virtual (RV), hologramas y escenarios híbridos y se indaga en la mejora de procesos educativos a partir de la aplicación de las tecnologías mencionadas, sobre las competencias digitales docentes y la calidad y evaluación de la educación digital.

Palabras claves: Innovación tecnológica; Innovación metodológica; Educación; Educación digital; Educación a Distancia; Realidad Virtual; Realidad Aumentada; Hologramas; Escenarios híbridos.

Contexto

Como viene sucediendo en los últimos años, las tecnologías siguen teniendo un impacto significativo en todos los ámbitos de nuestras vidas y la educación no es la excepción, por lo que se hace imprescindible evaluar la calidad de los procesos que se derivan de la utilización de las mismas.

Los repentinos escenarios en los que se desarrolló la educación durante la crisis sanitaria por COVID-19 determinaron la importancia de la definición e implementación de modelos híbridos de aprendizaje en Educación Superior. La educación híbrida es el genuino resultado de la introducción de las pedagogías informáticas en la educación presencial y la articulación de múltiples modalidades educativas no fragmentadas. No es simplemente una modalidad de la enseñanza presencial con apoyo en plataformas digitales (Rama, 2021: 120) [1], sino más bien una

integración indisoluble de la presencia y la no presencia en las aulas junto a la coherente modificación del diseño de los cursos impartidos y de la planificación docente (Barragán De Anda et al, 2021: 151) [2]. En este sentido, la modalidad híbrida no sólo amplía la accesibilidad, sino que además otorga una flexibilidad temporal, espacial y pedagógica que admite trayectos académicos acordes tanto a las habilidades y a los conocimientos que se desea desarrollar –y que son demandados por los diversos campos profesionales– como a las personas que participan de los procesos de enseñanza-aprendizaje (Rama, 2021: 121. Barragán De Anda et al, 2021: 151) [1][2].

Por otro lado, la introducción de tecnologías que permiten reproducir nuestra realidad o, incluso, crear nuevas realidades donde el usuario se vuelve partícipe activo y puede interactuar con otros usuarios o con objetos de ese mundo virtual (RA, RV, Holografía), están siendo utilizadas cada vez en más sectores de nuestra sociedad, incluyendo en espacios educativos. Sin embargo, no se está evaluando mucho respecto a las mejoras de los procesos y teniendo en cuenta la calidad de los mismos.

1. Introducción

El contexto planteado requiere que se desarrollen y se definan competencias digitales virtuales necesarias para los nuevos escenarios educativos y que se diseñen propuestas mediante una metodología de investigación que permita arribar a conclusiones y/o aplicaciones para la mejora de la calidad de la educación digital.

Este trabajo surge como una evolución de proyectos acreditados en convocatorias a Subsidios de Investigación Bianuales de la UNNOBA como *Tecnologías Emergentes* (SIB 2019), *Tecnologías exponenciales en*

contextos de realidades mixtas e interfaces avanzadas (SIB 2015) y *El desafío de las nuevas tecnologías de la información y las comunicaciones en los contextos educativos* (SIB 2013).

Sus **objetivos generales** son (a) analizar e inferir cómo la innovación tecnológica y metodológica impacta en la educación digital para definir estrategias, competencias digitales y procesos que permitan su utilización con calidad; (b) promover la formación y finalización de los estudios de grado y postgrado de los miembros del proyecto y (c) fomentar la interacción con otros grupos de I+D+i del país y del exterior en la temática del proyecto. Asimismo, son objetivos específicos del primer punto (a.1) analizar la utilización de tecnologías aplicadas en la educación, sus características, funciones, normativas, desarrollos y metodologías aplicadas; (a.2) identificar, adaptar y proponer competencias digitales, teniendo en cuenta a las y los participantes dentro del proceso educativo; (a.3) identificar, definir, diseñar e implementar estrategias de mejora en los procesos educativos digitales; y (a.4) determinar el impacto de la implementación de las estrategias de mejora en los procesos educativos digitales.

Finalmente, la **hipótesis** de trabajo de la cual se parte es que la innovación tecnológica y metodológica impacta en la educación digital y requiere de la definición de estrategias, competencias digitales y procesos que permitan su utilización con calidad.

2. Líneas de investigación y desarrollo

En primer lugar, el proyecto se propone estudiar el impacto que tienen en la educación, las tecnologías que permiten reproducir nuestra realidad o, incluso, crear nuevas realidades (RA, RV, Holografía). En tal

sentido, se analizará el impacto que tienen las tecnologías aplicadas a procesos educativos. En particular el trabajo estará centrado en realidad aumentada, realidad virtual y hologramas. Cuando se habla de RA, podemos decir que el mundo real y el mundo virtual se entremezclan para crear una realidad mixta en tiempo real [3]. Es un recurso tecnológico que ofrece experiencias interactivas al usuario a partir de la combinación entre la dimensión virtual y la física, con la utilización de dispositivos digitales y se aplica en contextos y campos diversos.

Por otra parte, la realidad virtual (RV) es un sistema tecnológico, basado en el empleo de distintos dispositivos, cuyo fin es producir una apariencia de realidad que permita al usuario tener la sensación de estar presente en ella. Su aplicación, aunque centrada inicialmente en el terreno de los videojuegos, se ha extendido a otros muchos campos, como la medicina, educación generando aulas virtuales o simulaciones de vuelo y laboratorios entre otros [4].

Los continuos avances tecnológicos influyen en las diferentes formas de enseñar y aprender. De esta manera, la tecnología se ha convertido en una ayuda indispensable tanto para profesores como estudiantes. Los hologramas, en particular cuando se utilizan en educación, pueden ser empleados para emular la presencia física del profesor, disertante, entre otros; en un espacio físico donde se encuentran las y los estudiantes presencialmente, de manera de generar un acercamiento y una comunicación más fluida.

Esta línea buscará continuar la evaluación del impacto del uso de estas tecnologías en el proceso educativo.

En segundo lugar, el proyecto propone investigar los escenarios híbridos en ámbitos educativos. Los modelos híbridos de aprendizaje combinan la enseñanza presencial

con la educación en línea, creando así modelos educativos más flexibles. El aprendizaje híbrido, semipresencial o combinado, es un modelo de enseñanza y aprendizaje que entrelaza elementos de la clase presencial y la clase virtual o el aprendizaje en línea. Ambas modalidades deben estar plenamente integradas en la propuesta y cuidadosamente planificadas, para que esta metodología funcione de la mejor manera posible. El modelo de aprendizaje híbrido es capaz de brindar una instrucción más personalizada a un mayor número de estudiantes, posibilitando y extendiendo el aula en encuentros que combinan sincronía y asincronía, aprovechando las bondades que brinda Internet. [1][2] Las aplicaciones y diferentes combinaciones constituyen el objeto de estudio de esta segunda línea.

En tercer lugar, el proyecto se propone indagar en las competencias digitales en educación. Son competencias requeridas en la formación del ciudadano del siglo XXI para afrontar el aprendizaje a lo largo de la vida. Sobre este tema es necesario citar el trabajo pionero realizado en el marco del proyecto DeSeCo (OCDE, 2005) en el cual se determinan las competencias clave para el ciudadano del siglo XXI: usar las herramientas de forma interactiva (lenguaje, símbolos, conocimiento, información y tecnologías en general); interactuar con grupos heterogéneos (relacionarse, cooperar, trabajar en equipo y resolver conflictos); y tercero, actuar de manera autónoma (en cualquier contexto, conducir planes de vida y defender derechos e intereses). En la propuesta reelaborada (OCDE, 2009) se definen las tres dimensiones como información, comunicación e impacto ético-social.

De forma general, las definiciones de competencia digital pueden clasificarse en relación a dos grandes perspectivas: por un

lado, las que enfatizan el componente tecnológico; y, por otro lado, las que hacen hincapié en la dimensión informacional o comunicativa (Duran et al, 2019) [5].

En cuarto lugar, el proyecto procura la evaluación del impacto de la utilización de tecnologías y metodologías innovadoras en los procesos educativos. Esta línea nos permitirá analizar, evaluar, definir y redefinir el uso de tecnologías y metodologías innovadoras en los procesos educativos para que los mismos se desarrollen con calidad.

3. Resultados obtenidos/esperados

Se pretende difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes revistas especializadas, congresos, jornadas y workshops, nacionales e internacionales, relacionados con la mejora de los procesos, productos y servicios mediante la innovación.

Se espera también que los aspectos investigados contribuyan a propiciar el fortalecimiento en la formación de recursos humanos, en su rol de investigadores o partícipes activos en equipos de investigación, transferencia e innovación y en colaboración con grupos de investigación de otros proyectos aprobados. Se anhela el fomento de la culminación de estudios superiores de grado y posgrado, la promoción de la escritura, exposición y defensa de Trabajos Finales de Grado y Postgrado, y la realización de Prácticas Profesionales Supervisadas y Tesinas de Grado.

4. Metodología de investigación

La metodología aplicada posee nueve etapas que vienen desarrollándose en estudios anteriores (SIB 2019, 2015, 2013). Estas son:

1. Análisis del estado actual y fundamentos teóricos de los temas propuestos.
2. Realización de talleres o seminarios de discusión internos y externos para identificar problemas tecnológicos. Es importante el intercambio de información con los grupos del país y del exterior que trabajan en proyectos similares.
3. Recopilación y estudio de bibliografía existente, publicaciones realizadas por otros equipos de investigación, revistas y demás producciones digitales disponibles en la Web sobre la temática de estudio.
4. Focalización del objetivo de investigación descriptiva para conocer situaciones, costumbres y actitudes predominantes a través de la descripción exacta de las actividades, objetos, procesos y personas.
5. Investigación de alternativas resolutorias para optimizar el empleo de recursos humanos y tecnológicos y obtener una solución específica para el problema concreto. Esta etapa resulta particularmente importante en la formación de recursos humanos que se inician en I/D y es beneficioso el intercambio con otros grupos con proyectos de investigación similares.
6. Integración del trabajo al equipo de investigación en el marco del proyecto al que pertenece.
7. Análisis de hipótesis, experimentación y obtención de resultados e indicadores.
8. Evaluación de resultados para su eventual publicación y transferencia al medio.
9. Publicación de resultados obtenidos.

5. Formación de recursos humanos

En cuanto a la formación del equipo de trabajo, la mayoría de los integrantes posee una Licenciatura en Sistemas o en Informática, o es Analista de Sistemas, a excepción de una Licenciatura en Ciencias de la Comunicación

Social con orientación en Comunicación y Procesos Educativos. Tanto este contraste como el trabajo colaborativo con otros equipos de investigación apela a la interdisciplinariedad requerida para alcanzar los resultados buscados. Asimismo, el 50% cuenta con estudios de Maestría o Especialización en curso, el 25% es Magister y el 8% posee un Doctorado en Ciencias Informáticas.

En cuanto a los becarios de grado, hay un Analista de Sistemas egresado, dos personas con Licenciaturas en Sistemas en curso, una con Ingeniería en Sistemas y otra con Analista en Sistemas camino a concluir sus estudios. Este trabajo los ayudará en el camino a su titulación.

5. Bibliografía y referencias

[1] RAMA, Claudio (2020). “La nueva educación híbrida”, Cuadernos de Universidades, no. 11 (2021), pp. 139, México D.F.: Unión de Universidades de América Latina y el Caribe (UDUAL). ISBN de la colección: 978-607-8066-35-3. Ficha catalográfica del título de la serie: Cuadernos Universidades, México: UDUAL.

[2] BARRAGÁN DE ANDA, A. B., ÁVILA GONZÁLEZ, C., BELMONTE HERRERA, A. M., CAMARENA CADENA, M. J. y GÓMEZ VALENZUELA, R. (2021). “Ambientes híbridos de aprendizaje en estudios de posgrado,” Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET), no. 28, pp. 149-156, La Plata: Universidad Nacional de La Plata (UNLP). doi: 10.24215/18509959.28.e18.

[3] PÉREZ, S. M., ROBLES, B. F., & OSUNA, J. B. (2021). La realidad aumentada como recurso para la formación en la educación superior. *Campus Virtuales*, 10(1), 9-19.

[4] CASTRO LOZANO, C. de (2012). El futuro de las tecnologías digitales aplicadas al aprendizaje de personas con necesidades educativas especiales. *Revista de Educación a Distancia (RED)*, (32).

[5] DURÁN, M. C., PRENDES, M.P.E. y GUTIÉRREZ, I. P. (2019). Certificación de la Competencia Digital Docente: propuesta para el profesorado universitario. *RIED. Revista Iberoamericana de Educación a Distancia*, 22(1), pp. 187-205. doi: <http://dx.doi.org/10.5944/ried.22.1.22069>.

Evaluación de la Calidad de Objetos de Aprendizaje, para la enseñanza de la Informática, desde el Enfoque de los estudiantes. Medición del significado afectivo

Nevelin Irene Salazar, Elena Beatriz Durán y Germán Lescano

**Instituto de Investigación en Informática y Sistemas de Información (IISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)
Avda. Belgrano (S) 1912 – Santiago del Estero CP 4200
[nsalazar, eduran, gelescano] unse.edu.ar**

RESUMEN

Desde el año 2012, a través de distintos proyectos de investigación, las autoras de este trabajo han venido consolidando la línea de investigación sobre diseño y construcción de Objetos de Aprendizaje (OA) como recursos digitales abiertos para fomentar el aprendizaje activo.

Continuando con esa línea de investigación, en esta propuesta se aborda la evaluación de la calidad de los Objetos de Aprendizaje. Se entiende que la falta de controles de calidad de estos recursos definitivamente afecta el proceso de enseñanza y aprendizaje. Esto, refuerza la necesidad de evaluarlos adecuadamente para asegurar sus niveles de calidad. En consecuencia, en esta investigación se busca diseñar un modelo para evaluar la calidad de los OA, desde el enfoque de los estudiantes empleando el diferencial semántico. El diferencial semántico es una de las técnicas más utilizadas en el llamado “diseño emocional”. La misma aporta información sobre las emociones que un objeto o concepto genera, obteniendo el valor connotativo y captando el significado afectivo que el usuario tiene de él. En particular, la propuesta será evaluada con OA sobre temas de Informática

Palabras clave: *Objetos de Aprendizaje, Evaluación, Calidad, Diferencial Semántico, Computación Ubicua, Técnicas de Inteligencia Artificial.*

CONTEXTO

En esta investigación se presenta una de las líneas de investigación del proyecto “Modelos basados en Inteligencia Artificial y Computación Ubicua para la resolución de Problemas en Educación y otros dominios”, presentado en la convocatoria 2022-2025 en la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT – UNSE). La investigación es una continuación de la línea iniciada en 2012 en el marco de los siguientes proyectos: “Sistemas de Información Web Personalizados, basados en Ontologías, para soporte al Aprendizaje Ubicuo”, CICyT-UNSE, años 2012-2016; “Sistemas Adaptativos Multi-Agente de soporte al Aprendizaje Ubicuo y al Aprendizaje Colaborativo”, PICTO-UNSE, años 2014-2016 y “Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas”, CICyT-UNSE, años 2017-2021.

Resultados previos, obtenidos por el equipo de investigación, tales como (Salazar & Durán, 2014), (Salazar & Durán, 2016), (Salazar & Durán, 2017), (Salazar &

Durán, 2019), servirán de base para alcanzar los objetivos de la presente investigación.

1. INTRODUCCIÓN

En la actualidad el e-learning se ha posicionado como uno de los principales modelos de educación cambiando los procesos tradicionales de enseñanza y de aprendizaje y proporcionando numerosas ventajas de carácter pedagógico y de acceso. Este esquema genera nuevas necesidades tales como el desarrollo y mejora de materiales educativos, exclusivamente expresados como OA. Los OA se distinguen de otros recursos por su predisposición a la reutilización en múltiples contextos, además de su disponibilidad en diferentes ambientes (Morales et al., 2005).

Trabajar con OA se ha convertido en una oportunidad más para identificar el potencial de las TIC como mediadores del aprendizaje (Triana Muñoz et al., 2016). Al respecto, Wong (2012) puntualiza que cuando se utiliza una computadora o material de aprendizaje digital en una situación de aprendizaje se espera que se haga para introducir un valor añadido a este proceso diferente al que se lograría con el material impreso.

Los OA son una herramienta educativa que puede insertarse en propuestas curriculares y metodologías de enseñanza y aprendizaje de diversa índole. Se consolidan como recursos digitales para fomentar el aprendizaje individualizado, independiente del tiempo y lugar. El desarrollo de los OA se basa en una estrategia orientada al aprendizaje del estudiante y, para ello, su desarrollo debe regirse por un estándar para la estructuración de los datos. Es importante realizar acciones que permitan entregar al usuario final recursos de calidad; para ello se requiere definir una serie de parámetros y una escala de evaluación. La razón principal para preocuparse por la calidad en los OA

es que se trata de recursos de apoyo al aprendizaje. Aun cuando no es el único factor que afecta a los procesos de enseñanza y de aprendizaje, la carencia de controles de calidad de los recursos educativos puede afectar el proceso y sus resultados en los estudiantes (Vidal et al., 2008).

El concepto de calidad de OA y el desarrollo de mecanismos apropiados para su evaluación, han ganado relevancia con el fin de garantizar el acceso a recursos de aprendizaje de calidad. Existen numerosas investigaciones que plantean la evaluación de la calidad de los OA desde diferentes enfoques considerando elementos *tecnológicos* (Pinto et al., 2015), (Gordillo et al., 2018), (Cepeda et al., 2017),; *pedagógicos* (Massa et al., 2012a), (Tabares, et al., 2017), (Baldassarri & Alvarez, 2016); *el evaluador*; (Massa, Di Giusti, Pesado, 2012a); (Tabares et al., 2017); ((Pinto et al., 2015); (Baldassarri & Alvarez, 2016), (Gordillo et al., 2018); *los metadatos* (Hughes, 2004), (Bruce & Hillmann, 2004), (Nichols, 2008), (Ochoa; 2008), entre otros. Todos ellos brindando información sobre la calidad de los OA.

Para llevar adelante la evaluación de calidad de los OA, se han diseñado innumerables instrumentos, aunque actualmente no se ha llegado a un consenso de todos los elementos que intervienen en la medición de la calidad de los OA. Se han planteado varias propuestas para establecer diferentes dimensiones y criterios a considerar a la hora de evaluar su calidad.

Particularmente, si nos situamos en los procesos de evaluación de calidad desde el enfoque de los usuarios, Massa (2012b) afirma que realizar una evaluación a un recurso de aprendizaje, desde esta perspectiva, puede definir muchas cuestiones acerca de la valoración del mismo. Por ejemplo, si estos manifiestan dificultades para comprenderlo, o no ha facilitado su aprendizaje a pesar de ser calificado como de alta calidad por expertos en una evaluación heurística. En el instrumento de evaluación de la calidad de

los OA desde la perspectiva de los estudiantes diseñado en esta investigación, se introduce sintéticamente el concepto de Diferencial Semántico (DS) para evaluar la reacción global de los estudiantes en el uso de OAs.

El DS fue desarrollado originalmente por Osgood, Suci y Tannenbaum (1957) para explotar las dimensiones del significado. Pero hoy en día consiste en una serie de adjetivos bipolares que califican a un concepto/objeto de actitud, ante los cuales se solicita la reacción del sujeto, quien debe calificar al concepto/objeto de actitud en un conjunto de adjetivos bipolares. Entre cada par de adjetivos se presentan varias opciones y el sujeto selecciona aquella que refleje su actitud en mayor medida.

El DS es un instrumento de medida para obtener el significado connotativo de objetos, sujetos, situaciones, sucesos a través de conceptos, y es utilizado para determinar la actitud hacia el objeto por el usuario (Osgood et al, 1957). El significado afectivo o connotativo de un objeto, se refiere a que el significado es subjetivo y puede presentar varios significados; donde la connotación depende del contexto. Luego, el significado afectivo o connotativo no comunica únicamente información, sino que también aporta sentimientos y sensaciones.

Estudiar la actitud de los estudiantes frente a los OA a través del DS, permitiría obtener información sobre como los estudiantes perciben el proceso de enseñanza y aprendizaje a través de los OA. De esta manera para un estudiante aprender a través de OA puede ser fácil o difícil, dinámico o estático, motivador o desmotivador, aburrido o divertido entre otras características, según la escala de DS que se construya.

Por lo tanto, centrarse en la calidad del OA desde la perspectiva de los usuarios pretende conocer si el OA ha sido adecuado como unidad de enseñanza, y deben definirse una serie de criterios o indicadores

que se correspondan con la valoración buscada.

Por lo expuesto, en esta investigación se plantea desarrollar un instrumento de evaluación de la calidad de los OA desde el enfoque de los estudiantes, con el fin de determinar la satisfacción del usuario con el uso de OAs. La propuesta será validada con OAs contruidos para el aprendizaje de temas de Informática del nivel universitario.

Cabe destacar que no se han encontrado trabajos vinculados con el desarrollo de instrumentos de evaluación de calidad de los OA que incluyan como dimensión de evaluación el significado afectivo lo cual demuestra la originalidad de esta propuesta.

2. DESCRIPCIÓN DE LA LINEA DE INVESTIGACION

La pregunta que guiará el desarrollo de la investigación es la siguiente: *¿es posible evaluar la calidad de los OA analizando la dimensión afectiva de las opiniones emitidas por los alumnos?*

Tomando en cuenta los interrogantes enunciados, se establece como objetivo general de esta investigación: Evaluar la calidad de los OA mediante el análisis de la dimensión afectiva de las opiniones emitidas por los alumnos empleando el diferencial semántico.

El objetivo general de esta propuesta pretende ser alcanzado sobre la base de los siguientes objetivos específicos:

- Elaborar un marco de referencia sobre indicadores que definen la calidad de los OA, y en particular desde la perspectiva de los estudiantes.
- Crear un instrumento de evaluación basado en el DS, que permita medir el significado afectivo de las opiniones emitidas por los estudiantes de carreras de Informática hacia los OA.
- Validar experimentalmente el instrumento de evaluación de calidad de OA propuesto.

Finalmente se proponen las siguientes actividades generales para el desarrollo de la investigación:

- Búsqueda de bibliografía: en esta actividad se buscará material bibliográfico sobre antecedentes respecto a la evaluación de la calidad de los OA desde el enfoque de los usuarios, instrumentos diseñados por otros autores identificando y analizando los elementos que constituyen la calidad desde este enfoque.
- Construcción del modelo de evaluación de calidad de OA: en función de los resultados obtenidos en la actividad anterior, se realizará un estudio, análisis y selección de indicadores para medir la calidad de los OA. También se generará la lista de adjetivos bipolares resultantes de la aplicación del DS a los indicadores definidos.
- Trabajo experimental: Construcción y evaluación del modelo de indicadores basados en el DS, para medir la calidad de los OA. Para ello se realizarán experimentaciones con estudiantes en las diferentes asignaturas de carreras de Informática de la Universidad Nacional de Santiago del Estero, donde se empleen OA como recursos educativos.
- Evaluación de Resultados: se realizará el procesamiento de los datos resultantes del trabajo experimental. Luego se analizarán posibles técnicas que permitan validar el instrumento creado.
- Divulgación de resultados: en esta etapa presentará los resultados en eventos científicos nacionales e internacionales y publicaciones en revistas indexadas.

3. RESULTADOS ESPERADOS Y OBJETIVOS

Esta investigación pretende favorecer el desarrollo de conocimiento científico-tecnológico de relevancia que contribuya a

la evaluación de calidad de los Objetos de Aprendizaje en el área de la enseñanza de la informática.

Algunos de los resultados a los que se espera arribar al concretar este trabajo son:

- Estado del arte referido a aspectos de calidad de los OA desde el enfoque de estudiantes.
- Un modelo de indicadores para la evaluación de la calidad de recursos de aprendizaje, específicamente Objetos de Aprendizaje.
- Resultados experimentales respecto a la validación de la propuesta.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación está compuesto por dos docentes de investigación formados y otra en formación, quien está desarrollando su tesis para obtener el título de Magister en Informática Educativa en el marco de esta investigación.

5. REFERENCIAS

Baldassarri, S. & Álvarez, P. (2016). M-eRoDes: Involucrando a los estudiantes en la creación y evaluación colaborativa de objetos de aprendizaje. In: Actas de las XXII Jornadas sobre la Enseñanza Universitaria de la Informática, pp. 195–202, Universidad de Almería.

Bruce, T. & Hillmann, D. (2004) The Continuum of Metadata Quality: Defining, Expressing, Exploiting. Metadata in Practice.

Cepeda, O., Gallardo, I. & Rodríguez, J. (2017). Aspectos e indicadores para evaluar la calidad de los OA. Revista de Universidad y Sociedad del Conocimiento

Gordillo, M., Barra, A. & Quemada, V. (2018). Estimación de la Calidad de Objetos de Aprendizaje en Repositorios de Recursos Educativos Abiertos Basada en las

Interacciones de los Estudiantes. Educación XXI 21(1), 285–302

Hughes, B. (2004). Metadata Quality Evaluation: Experience from the Open Language Archives Community. Digital Libraries: International Collaboration and Cross-Fertilization.

Massa, S., De Giusti, A. & Pesado, P. (2012a). Objetos de Aprendizaje: Metodología de Desarrollo y Evaluación de la Calidad. 315. La Plata, Buenos Aires, Argentina: Universidad Nacional de la Plata: Facultad de Informática.

Massa, S. M. (2012b) Objetos de Aprendizaje: Metodología de Desarrollo y Evaluación de la Calidad. Universidad Nacional de La Plata.

Morales, E., Gil, A. & García, F. (2007) Arquitectura para la Recuperación de Objetos de Aprendizaje de calidad en Repositorios Distribuidos. SCHA: Sistemas Hipermedia Colaborativos y Adaptativos. II Congreso Español de Informática CEDI.

Salazar, N. & Durán, E. (2014). Objeto de Aprendizaje para la enseñanza de la Simulación. En Anales IX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2014). Disponible en <http://sedici.unlp.edu.ar/handle/10915/38644>.

Salazar, N. & Durán, E. (2016). Objetos de aprendizaje personalizados para entornos ubicuos. Una revisión de antecedentes. En actas de XI Jornadas de Ciencia y Tecnologías de facultades de Ingeniería del NOA. UNJU.

Salazar, N. & Durán, E. (2017). Diseño de un objeto de aprendizaje para simulación basado en pedagogías emergentes, XII Congreso Latinoamericano de Tecnologías del Aprendizaje (LACLO), pp. 1-4, doi: 10.1109/LACLO.2017.8120943.

Salazar, N. & Durán, E. (2019). Evaluación de la calidad de los objetos de aprendizaje: un estudio de caso, XIV Congreso Latinoamericano de Tecnologías del Aprendizaje (LACLO), pp. 251-258, doi: 10.1109/LACLO49268.2019.00050.

Ochoa, X. (2008) Learnometrics: Metrics for Learning Objects Katholieke Universiteit Leuven.

Osgood Ch.E. Suci G.J. & Tannenbaum P. (1957) The Measurement of meaning. Universidad de Illinois.

Pinto, M., Gómez, C., Fernández, A. & Vinciane, A. (2015). Evaluareed: desarrollo de una herramienta para la evaluación de la calidad de los recursos educativos electrónicos. Investigación Bibliotecológica 31(72), 227–248. México

Tabares, M., Duque, M. & Ovalle, C. (2017). Modelo por capas para evaluación de la calidad de Objetos de Aprendizaje en repositorios. Revista Electrónica de Investigación Educativa 19(3), 33–48

Triana, M., Ceballos, J. & Villa, J. (2016) Una dimensión didáctica y conceptual de un instrumento para la Valoración de Objetos Virtuales de Aprendizaje. El caso de las fracciones Entramado 12(2), 166–186. Universidad Libre Cali, Colombia.

Vidal, C., Segura, A. & Prieto, M. (2008) Calidad en Objetos de Aprendizaje. V Simposio Pluridisciplinar sobre Diseño y Evaluación de Contenidos Educativos Reutilizables SPEDECE 08. Salamanca, España

Wong, KY, Boey, KL, Lim-Teo, SK et al. (2012) La preparación de profesores de matemáticas de primaria en Singapur: programas y resultados del estudio TEDS-M. Educación matemática ZDM 44, 293–306. Disponible en: <https://doi.org/10.1007/s11858-011-0370-1>

Marco para desarrollar nuevas unidades de proyectos e incorporar docentes y alumnos en actividades de investigación

Bianchini Germán¹, Caymes-Scutari Paola^{1,2}, Ontiveros Patricia³, Poch Miguel³, Rotella Carina³, Tagarelli Sandra⁴, Alario Rocío¹, Brancolini Alessandro¹, Cáceres Felipe¹, Estrella Micaela¹, Isgró Valentino¹, Ledesma Emiliano¹, Masuet Juan Pablo¹, Mazurán Clara¹, Tintenfich Dylan¹

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido,
Departamento de Ingeniería en Sistemas de Información. Facultad Regional
Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Laboratorio de Gobierno Electrónico, Departamento de Ingeniería en Sistemas de
Información. Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza

⁴Laboratorio de Analítica de Datos, Departamento de Ingeniería en Sistemas de
Información. Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza

gbianchini@frm.utn.edu.ar, pcaymesscutari@frm.utn.edu.ar, pontiveros@frm.utn.edu.ar,
miguelpoch@yahoo.com.ar, carinarotella@gmail.com, stagarelli@gmail.com,
rocioalario1607@gmail.com, alessandro.brancolini@alumnos.frm.utn.edu.ar,
felipe.caceres@alumnos.frm.utn.edu.ar, micaelaestrellamoraes@gmail.com,
valentinoisgro@gmail.com, emilianoledesma250@gmail.com,
juan.masuet@alumnos.frm.utn.edu.ar, maria.mazuran@alumnos.frm.utn.edu.ar,
tintenfichdylan@gmail.com

RESUMEN

El presente proyecto, que es la continuación de una primera experiencia realizada desde el 2020 al 2022, propone la formación de investigadores como objetivo primordial y a la vez define un marco globalizador que permita la coexistencia de diversas unidades de proyectos que de otra manera no tendrían la posibilidad de llegar a la instancia de conformar un proyecto independiente. Dicho Marco constituirá un espacio de promoción de ideas e iniciativas para la iniciación en la investigación, tanto para estudiantes de grado como para docentes. Ofrece la posibilidad de realizar prácticas

de investigación específicas como un primer peldaño que permita acceder a distintos elementos (categorización, recursos económicos, acompañamiento de los directores del proyecto y miembros con más experiencia, posibilidad de obtención de becas y, fundamentalmente, un contexto que guíe y enmarque la experiencia para favorecer su concreción). Es importante remarcar que normalmente no todo docente (menos aún los estudiantes) logra involucrarse y participar formalmente dentro de un proyecto de investigación dado que los requisitos para la conformación de grupos, la solicitud de financiamiento, y/o su incorporación a actividades de

investigación suelen convertirse en un obstáculo. Por ende, permitir la incorporación de los mismos, guiar su trayecto formativo y posibilitar los medios económicos mínimos para desarrollar y dar a conocer sus propuestas, es el objetivo final del presente proyecto.

Palabras clave: Investigación, Alumnos, Docentes, Inclusión.

CONTEXTO

El presente proyecto, que se encuentra en proceso de evaluación, se desarrollará en el LICPaD (Laboratorio de Investigación en Cómputo Paralelo /Distribuido), contando con la participación tanto de docentes del propio laboratorio, como de docentes de los grupos ADA-Lab (Laboratorio de Analítica de Datos) y GE-Lab (Laboratorio de Gobierno Electrónico), todos grupos pertenecientes al Departamento de Ingeniería en Sistemas de Información de la UTN-FRM. Mediante este trabajo se busca conjugar la formación teórica y la práctica de investigadores iniciales para que éstos atraviesen todas las etapas necesarias que involucra el proceso científico en un ámbito de investigación. Para esto se planteó utilizar un trabajo de investigación específico como meta del proceso de aprendizaje, enfocado en las competencias específicas que deben adquirir los graduados de Ingeniería en Sistemas de Información, además de impulsar el pensamiento crítico de docentes y alumnos frente a nuevos retos. En el caso del LICPaD, los participantes de la formación son alumnos y becarios, mientras que en el caso de ADA-Lab y GE-Lab, los participantes en formación son los propios docentes.

1. INTRODUCCIÓN

La formación de investigadores –y la ciencia en general–, ha tomado especial importancia dado que se ha evidenciado que aporta significativamente al desarrollo de sociedades que procuran el bienestar a través del desarrollo científico, y que adicionalmente están preparadas para afrontar los desafíos de la naturaleza, la tecnología y el contexto mundial, procurando así el desarrollo social y cultural del país [1, 2]. En todo el mundo se vienen registrando grandes avances en la formación de investigadores; por ejemplo, en algunos de los países que forman parte de la OCDE (Organización para la Cooperación y el Desarrollo Económicos), tales como República Checa, Portugal, Austria, Dinamarca y Eslovenia, la formación de recursos humanos en investigación se ha incrementado exponencialmente [3].

En el continente americano, un país como los Estados Unidos de América ocupa el primer lugar en cuanto a número de investigadores científicos (según datos publicados por la UNESCO en 2015, en EEUU el 10% de cada 1000 habitantes de su población económicamente activa (PEA) se dedica a actividades relacionadas con la investigación), al que sigue Canadá con un 8% por cada 1000 de la PEA [4].

Al sur del continente la formación de investigadores se desarrolla principalmente en las universidades, fundamentalmente a través de programas de doctorado. Si bien Argentina es la nación con mayor número de investigadores, con un 2,4% de científicos por cada 1000 habitantes de su PEA [5] (por sobre Chile con un 2% y Brasil con un 1,3%) existe un rezago considerable en cuanto al número de investigadores –y a los recursos económicos que el estado dedica– y a su vez a la distribución a lo largo y ancho del país, y en consecuencia a la formación de profesionales en esta

área. Además, es importante remarcar que parte de las deficiencias en la formación de investigadores ha tenido origen en los niveles iniciales del sistema educativo, pues sólo hace muy poco tiempo, y en general como actividades optativas o complementarias, los planes nacionales de educación han comenzado a contemplar la formación para la investigación como un elemento importante en la educación de niños y jóvenes.

Debido a esto, diversas instituciones a lo largo y ancho del mundo han comenzado a desarrollar programas para el fomento de vocaciones científicas y de investigación a edades tempranas [6,7,8,9]. Por lo que es claro que el camino es implementar estrategias enfocadas en la formación de investigadores. Esta necesidad de caminar cada vez más en el sendero de la investigación también se aprecia en los procesos de acreditación de las diversas carreras universitarias, en donde un aspecto importante en la evaluación de las mismas son los proyectos de investigación y los aportes que el plantel docente –y en lo posible el alumnado– pueden realizar durante su vida académica.

Si bien en la UTN existen recursos para la formación de investigadores (un ejemplo son los cursos PROFORVIN [10], que anualmente se dictan con una serie de contenidos que brindan información a los participaciones acerca de cuestiones como las Buenas prácticas de la vinculación tecnológica, la Gestión de la Ciencia, la tecnología y la innovación o Inglés para la publicación de artículos científicos, por citar algunos ejemplos), hallamos que hay una brecha y un desconocimiento, proveniente de la desinformación y la falencia de instrumentos específicos que permitan profesionalizar y financiar la investigación, lo que mantiene alejados a

los potenciales investigadores de estos recursos y fuentes de información. Por ello, en el presente proyecto se han abierto las puertas a los interesados en llevar a cabo sus propias unidades de proyectos y abrirse camino en el ámbito de la ciencia, favoreciendo el desarrollo de las mismas, si bien a la escala que los recursos lo permitan.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Como se ha comentado, cada unidad de proyecto abordará diferentes temas en relación a la formación y conocimientos previos de los integrantes. Dentro del área del Cómputo Paralelo, uno de los equipos de trabajo, conformado por dos estudiantes avanzados, se enfocará en explorar cómo el cómputo paralelo puede mejorar el rendimiento de una determinada Metaheurística y su aplicación en Machine Learning. Otra unidad de proyecto, compuesta por estudiantes y becarios de segundo año, proyecta desarrollar un algoritmo en paralelo para la resolución del problema del viajante, tema en el cual se comenzó a trabajar en un proyecto previo.

Los docentes intervinientes orientarán sus temas al uso de tecnología Blockchain para emisión de Certificados de Extensión Universitaria, tema en el que han avanzado en la programación de diversos módulos, y también comenzarán una unidad de proyecto que pretende fomentar el desarrollo de proyectos extensionistas en todas las universidades del país, ligado a una convocatoria de la Secretaría de Políticas Universitarias (SPU).

Finalmente, en otra línea, se pretende abordar un estudio sobre la inteligencia artificial y su aplicación a los servicios de internación domiciliaria, para tratar de generar un enfoque integral entre el

ámbito del equipamiento y el de los servicios médico-asistenciales.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el proyecto antecedente se abordaron diversas líneas temáticas, que en algunos casos fueron modificándose a lo largo del tiempo, organizadas en subgrupos de trabajo. Dentro del área del Cómputo Paralelo, las líneas elegidas fueron: Paralelismo en algoritmos de aprendizaje para redes neuronales, Redes neuronales paralelas aplicadas a la visión computacional, Predicción del avance del dengue en zonas endémicas y Multiplicación de Matrices.

Los docentes intervinientes del subgrupo ADA-Lab orientaron sus temas de investigación hacia Indicadores de procesos y medición, y los docentes del subgrupo GE-Lab primero se centraron en la temática de aplicación de Blockchain Federal (BFA) sobre la gestión de documentos judiciales, y luego en la aplicación de la tecnología de Firma Digital Remota en actas de exámenes de la UTN-FRM, la cual finalmente quedó como línea abierta a ser desarrollada en un futuro. Además, en simultáneo se comenzó a esbozar un sistema para la autenticación facial de alumnos.

En lo concerniente a los resultados de las investigaciones, obtenidos de las tareas individuales o grupales, si bien la contribución principal que se propició fue la transferencia de formación por parte de docentes-investigadores hacia docentes y alumnos de grado y al medio educativo, también se impulsó que los alumnos aprendiesen y experimentasen el trabajo en colaboración y por metas, puesto que el trabajo de investigación requiere normalmente de espacios inter y transdisciplinarios.

En concreto, en el presente proyecto los objetivos generales de investigación son primariamente consolidar la actividad de investigación, formar recursos humanos y apoyar la incorporación de los mismos en el plantel docente de la carrera de Ingeniería en Sistemas de Información, lo que por añadidura permitirá apoyar el crecimiento de nuevas unidades de proyectos y líneas de investigación, promover la información y divulgación científico-tecnológica, como así también el seguimiento y evaluación de la producción y la promoción de la innovación tecnológica.

Teniendo esto en consideración, la propuesta constituye un proyecto marco, o meta-proyecto, que aborda dos niveles de proyectos en simultáneo: en una capa, están las unidades de proyecto, que se enfocan en problemas particulares concretos de cada grupo o subgrupo. En la otra capa, nos enfocamos en el problema de investigación como actividad *per se*, y por tanto en la formación y guía de los participantes, en busca de su desarrollo y del éxito de su unidad de proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente proyecto se enfoca precisamente en contribuir al crecimiento y desarrollo de investigadores nóveles y afianzar el desarrollo de docentes-investigadores que en una instancia previa realizaron los primeros pasos en investigación y actualmente abordan temáticas en las unidades de proyecto. Como se aprecia en el listado de participantes, además de los dos investigadores formados que conducen y coordinan el proyecto, participan cuatro docentes y nueve estudiantes, entre alumnos y becarios. En este último caso,

una parte cuenta con el beneficio de becas de investigación (ya sean BIS o I+D), que, aunque económicamente son de montos muy bajos, les brindan una formalidad y una mínima ayuda que les permite dedicar cierta cantidad de horas semanales a la tarea de investigar para así cumplir los objetivos formales establecidos. Además, dicha actividad los habilita a la realización de unidades de proyectos, en algunos casos incluso propuestos por ellos mismos, como así también contar con mínimos fondos para lograr su concreción, lo que es extremadamente valioso para aquellos que quieran iniciarse en el camino de la investigación, estén en la etapa de desarrollo de su proyecto de fin de carrera o manifiesten una vocación científica. En el caso de los docentes en formación, algunos cuentan con una dedicación simple para tareas de investigación, pero requieren una guía para poder abordar sus proyectos y a más largo plazo lograr independencia para obtener sus propias líneas de financiamiento.

5. BIBLIOGRAFÍA

- [1] Cazaux, Diana. (2008). La comunicación pública de la ciencia y la tecnología en la “sociedad del conocimiento”. Las universidades argentinas y la divulgación de su producción científica a través de internet. Bitácora-e: revista electrónica latinoamericana de estudios sociales, históricos y culturales de la ciencia y la tecnología, 1, pp. 66-87
- [2] Aldana, Maximino. (2012). ¿Qué le falta a la ciencia en México? Revista Temas, 69, pp. 26-30.
- [3] Sanz Menéndez, Luis & Cruz Castro, Laura. (2010). Análisis sobre ciencia e innovación en España. España: Instituto de Políticas y Bienes Públicos del CSIC.
- [4] CONACYT. (2014). Programa Especial de Ciencia, Tecnología e Innovación 2014-2018. México: Gobierno de la República.
- [5] Sebastián, Jesús. (2003). Estrategias de cooperación universitaria para la formación de investigadores en Iberoamérica. Madrid, España: Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura.
- [6] Centro Universitario de Los Lagos. Universidad de Guadalajara. <http://www.lagos.udg.mx/> (fecha de consulta: abril de 2019)
- [7] Facultad de Medicina, Universidad de Navarra. <https://www.unav.edu/web/facultad-de-medicina/la-investigacion-en-la-facultad> (fecha de consulta: abril de 2019)
- [8] Universidad Maimónides. <https://www.maimonides.edu/investigacion/> (fecha de consulta: abril de 2019)
- [9] Escalante Héctor, Corrales Rafael, Barahona Miguel (2011). La investigación científica en los estudiantes de grado y posgrado. Revista Portal de la Ciencia, UNAH. Nro.11. pp.19-25.
- [10] Cursos PROFORVIN. <https://www.utn.edu.ar/es/secretaria-academica/sacad-posgrado/proforvin/sacad-cursos-proforvin>

HERRAMIENTAS COLABORATIVAS MULTIPLATAFORMA PARA EL APRENDIZAJE BASADO EN COMPETENCIAS EN LA INGENIERÍA DE SOFTWARE

Nicolás Battaglia, Carlos Neil, Marcelo De Vincenzi,
Fabian Maffei, Nelson Garrido, Facundo Oliva, Jorge Lomoro

Universidad Abierta Interamericana. Facultad de Tecnología Informática.
Centro de Altos Estudios en Tecnología Informática. Buenos Aires. Argentina
{Nicolás.Battaglia, Carlos.Neil, MeDevincenzi}@uai.edu.ar,
{Fabian.Maffei, Nelson.Garrido, JuanFacundo.Oliva, Jorge.Lomoro}@alumnos.uai.edu.ar

RESUMEN

Las instituciones educativas han adoptado la educación híbrida y la educación a distancia, por este motivo, han tenido que enfrentar diversos desafíos tales como reinventar sus modelos pedagógicos para que el proceso de enseñanza aprendizaje continúe de manera ininterrumpida y capacitar su cuerpo docente. Todo esto no hubiese sido posible sin los cambios constantes de las TIC que nos hicieron reflexionar sobre la efectividad del método tradicional de enseñanza: los graduados deben ser competentes, deben conocer los saberes de un área de conocimiento y aplicarlos en futuros contextos profesionales. Surge entonces un doble desafío: diseñar una infraestructura para poder brindar servicios de educación mediada por la tecnología y que los estudiantes puedan aprender en forma colaborativa para acreditar sus competencias. En esta línea, el aprendizaje basado en competencias debe verse reflejado en los entornos colaborativos de aprendizaje en todas sus actividades, en el marco del alineamiento constructivo. El proyecto propone la integración de plataformas tecnológicas para el aprendizaje colaborativo y evaluación basada en competencias, en cursos vinculados con la ingeniería de Software.

Palabras clave: aprendizaje colaborativo, aprendizaje basado en competencias, ingeniería de software, trabajo colaborativo, UML.

CONTEXTO

El proyecto de investigación se desarrolla en la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI) con estudiantes de 2° a 5° año de la carrera de ingeniería en sistemas informáticos que trabajan de manera articulada en las siguientes asignaturas, Análisis y diseño de sistemas I, II (2° año); Análisis y diseño de sistemas III, Bases

de datos, Trabajo de diploma e Ingeniería de software I (3° año) y Seminario de trabajo final de ingeniería y Trabajo final de ingeniería (5° año).

1. INTRODUCCIÓN

La constante innovación tecnológica en TIC genera profundos cambios en la sociedad, en su forma de vivir y pensar. La presencia generalizada de computadoras y dispositivos que permiten tener información constante, en cualquier momento y en cualquier lugar demuestra de forma sustentable que la informática se ha vuelto ubicua [1]. Estos avances en la tecnología crean nuevos contextos de comunicación e interacción entre usuarios, generando nuevos conjuntos de herramientas, soportes y canales que brinden un mejor tratamiento de la información, mejorando las condiciones técnicas y el tiempo de respuesta, dejando disponibles nuevos dispositivos que reemplazan a otros más antiguos y aburridos [2], la capacidad de ubicuidad que genera este avance tecnológico, permite disponer de contenidos a lo largo del día, en cualquier momento y en cualquier lugar; obligando a disponer de nuevas infraestructuras tecnológicas [3].

Desde el comienzo del proyecto de investigación [4]–[8] afirmamos que el cambio y la evolución generó la necesidad de repensar el sistema educativo. Uno de los cambios más relevantes de los últimos tiempos es la capacidad actual de poder acceder a información de manera ubicua. Esto está vinculado, sin dudas, con la evolución en las TIC [9].

De esta manera, gracias a Internet y a los dispositivos móviles, el estudiante puede aprender de manera virtual, utilizar herramientas sincrónicas (videoconferencias, chat, etc.) y asincrónicas (foros y redes que promueven la interacción con otras personas), facilitándoles el aprendizaje colaborativo. Por este motivo, el trabajo colaborativo, las TIC y la enseñanza se pueden integrar para crear entornos de

aprendizaje colaborativo asistido por computadora. Esta integración se basa en el impacto social del trabajo colaborativo y las TIC en la educación que transforma la manera tradicional de enseñanza/aprendizaje. Esto conlleva varios cambios, entre ellos el Aprendizaje Basado en Competencias (ABC) y centrada en el estudiante y el impacto de las TIC en los procesos de enseñanza y aprendizaje, para lo cual es necesario también modificar y adaptar los criterios y estrategias de evaluación motivados por esta disrupción [10].

En resumen, el proyecto de investigación se fundamenta en los siguientes pilares: el aprendizaje colaborativo basado en competencias, la matriz de competencias, la evaluación mediante rúbricas analíticas y el alineamiento constructivo.

2. APRENDIZAJE COLABORATIVO

El trabajo colaborativo apoyado en el concepto de Zona de Desarrollo Próximo (ZDP) propuesto por Vygotski a principios del siglo XX [11] y, por el otro, las TIC en el marco de las propuestas de *e-learning* o *blended learning*. Esto sumado al desarrollo creciente de las TIC junto con el concepto de trabajo colaborativo, conforman los entornos de Trabajo Colaborativo Asistido por Computadora (CSCW). Este concepto, integrado en entornos de enseñanza y aprendizaje colaborativo, dio origen a los entornos de Aprendizaje Colaborativos Asistidos por Computadora (CSCL) [12], [13]. En el ámbito académico, los nuevos estándares para la acreditación de carreras de ingeniería en la República Argentina [14], [15] que proponen el ABC y están centrado en el estudiante, dan cuenta de la necesidad de la integración de ambos modelos, además de considerar todas las herramientas de seguimiento y evaluación que sean requeridas por este paradigma. Por otro lado, el estudio teórico de la Ingeniería de Software (IS) no es suficiente para comprender y resolver los problemas de cooperación y colaboración que surgen durante el desarrollo de un proyecto informático [16]. Los estudiantes suelen centrar su esfuerzo en aspectos técnicos y asumen que los inconvenientes vinculados al trabajo en equipo no impactarán en el proyecto. En virtud de resolver el problema planteado previamente, Daniele et al. [17] proponen, para mejorar los aspectos comunicacionales, la integración de las plataformas CSCL con entornos especializados en la resolución de problemas prácticos y técnicos en un proyecto de desarrollo de software; además, esta propuesta potencia el aprendizaje de la IS y permite el

desarrollo de competencias profesionales y sistemáticas en el desarrollo de software [18]. De la misma manera, tal como sucede en los procesos tradicionales de enseñanza, en los entornos CSCL, la evaluación cumple un rol muy importante. Según el estudio realizado en [19], en el modelo de enseñanza y aprendizaje de la IS mediado por tecnología, los estudiantes cumplen un papel de mayor protagonismo y responsabilidad y, en consecuencia, los docentes asumen un nuevo rol cuando es el estudiante quien adquiere mayor autonomía en su proceso de aprendizaje a través de entornos virtuales. Esta autonomía es una característica distintiva de la enseñanza centrada en el estudiante. En este contexto, la evaluación es entendida como un proceso que promueve el aprendizaje con una finalidad formativa, más que como un proceso de control de resultados. Por consiguiente, un proceso específico de enseñanza y aprendizaje en un entorno virtual colaborativo requiere también de un proceso de evaluación, autoevaluación y coevaluación (evaluación 360°) y seguimiento acorde al modelo de aprendizaje colaborativo planteado por el concepto CSCL [10]. El énfasis en la educación centrada en el estudiante permite retomar el modelo de ABC, que también promueve reformular y adaptar los criterios y estrategias de evaluación motivados por la disrupción tecnológica que da origen a los entornos CSCL.

3. MATRIZ DE COMPETENCIAS

La universidad debe preparar al estudiante en un entorno de competencias genéricas y específicas que les permitan afrontar los problemas que plantea esta sociedad cambiante [20]. En la búsqueda de una alternativa a las pruebas de aptitud e inteligencia tradicionales, desarrolló el concepto de “competencia” definido como una característica subyacente de una persona que le permite demostrar un desempeño superior en un determinado rol o situación, haciendo la diferencia entre personas con desempeño excelente versus personas con desempeño promedio. Por otro lado, el concepto de competencia engloba todo un conjunto de conocimientos, procedimientos y actitudes que se combinan, coordinan e integran, para que el individuo logre «saber conocer», «saber hacer», «saber ser» [21]. El dominio de estos saberes le permite ser «capaz de» actuar con eficacia y eficiencia en situaciones que se le presenten [22]. En este sentido, la competencia no reside en los recursos (capacidades) sino en cómo éstos se movilizan para lograr un objetivo. Además, el aprendizaje por competencias ha promovido toda

una renovación de las teorías psicopedagógicas de los aprendizajes, dando lugar a la evolución de los esquemas de referencia de la formación de profesionales [22].

Un instrumento que permite obtener una visión general de cómo los diferentes espacios curriculares aportan al perfil de egreso es la matriz de competencias que representa, para cada una de ellas, los niveles de dominio de los distintos espacios curriculares del plan de estudios. Con el fin de evaluar el proceso de desarrollo de las competencias, se han establecido niveles de dominio que permiten establecer los logros en el aprendizaje durante un módulo determinado, en la parte intermedia y al final de un ciclo [23]. En tal sentido, la matriz de competencia representa la forma en que cada asignatura tributa al logro de una competencia considerando los distintos niveles de complejidad, integración y autonomía que alcanza el estudiante. De este modo, cada asignatura de un plan de estudios se orienta a promover determinados resultados de aprendizaje conforme a estas especificaciones. Esto implica que las competencias se deben desarrollar gradualmente y en forma progresiva y planificada a lo largo de la carrera. Esta matriz se diferencia de la de tributación [24] donde el foco está centrado en los espacios curriculares, en lugar de los estudiantes como lo hace la matriz de competencias, a partir de concepto de nivel de dominio.

En particular, en la Universidad Abierta Interamericana (uai.edu.ar) se definieron, para todas las carreras, matrices de competencias con tres niveles de dominio que el alumno debe transitar es su recorrido académico de modo tal que la concatenación de los resultados de aprendizaje de las asignaturas que aportan a una competencia se desarrolle en su máximo nivel (de 1 a 3) al finalizar el plan de estudios. En la matriz de competencia queda explícito cómo cada asignatura de un plan de estudios se orienta a promover determinados resultados de aprendizaje.

4. EVALUACIÓN DE COMPETENCIAS

Tal como planteó Perrenoud [25], las competencias son invisibles y sólo son abordables a través de los desempeños observables, por lo tanto, en línea con Zapata [26] éstas deben ser observables y evaluables y esto sucede, básicamente, por medio de las conductas que los sujetos generan ante diversos problemas en contextos cambiantes. Esto plantea la necesidad de la sustitución o coexistencia de

procedimientos, herramientas y formas de evaluar los saberes por medio de procesos descriptivos o conceptuales (exámenes, pruebas de destreza, problemas teóricos, etc.) por otros métodos donde se tenga en cuenta la simulación real de casos prácticos, con metodologías de indagación, de investigación formativa, con elaboración de proyectos, etc.

Como consecuencia, el ABC requiere un sistema de evaluación variado, debido a que cada competencia tiene componentes muy distintos que necesitan procedimientos diversos para ser evaluados correctamente [27]. Si bien existen diversas metodologías de evaluación de competencias [28], en la actualidad uno de los mejores instrumentos, son las rúbricas, debido a su facilidad de uso y a la oportunidad de obtener aspectos complejos, imprecisos y subjetivos [19].

Según Zapata [29], cuando se evalúa en un entorno virtual de aprendizaje, es necesario también la comunicación de los resultados que ayudan al estudiante a comprender el nivel de adquisición de una competencia determinada. En estos entornos, es fundamental el concepto tradicional de retroalimentación o *feedback* propuesto por los entornos virtuales. Sin embargo, en los entornos virtuales colaborativos, es necesario medir también el nivel de participación de un estudiante dentro de un equipo de trabajo, por ejemplo, por medio de la cantidad de mensajes enviados, tomando en cuenta aspectos semánticos y de relevancia en cuanto a cómo son estos mensajes, su influencia en otros estudiantes, si son respondidos, etc.

i. Rúbricas analíticas

La necesidad de instrumentar la evaluación para que facilite la mejora del aprendizaje implica, primero, plantearla como una tarea de aprendizaje; segundo, involucrar a los estudiantes y, finalmente, ofrecer sus resultados como retroalimentación [30]. Estos tres objetivos pueden abordarse mediante el uso de rúbricas que, como instrumento de evaluación, permiten valorar los aspectos complejos, imprecisos y subjetivos de un tema/problema, aportando una evaluación fácilmente interpretable, justa y transparente para profesores y estudiantes [31].

Las rúbricas no sólo evalúan, también enseñan; la autoevaluación basada en su uso contiene los criterios de evaluación que el estudiante ha de utilizar para valorar su trabajo [32]. Como instrumento de evaluación, brinda un conjunto de criterios de calidad relacionados con las competencias a evaluar y, mediante descriptores y diferentes niveles de logro, ponen de manifiesto no sólo el incremento cuantitativo de los estudiantes, sino también el salto

cualitativo; demuestran cuánto han aprendido y lo bien que lo han hecho [33]. La rúbrica no es solamente una herramienta de evaluación sino, también, una estrategia de aprendizaje; en el proceso de autoevaluación, el estudiante no resuelve un problema y luego se autoevalúa, lo hace con las consideraciones establecidas en los criterios de evaluación que le permite reflexionar sobre el mismo y actuar en consecuencia. Por otro lado, su utilización se potencia en los contextos virtuales de enseñanza y aprendizaje, ya que su diseño y uso empoderan al docente, al estudiante y a la institución promoviendo la cultura de la evaluación auténtica en la educación [34].

Por otro lado, la evaluación no debe ser un proceso distinto del aprendizaje, es una oportunidad para promoverlo [35]. La evaluación 360° tiene como objetivo desarrollar un modelo que cubra todos los aspectos relevantes de la evaluación de un proceso a partir de la retroalimentación, permitiendo elevar la calidad y la mejora continua tanto del desempeño docente como del estudiante, es decir, de la enseñanza y del aprendizaje [36]. En este sentido, las rúbricas favorecen la autoevaluación debido a que poseen dos características fundamentales: los criterios de evaluación y los estándares de calidad [32].

5. ALIENAMIENTO CONTRUCTIVO

Los resultados de aprendizaje, las actividades de enseñanza y aprendizaje y los medios de evaluación se deben definir de manera coherente (se "alinean") para mejorar el aprendizaje del estudiante [37]. El alineamiento constructivo es un modelo conceptual que persigue unos fines pedagógicos muy claros: mejorar la manera de enseñar de modo que todos los miembros de la clase tengan la oportunidad de aprender más profundamente. El alineamiento constructivo requiere que los profesores conozcan, con claridad y precisión, lo que se consideran los tres elementos centrales del planeamiento educacional. Los resultados de aprendizaje: ¿Qué esperamos que nuestros estudiantes logren en nuestras carreras, cursos o clases? Las actividades de enseñanza y aprendizaje: ¿Qué van a hacer nuestros estudiantes para alcanzar los resultados esperados y qué vamos a hacer nosotros para apoyarlos? y los medios de evaluación: ¿Cómo vamos a evaluar si nuestros estudiantes alcanzaron los resultados esperados?

Se espera también que los docentes coordinen estos tres pilares, es decir, que garanticen su cohesión y consistencia internas y que tengan plena conciencia de que, si cambian uno de ellos,

los otros dos elementos del plan deben ajustarse. Esto significa que los métodos de enseñanza y los sistemas de evaluación se definen paralela e integradamente en relación con las competencias a alcanzar

6. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente trabajo está identificado como proyecto de investigación y desarrollo en el Centro de Altos Estudios en Tecnología Informática (CAETI), dentro de la línea de investigación Sociedad del Conocimiento y Tecnologías aplicadas a la educación. Además, el proyecto se ramifica en diferentes áreas temáticas específicas, a saber: 1) Analítica de aprendizaje: entendemos este concepto como "la medición, recopilación, análisis e informe de datos sobre los estudiantes y sus contextos, con el fin de comprender y optimizar el aprendizaje y los entornos en los que se produce". 2) Estilos de aprendizaje y personalización de la enseñanza: el aprendizaje personalizado es un enfoque educativo cuya finalidad es que éste el aprendizaje se ajuste a las fortalezas, necesidades, habilidades e intereses de cada estudiante. La identificación de estilos de aprendizaje permite adaptar las estrategias de enseñanza a las características del estudiante. El objetivo es identificar, para cada estudiante, su estilo de aprendizaje para personalizar las estrategias de enseñanza. 3) Estudios empíricos sobre modelos de enseñanza aprendizaje: la herramienta UAI Case [7] se enfoca en la enseñanza y el aprendizaje del modelado de software utilizando el estándar UML. Existen muchos enfoques de enseñanza aprendizaje, desde los tradicionales hasta los más innovadores. El objetivo es realizar estudios empíricos que permitan establecer la ganancia de aprendizaje cuando se utiliza UAI Case respecto de otras estrategias de enseñanza. 4) Inteligencia artificial aplicada a la educación: el aprendizaje en entornos virtuales permite al estudiante organizar sus tiempos y no depender de un encuentro sincrónico para cumplir los objetivos, sino que puede hacerlo de manera desconectada, en cualquier momento y lugar. Sin embargo, el estudiante necesita obtener *feedback* del profesor y para eso depende de sus tiempos y espacios para poder responder consultas y corregir actividades. En base a esto, consideramos que la incorporación de herramientas cognitivas de educación, permitirán asistir a los estudiantes en tiempo real durante el proceso de aprendizaje y de esta manera permitir una mayor independencia del estudiante para poder realizar

sus actividades. 5) Gamificación: la ingeniería de software es una actividad centrada en las personas [38], que no solo incluye aspectos de tipo técnico o tecnológico, sino que además está fuertemente influenciado por las competencias blandas que poseen los miembros del equipo. Bajo esta perspectiva, el proceso de desarrollo de software enfrenta desafíos relacionados con el tratamiento de factores sociales y humanos que influyen en el desempeño de los equipos de trabajo y su repercusión en el éxito o no de los proyectos de desarrollo de software [39]. La investigación se centrará en la adecuación de estrategias didácticas basadas en la gamificación [40], que se puedan diseñar y construir en un entorno virtual colaborativo, para la formación en ingeniería de software [41]. Utilizando las estrategias didácticas mediadas por recursos TIC, que apunten a mejorar la enseñanza, aprendizaje y evaluación de competencias específicas de la ingeniería de software. 6) Mecanismo estandarizado de propuesta de mejoras en las herramientas de enseñanza virtual en Plataformas educativas: dado que las plataformas de educación virtual son un recurso utilizado en la enseñanza superior y universitaria, es menester tener un estándar de calidad en su uso e implementación [42]. Existen plataformas comerciales que ya vienen preparadas para ser multifuncionales [43] pero estas deberían ser útiles para el desarrollo de las actividades de enseñanza y aprendizaje [44]. El objetivo es crear un mecanismo estandarizado de propuesta de mejoras en las herramientas de enseñanza virtual ubicua (específicamente, en plataformas educativas) para obtener una perspectiva general de los aspectos básicos que debe tener una plataforma para satisfacer los requisitos de la educación virtual de calidad [45].

7. RESULTADOS OBTENIDOS/ESPERADOS

i. Resultados obtenidos:

En el marco del proyecto de investigación se han presentado, el año 2022, los siguientes trabajos de investigación: “Marco metodológico para el diseño de rúbricas analíticas” [46]. “Tecnología asistiva para personas con discapacidad en miembros superiores: un mapeo sistemático de la literatura” [47]. “Herramientas para determinar estilos de aprendizaje basadas en Inteligencia artificial” [48]. Además, se desarrolló una herramienta para diseñar rúbricas analíticas y un repositorio para que sean compartidas por la comunidad educativa ([UAIRubric](#))

ii. Resultados Esperados:

Se esperan los siguientes resultados: 1) Diseñar herramientas que permitan integrar el concepto de ABC junto con herramientas que faciliten evaluar y realizar seguimientos en entornos colaborativos (UAI Case). 2) Generar indicadores en tiempo real que permitan medir la evolución de los saberes y competencias en el ABC. 3) Obtener un conjunto de indicadores que permitan medir el rendimiento de un equipo virtual de trabajo durante el proceso de enseñanza y aprendizaje de la IS. 4) Realizar la evaluación empírica del prototipo de la herramienta UAI Case por medio de la definición de un método de evaluación basado en métricas. 5) Desarrollar sistemas de tutorías inteligentes mediante la implementación de tutores cognitivos que imiten el rol del profesor, guíen el desarrollo del aprendizaje y ofrezcan pistas a los estudiantes cuando están atascados en un problema. 6) Utilizar analíticas del aprendizaje que permitan, mediante la medición, recopilación y análisis de datos de los estudiantes durante el proceso de enseñanza aprendizaje, detectar los errores más comunes y proporcionen una respuesta en tiempo real al estudiante. 7) Personalizar las estrategias de enseñanza a partir de los estilos de aprendizaje de los estudiantes utilizando inteligencia artificial. 8) Proponer mejoras en las herramientas de enseñanza virtual en Plataformas educativas mediante un mecanismo estandarizado. 9) Desarrollar una herramienta que facilite detectar, en un ambiente de trabajo con varias pantallas, hacia cual el operador está dirigiendo su mirada. 10) plantear lineamientos para la escritura de resultados de aprendizaje a partir de la matriz de competencias.

8. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado, por 13 estudiantes de posgrado (doctorado en ciencias informáticas, maestría en tecnología informática y especialización en ingeniería de software) quienes están desarrollando sus respectivas tesis y trabajos finales en temas vinculados al proyecto: Fernando Parra, Jorge Lomoro, Silvia Poncio Gabriela Iannantuoni, Jorge Zarate, Marcelo Monferrato, Juan Facundo Oliva, Charles Maldonado, Nelson Garrido, Fabian Maffei y Darío Kiryczun y los alumnos de grado Facundo Romeu y Juan Manuel Stecklain.

9. BIBLIOGRAFÍA

- [1] S. G. Figueroa and R. D. Cordero, "Informática ubicua: su aplicación en el contexto universitario," in *XIV Workshop de Investigadores en Ciencias de la Computación*, 2012.
- [2] A. Bartolomé, J. Cabero, M. Cebrian, L. Serna, and F. Martínez, "Nuevas tecnologías, comunicación y educación," *edutec.es*, 1996.
- [3] J. L. Filippi, G. J. Lafuente, and R. A. Bertone, "Diseño de un ambiente de aprendizaje colaborativo," *V Congreso de Tecnología en Educación y Educación en Tecnología*. 2010.
- [4] C. Neil, M. De Vincenzi, N. Battaglia, and R. Martínez, "Herramientas Colaborativas Multiplataforma en la Enseñanza de la Ingeniería de Software," In *XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina)*., 2016.
- [5] M. de Vincenzi, C. Neil, N. Battaglia, and R. Martínez, "Uaicase: Enseñanza De Ingeniería De Software En Entornos Virtuales Colaborativos," *XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste)*., pp. 1124–1129, 2018.
- [6] N. Battaglia *et al.*, "Desarrollo y evaluación de competencias en la ingeniería de software en un entorno virtual de aprendizaje colaborativo," in *XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza)*, 2022.
- [7] N. Battaglia, C. Neil, M. de Vincenzi, R. Martínez, and J. P. Beltramino, "UAI case: desarrollo y evaluación de competencias en la ingeniería de software en un entorno virtual de aprendizaje colaborativo," *XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)*., 2019.
- [8] Nicolás Battaglia, Carlos Neil, and Marcelo De Vincenzi, "Desarrollo y evaluación de competencias en la ingeniería de software en un entorno virtual de aprendizaje colaborativo," in *XXIII Workshop de Investigadores en Ciencias de la Computación*, 2021, pp. 875–880.
- [9] G. Lavigne, M. Ovando, J. Sandoval, and L. M. Salas, "Exploración preliminar del aprendizaje colaborativo dentro de un entorno virtual," *Actualidades Investigativas en Educación*, 2012.
- [10] N. Battaglia, "Integración de una Herramienta CASE en un Entorno Académico Colaborativo para la Enseñanza de Ingeniería de Software," Universidad Abierta Interamericana, 2017.
- [11] L. Vygotsky, "The collected works of LS Vygotsky: Problems of the theory and history of psychology," 1997.
- [12] G. Stahl, T. Koschmann, and D. Suthers, "Computer-supported collaborative learning: An historical perspective," 2006.
- [13] N. Battaglia, C. Neil, R. Martínez, G. Dana, and M. de Vincenzi, "Learning of Software Engineering on Collaborative Virtual Enviroments," in *Proceedings - 2017 7th World Engineering Education Forum, WEEF 2017- In Conjunction with: 7th Regional Conference on Engineering Education and Research in Higher Education 2017, RCEE and RHED 2017, 1st International STEAM Education Conference, STEAMEC 201*, 2018, pp. 548–552. doi: 10.1109/WEEF.2017.8467072.
- [14] CONFEDI, *Propuesta de estándares de segunda generación para la acreditación de carreras de ingeniería en la República Argentina "Libro Rojo de CONFEDI"*. 2018.
- [15] C. Neil, N. Sotomayor, R. Muñoz, P. Cristaldo, B. Parra de Gallo, and J. C. Calloni, *Aportes para la implementación de un modelo de formación orientado a competencias*, 1ra. Ed. Ciudad Autónoma de Buenos Aires: Universidad Abierta Interamericana, 2021.
- [16] P. Bouillon, J. Krinke, and S. Lukosch, "Software engineering projects in distant teaching," in *Proceedings - 18th Conference on Software Engineering Education and Training, CSEE and T 2005*, 2005. doi: 10.1109/CSEET.2005.31.
- [17] M. Daniele, M. Uva, P. Martelloto, and G. Picco, "Aplicación de herramientas CASE a la enseñanza de Ingeniería de Software: Gestión de la Configuración de Software y Testing Funcional," 2010.
- [18] N. Battaglia, C. Neil, M. De Vincenzi, and J. P. Beltramino, "Competency-based learning in collaborative virtual platforms," *International Institute of Innovation and Technology. 10th IIITEC International Symposium on Innovation*

- and Technology ISIT2019. Cuzco, Perú., 2019.
- [19] N. Battaglia, R. Martínez, M. Otero, C. Neil, and M. De Vincenzi, "Autoevaluación Colaborativa por medio de Rubricas en Entornos Virtuales de Enseñanza y Aprendizaje," 2016.
 - [20] D. C. McClelland, "Testing for competence rather than for 'intelligence'.,," *Am Psychol*, 1973, doi: 10.1037/h0034092.
 - [21] S. Tobón, "La Formación Basada en Competencias en la Educación Superior: El enfoque complejo.,," 2009.
 - [22] J. T. Fernández and C. R. Bueno, "Evaluación de competencias profesionales en educación superior: Retos e implicaciones," *Educacion XXI*, 2016, doi: 10.5944/educXXI.12175.
 - [23] S. T. Tobón, J. H. P. Prieto, J. Antonio, G. Fraile, and P. Hall, "Secuencias didácticas: aprendizaje y evaluación de competencias," 2010.
 - [24] A. González *et al.*, "Propuesta para evaluación del logro de perfiles de egreso," *Evaluación del logro de perfiles de egreso: Experiencias universitarias*, pp. 121–162, 2017.
 - [25] P. Perrenaud, *Construir competencias desde la escuela*. JC Sáez, 2008.
 - [26] A. Fernández March, "La evaluación orientada al aprendizaje en un modelo de formación por competencias en la educación universitaria," *REDU. Revista de Docencia Universitaria*, 2011, doi: 10.4995/redu.2010.6216.
 - [27] A. Villa Sánchez and M. Poblete Ruiz, *Aprendizaje Basado En Competencias Una Propuesta Para La Evaluación*. 2007.
 - [28] J. Cubero-Ibáñez, M. S. Ibarra-Sáiz, and G. Rodríguez-Gómez, "Propuesta metodológica de evaluación para evaluar competencias a través de tareas complejas en entornos virtuales de aprendizaje," *Revista de Investigacion Educativa*, vol. 36, no. 1, pp. 159–184, 2018, doi: 10.6018/rie.36.1.278301.
 - [29] M. Zapata, "Evaluación de competencias en entornos virtuales de aprendizaje y docencia universitaria," *Revista de Educación a Distancia. Sección de Docencia Universitaria en la Sociedad del Conocimiento.*, vol. 1, no. 1DU, 2010.
 - [30] I. Á. Valdivia, "Evaluación del aprendizaje en la universidad: Una mirada retrospectiva y prospectiva desde la divulgación científica," *Electronic Journal of Research in Educational Psychology*, vol. 6, no. 14, pp. 235–272, 2008.
 - [31] M. García Hípola, "Diseño de la auto, co-evaluación y rúbrica como estrategias para mejorar el aprendizaje," *VI Jornadas sobre Innovación Docente en Arquitectura (JIDA)*, 2018, doi: 10.5821/jida.2018.5433.
 - [32] E. Panadero, J. A. Tapia, and J. A. Huertas, "Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education," *Learn Individ Differ*, vol. 22, no. 6, pp. 806–813, 2012, doi: 10.1016/j.lindif.2012.04.007.
 - [33] M. P. G. Sanz, "La evaluación de competencias en Educación Superior mediante rúbricas: un caso práctico," *Revista electrónica interuniversitaria de formación del profesorado*, vol. 17, no. 1, p. 106, 2014.
 - [34] N. Battaglia, C. Neil, and M. De Vincenzi, "Software Engineering Competence-Based Learning in Collaborative Virtual Environments," *V IEEE Congreso Mundial de Educación en Ingeniería (EDUNINE2021)*, 2021.
 - [35] V. Torres-Sanz, J. Tramullas, J. A. Sanguesa, P. Garrido, and F. J. Martinez, "Rúbricas como estrategia de evaluación en entornos TICS [Rubrics as an evaluation strategy in ICT environments]," 2017.
 - [36] Y. Jiménez Galán, M. González, and J. Hernández, "Modelo 360 ° para la evaluación por competencias (enseñanza-aprendizaje)," *Innovación Educativa, IPN*, vol. 10, no. 53, pp. 43–53, 2010.
 - [37] J. Bigg, "Calidad del aprendizaje universitario. Cómo aprenden los estudiantes." Madrid: Narcea, 2003.
 - [38] L. María, M. Suárez, V. Daniel, G. Vera, and M. Suarez, "Gamificación en la enseñanza de Técnicas de Elicitación de Requisitos," *researchgate.net*, 2021.
 - [39] L. Machuca-Villegas, G. P. Gasca-Hurtado, L. M. R. Tamayo, and S. M. Puente, "Elementos de gamificación en el contexto de ingeniería de software," *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, vol. 2020, no. E27, pp. 718–732, 2020.
 - [40] M. Souza, R. Moreira, and E. Figueiredo, "Playing the Project: Incorporating Gamification into Project-based Approaches for Software Engineering

- Education,” in *sol.sbc.org.br*, 2020, pp. 71–80. doi: 10.5753/wei.2019.6618.
- [41] M. Del Carmen Pegalajar Palomino, “Implications of gamification in Higher Education: A systematic review of student perception,” *Revista de Investigacion Educativa*, vol. 39, no. 1, pp. 169–188, 2021, doi: 10.6018/RIE.419481.
 - [42] J. D. Britos *et al.*, “Los MOOC como propuesta para la estandarización de la calidad educativa,” in *XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2016)*, 2016.
 - [43] J. Onrubia, “Aprender y enseñar en entornos virtuales: actividad conjunta, ayuda pedagógica y construcción del conocimiento,” *Revista de Educación a Distancia (RED)*, no. 50, 2016, doi: 10.6018/red/50/3.
 - [44] F. Molina, N. Da, S. Ignacio, A.-C. Juan, and J. Goyeneche, “Evaluación y monitoreo de plataformas educativas,” *colibri.udelar.edu.uy*, 2021.
 - [45] Aguaded. Ignacio and R. Medina-Salguero, “Criterios de calidad para la valoración y gestión de MOOC,” *Revista Iberoamericana de Educación a Distancia*, vol. 18, no. 2, pp. 119–143, 2015.
 - [46] C. Neil, N. Battaglia, M. de Vincenzi, Edutec. Revista Electrónica, and undefined 2022, “Marco metodológico para el diseño de rúbricas analíticas,” *Edutec Revista Electrónica edutec.es*, vol. 80, pp. 198–215, 2022, doi: 10.21556/edutec.2022.80.2425.
 - [47] N. Garrido, C. Neil, and N. Battaglia, “Tecnología asistiva para personas con discapacidad en miembros superiores: un mapeo sistemático de la literatura,” *Memorias de las JAIIO*, vol. 8, no. 14, pp. 50–64, 2022, Accessed: Feb. 23, 2023. [Online]. Available: <https://publicaciones.sadio.org.ar/index.php/JAIIO/article/view/326>
 - [48] F. Maffei, C. Neil, and N. Battaglia, “Herramientas para determinar estilos de aprendizaje basadas en Inteligencia artificial,” in *XVII Congreso de Tecnología en Educación & Educación en Tecnología-TE&ET 2022*, Jun. 2022.

EL PORTAFOLIO COMO METODOLOGÍA DE TRABAJO EN UN GRUPO DE INVESTIGACIÓN DE INFRAESTRUCTURA IT

Ferrara, Damián Antonio; Díaz Lapérgola, María Ayelén; Monfroglio, Romeo Lorenzo

GIDAS, Grupo de I&D Aplicado a Sistemas informáticos y computacionales
Facultad Regional La Plata - Universidad Tecnológica Nacional

dferrara@frlp.utn.edu.ar; madlapergola@frlp.utn.edu.ar; rmonfrolio@alu.frlp.utn.edu.ar

RESUMEN

Se presenta una metodología de relevamiento de información y análisis de procesos de investigación, en el campo de la innovación en educación en informática.

Se trata de un proyecto que se implementará en el equipo Transformación Digital en Organizaciones aplicando Tecnologías de Infraestructura Informática basada en Computación en la Nube (TDO) del laboratorio “GIDAS - Grupo de I&D Aplicado a Sistemas informáticos y computacionales” de la Universidad Tecnológica Nacional – Facultad Regional La Plata.

El mismo consta en el uso de la estrategia didáctica de Portafolio Digital, en el que los investigadores del equipo irán realizando sus propios registros de trabajo, exponiendo sus reflexiones y análisis, mientras van confeccionando el estado de situación de sus investigaciones. Esto permitirá que, al finalizar el ciclo de investigación en curso, se posean elementos y datos que promuevan la divulgación en la comunidad científico-tecnológica.

A su vez, promueve la metodología de autoevaluación, potenciando el espíritu crítico y la búsqueda de las mejores prácticas para el abordaje de resultados óptimos; además de la evaluación entre pares o coevaluación, en la que cada miembro del equipo, por medio de cuestionamientos, presentaciones de opiniones y sugerencias, colaboran en el desarrollo del análisis y la potenciación de cada proyecto.

Palabras claves: Portafolio digital, E-Portfolio, TICS, autoevaluación, coevaluación o evaluación entre pares

CONTEXTO

Este Proyecto de I&D (PID) se encuentra enmarcado en los Incentivos del Ministerio de Educación Nacional y homologado por la Secretaría de Ciencia Tecnología y Posgrado (SCTyP) del Rectorado de la Universidad Tecnológica Nacional (código de homologación: SITCBLP0008124TC). El mismo, recibe la denominación “Marco de Innovación Tecnológica para Transformación Digital de Organizaciones integrando Inteligencia Artificial y Cloud Computing en la Producción de Software”

Su homologación es por un período de tres años, que va desde 01/01/2021 hasta el 31/12/2023. En este PID los avances y resultados son financiados por becas de alumnos de la SCTyP, becas de alumnos para Investigación de la Secretaría de Asuntos Estudiantiles de la Facultad Regional La Plata, becas de estímulo a las vocaciones científicas (EVC) del Consejo Interuniversitario Nacional (CIN) del Ministerio de Educación Nacional y becas para graduados de iniciación a la investigación y desarrollo (conocidas como BINID en el programa de becas de la SCTyP de la UTN).

El equipo de Transformación Digital en Organizaciones aplicando Tecnologías de Infraestructura Informática basada en Computación en la Nube (TDO) trabaja en la optimización “del uso de TICs (Tecnologías de

la Información y la Comunicación) y, en forma inclusiva, la aplicación práctica -en diferentes contextos- de las TACs (Tecnologías del Aprendizaje y del Conocimiento) en entornos educativos para el mejoramiento del proceso enseñanza-aprendizaje de tecnologías de infraestructura informática. Para esto, es necesario realizar actividades de Investigación & Desarrollo en temas como Cloud Computing en conjunción con procesos de enseñanza-aprendizaje mediado por tecnología.” [1] Es aquí donde se enmarca la propuesta de implementación de la estrategia que presentaremos a continuación.

“GIDAS - Grupo de I&D Aplicado a Sistemas informáticos y computacionales” es una unidad científico-tecnológica homologada en el sistema científico de la Universidad Tecnológica Nacional y en cuyo espacio se desarrollan las actividades de investigación, diseño de prototipos y desarrollo experimental de este equipo. GIDAS cuenta con financiamiento tanto para acciones e iniciativas del grupo de investigación como para los PIDs que se desarrollan en él [2]. Este proyecto se encuentra alineado al área de aplicación Tecnologías para Educación y Aula Siglo XXI.

INTRODUCCIÓN

El equipo de investigación TDO ha incrementado considerablemente el número de integrantes en el último año, lo que aporta una amplia diversidad de conocimientos, habilidades y competencias en el gran grupo. Esto nos obliga a implementar algún tipo de estrategia para conocer el estado y evolución en lo referente a la producción de saberes entorno a la investigación científico-tecnológica.

De este modo, nos proponemos lograr una mayor homogeneidad del equipo en cuanto a entendimiento de las distintas tecnologías.

Buscamos, entonces, generar un aprendizaje basado en la investigación de tecnologías de

infraestructura IT, teniendo en cuenta las temáticas que más motiven a cada uno de los miembro del equipo, acordando áreas de investigación en conjunto y que, a partir de ellas, surjan producciones que concluyan en publicaciones académicas, presentaciones en congresos o, simplemente, colaborar con acciones de extensión dentro de alguna cátedra de la carrera de Ingeniería en Sistemas de Información que tenga referencia al tema seleccionado.

Para tal fin se implementará la estrategia didáctica de portafolio digital, la cual consiste en una serie de herramientas tecnológicas donde el investigador puede reunir trabajos y lograr un seguimiento transversal a todo el desarrollo del proceso, que lleve finalmente a la realización de una auto evaluación y evaluación entre pares durante la investigación.

Este tipo de estrategia permite un amplio modo de aplicación creativa del y los investigadores, que, a modo de bitácora, va registrando no sólo el estado del proceso de su trabajo sino las reflexiones, cuestionamientos, estados de situación del proyecto y producción de nuevos objetos de estudios o planteamientos de áreas de mejora para el logro de los objetivos iniciales. Se trata de una colección deliberada, no azarosa, de materiales de aprendizaje surgido en este caso de la investigación científico-tecnológica en el campo de las TDO.

Esta “colección de trabajos excepcionalmente buenos o relevantes en la trayectoria de una persona, en concreto” [3] responde a unos objetivos que pueden ser desarrollados de manera individual o colectiva, según el modo de trabajo seleccionado para la investigación.

La estrategia didáctica seleccionada puede ser utilizada como una herramienta de trabajo, como instrumento de presentación o de evaluación. El primero de ellos promueve la recopilación de todas las tareas para depositar y diagnosticar necesidades; sirve para ver áreas

fuertes y débiles del proceso, y requiere una evaluación periódica.

El portafolio como instrumento de presentación reclama la selección y exhibición de los mejores trabajos por parte de quien o quienes lo producen; busca mostrar el nivel más alto alcanzado en el proceso, haciendo una capitalización de lo que se considera más importante en la investigación, lo que es de carácter valorativo en los resultados y lo que se desea mostrar.

Finalmente, el tercer modo de uso, llamado de evaluación o de diagnóstico, la documentación de lo investigado, entorno a los objetivos y contenidos tratados, requiere mayor intervención reflexiva del o los productores del conocimiento; este tipo de portafolio demanda establecer criterios y modos de relevamiento de datos, que sean claros y conocidos por todo el equipo.

Si bien cada modelo presenta unas características específicas, no reviste inconvenientes para poder trabajar con varios a la vez. Este proyecto propone a cada uno de los miembros que seleccione una de las modalidades o incluso genere la producción del portafolio con más de una de ellas.

De este modo, se intenta llegar a los objetivos que encuadran el funcionamiento de este grupo de investigación: mejorar la calidad y experiencia educativa en la formación de ingenieros-investigadores.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

La modalidad de portafolios digitales, también conocida como E-portfolios [4], se basa en la elaboración a través de medios y servicios electrónicos. Este, a diferencia de la modalidad tradicional, se caracteriza por el desarrollo superador en la calidad de los servicios para una persona tanto como a una comunidad; y su formato permite que el estilo no siga una

secuencia lineal, combinando herramientas tecnológicas de todo tipo.

En el grupo de investigación TDO se lo utilizará como medio para ir digitalizando el proceso de investigación. Allí serán colocadas las distintas producciones, así como todo lo que se considere de importancia para el análisis: textos, gráficos, elementos multimediales, entre otros.

Estos E-portfolios serán utilizados por el equipo de investigación como herramienta de evaluación procesual. Para ello serán establecidos en el equipo los criterios de valoración generales y cada uno de los miembros deberá buscar los criterios específicos para analizar su propio proceso.

Cada miembro o equipo de trabajo tendría que ir presentando sistemáticamente los avances de su producción, utilizando los objetivos inicialmente planteados y exponiendo sus propias reflexiones al respecto. En cada reunión plenaria, el resto del grupo podrá presentar cuestionamientos, opiniones, sugerencias que le permita al o los investigadores seguir produciendo sus avances en el campo de lo investigado.

Para ello se podrá seleccionar las mejores prácticas logradas en el proceso y presentar información sobre logros y dificultades surgidas; o aportar evidencias de las modificaciones originadas en la práctica respecto de los propósitos planteados al inicio de la investigación.

De esta manera, todos los integrantes de la comunidad actuarán como colaboradores de los procesos que cada investigador desarrolle y que irá materializando en el E-portfolio durante el transcurso del ciclo.

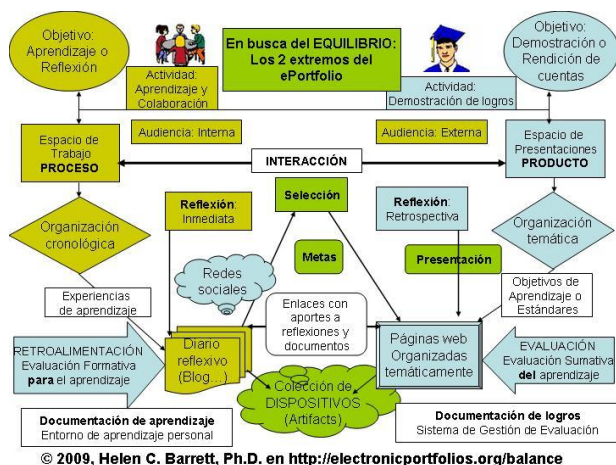


Figura 1: Audiencia interna y externa en un portafolio [4]

Para tal fin, se utilizará la aplicación Notion, que es una herramienta altamente capaz de cumplir los requisitos y desafíos que podemos encontrar a la hora de diseñar, administrar y compartir nuestros E-portfolios. Es un software de productividad individual y grupal, donde encontraremos, entre sus diversas opciones de personalización, algunos apartados que están hechos específicamente para el diseño de proyectos a largo y corto plazo.

Los workspaces de Notion son una zona donde todos los participantes pueden tener influencia en el trabajo que hace cada equipo. Estos pueden interactuar, evaluar y objetar sobre lo que ven en los demás grupos; se pueden encontrar plantillas predefinidas de trabajo, canvas y, la herramienta más importante, la barra de comandos (/). Esta nos da la posibilidad de poder no solo editar el texto, resaltar oraciones, diseñar matrices, crear hipervínculos a otras páginas, señalar fechas de entrega, objetivos y recursos importantes dentro y fuera de la aplicación, sino también, cómo Notion trabaja en conjunto con muchas otras aplicaciones; podemos hacer algunas opciones de pegado impresionantes, como, por ejemplo: PDF, GitHub, Seudocódigo, hasta incluso Twitter.

Dentro del grupo TDO, se usará para el establecimiento de fechas límites con

calendarios, creación de rutas de progreso del proyecto, divulgación grupal y juicio con los proyectos a compartir. Destacamos, además, que en la última actualización que esta aplicación presenta, se integra el uso de la inteligencia artificial (IA).

Finalmente, Notion promueve la creación de E-portfolios con una alta capacidad de seguimiento y modificación activa.

RESULTADOS ESPERADOS/OBTENIDOS

De este proyecto se espera que los investigadores del equipo TDO utilicen una herramienta de registro para sus tareas de investigación, que permita corroborar sus propios avances y, reflexionando de forma crítica sobre todo el trayecto de construcción de los trabajos de investigación, puedan lograr los objetivos iniciales propuestos.

Esto, también, permitirá tener un seguimiento y una evaluación del proceso por parte de cada miembro [5], retroalimentado desde sí mismo o de parte de un par. Esta última modalidad de evaluación conocida como “Evaluación entre pares” o “Coevaluación”, suele ser una modalidad poco utilizada en el campo universitario [6].

La misma propone compartir la responsabilidad de la examinación e intercambio de estrategias de cada miembro con sus pares. Permite una interpelación reflexiva para el desarrollo de la investigación, favoreciendo a una mejor administración y planificación, la participación crítica y constructiva de parte de todos.

La mencionada metodología tiene como finalidad crear un espacio para la valoración de los propósitos alcanzados, el análisis de las dificultades y áreas de mejora presentadas en el transcurso de la investigación; fomenta, por medio de un diálogo respetuoso, el intercambio de opiniones del equipo, con el fin de retroalimentar y mejorar la calidad de los

procesos de investigación y los resultados de la misma.

Así mismo, se espera un aumento y nivelación en los conocimientos de la comunidad de investigación conformada; además, de incrementar y potenciar el nivel de producciones escritas para la presentación a la comunidad científico-tecnológica.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación TDO está conformado por alumnos y egresados de la carrera de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional – Facultad Regional La Plata, además de profesionales de otras áreas interesados en temas de infraestructura IT.

La metodología de trabajo basada en portafolios digitales nos permitirá alcanzar, de manera eficaz y asertiva, los objetivos generales del equipo de investigación, proveer un espacio donde los estudiantes tengan acceso a recursos informáticos y lleven a cabo sus investigaciones; y haciendo que los distintos campos disciplinares que conforman esta comunidad científica encuentren un espacio de diálogo interdisciplinar, en post del avance de la ciencia y la tecnología.

Además de brindar un espacio que complemente lo desarrollado en las cátedras de la formación de ingenieros en el campo de los Sistemas de Información, proponemos un lugar para la búsqueda de nuevos enfoques de producción de saberes, basados en la manipulación de tecnologías, la búsqueda personal de inquietudes y problemáticas a desarrollar, dando lugar al análisis de los propios procesos de investigación y aprendizaje, la generación de interdiálogos colaborativos y constructivos del conocimiento, además de la motivación en la difusión de las investigaciones realizadas.

BLIBLIOGRAFÍA

[1] Portal Oficial del GIDAS, *Transformación Digital en Organizaciones aplicando Tecnologías de Infraestructura Informática basada en Computación en la Nube*, GIDAS: <https://gidas.frlp.utn.edu.ar/tdo>

[2] Portal Oficial del GIDAS: <https://gidas.frlp.utn.edu.ar>

[3] García-Doval, F. (2005). El papel de los portafolios electrónicos en la enseñanza-aprendizaje de las lenguas. Glosas didácticas, Revista Electrónica Internacional, n°14.

[4] Rey Sánchez, Ernesto & Escalera Gámiz, Águeda M^a. (2011) “El portafolio digital un nuevo instrumento de evaluación”. DIM: Didáctica, Innovación y Multimedia, Núm. 21, <https://raco.cat/index.php/DIM/article/view/247586>.

[5] Panadero, Ernesto & Alonso-Tapia, Jesús. (2013) “Autoevaluación: Connotaciones Teóricas y Prácticas. Cuándo Ocurre, Cómo se Adquiere y qué Hacer para Potenciarla en nuestro Alumnado”. Electronic Journal of Research in Educational Psychology, vol. 11, núm. 2, Universidad de Almería, España. <https://www.redalyc.org/pdf/2931/293128257012.pdf>

[6] González Velázquez, Lilia & González Velázquez, M^o. del Rosario (2014). “Evaluación de pares y coevaluación en estudiantes y docentes universitarios: una experiencia formativa para impulsar el modelo educativo”. International Journal of Developmental and Educational Psychology, vol. 2, núm. 1. <https://www.redalyc.org/pdf/3498/349851782053.pdf>

Enseñar programación en la universidad en tiempos de gran demanda de programadores

Gladys N. Dapozo, Cristina L. Greiner, Raquel H. Petris, Ana M. Company, María C. Espíndola

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste. 9 de Julio N° 1449. Corrientes. Argentina.
{gndapozo, cgreiner, rpétris, acompány, mcespíndola}@exa.unne.edu.ar

RESUMEN

Este proyecto de investigación abarca líneas de acción orientadas al estudio de métodos y herramientas para mejorar la enseñanza de la programación en la universidad y capacitar a los docentes de los niveles preuniversitarios en la utilización de estrategias educativas innovadoras para incorporar la programación en las escuelas. En este contexto, se observa una situación particular en las carreras de Informática: en los tres últimos años (2021-2023) se ha incrementado la cantidad de ingresantes, motivados por la demanda de programadores en el mercado laboral, asociada a buenos salarios. Esta expectativa, de inmediatez de resultados en la formación, puede afectar el proceso de enseñanza, por lo que se plantea analizar esta situación y proponer metodologías que articulen con las demandas de los estudiantes, resaltando que, más allá del código y de las herramientas, lo importante es que obtengan la formación necesaria para “resolver problemas”, utilizando los saberes propios de la disciplina.

Palabras clave: Enseñanza de la programación. Didácticas específicas. Mercado laboral

CONTEXTO

Las líneas de I/D corresponden al proyecto 21F016 “Modelos, metodologías y recursos para el desarrollo del pensamiento computacional”, acreditado por la SGCyT-UNNE, iniciado en el año 2022.

1. INTRODUCCIÓN

Existe consenso a nivel global sobre las competencias requeridas para el siglo XXI, así como de la importancia de la formación en las áreas STEM (Science, Technology, Engineering and Mathematics). Ante esta problemática, desde el 2008 se vienen implementando en los distintos países diversas políticas y estrategias para incrementar las vocaciones en TIC [1], [2].

A nivel de las políticas públicas, en Argentina, en el año 2018, el Consejo Federal de Educación ha declarado de importancia estratégica para el sistema educativo argentino la enseñanza y el aprendizaje de la programación desde niveles iniciales, y aprobó los Núcleos de Aprendizajes Prioritarios (NAP) de Educación Digital, Programación y Robótica [3].

En este marco, la iniciativa Program.AR, impulsada por la Fundación Sadosky, tiene como objetivo llevar la enseñanza y el aprendizaje de las Ciencias de la Computación a la escuela argentina. Incluye múltiples aspectos relacionados con la difusión y popularización de la disciplina, la generación de contenidos escolares y la formación docente, entre otros. Para la formación docente, Program.AR ha elaborado 2 cursos que incluyen contenidos, herramientas y actividades especialmente diseñadas para llevar la enseñanza de la programación a las escuelas, denominados “La Programación y su didáctica 1” y “La Programación y su didáctica 2”. La UNNE ha participado en varias convocatorias para esta capacitación,

desde el año 2013, y contribuye en forma activa con la formación docente en la región de influencia de la universidad [4]. En la actualidad, vinculada con la formación de profesionales informáticos, destaca una situación particular: la explosiva demanda de desarrolladores de software. En los últimos años, la industria tecnológica está protagonizando un boom a nivel mundial, la aceleración de la transformación digital de las empresas y la industria, como una de las consecuencias de la pandemia por coronavirus, ha generado una importante demanda de programadores, tanto de empresas radicadas en el país como en el exterior, con retribuciones que superan los ingresos promedios de otras actividades.

Según la Cámara de Software y Servicios Informáticos (CESSI) en su informe de evolución del empleo y cantidad de empresas de software (2015- 2021) muestra que el empleo registrado en Software ha crecido un 42,5% [5].

Conscientes de esta demanda, y su impacto en el desarrollo, los gobiernos impulsan capacitaciones específicas para proveer de recursos a la industria del software.

Ejemplos de iniciativas de este tipo son las siguientes:

Argentina Programa 4.0: Es un programa nacional, impulsado por el Ministerio de Economía de la Nación, que tiene como objetivo capacitar a personas en lenguajes y conocimientos sobre programación, testing y habilidades digitales para potenciar la empleabilidad en el sector del software y la tecnología [6].

Informatorio Chaco: Es un Laboratorio de Capacitaciones Informáticas que tiene por principal objetivo formar a jóvenes fomentando el desarrollo de sus *hard skills* (programación, herramientas y técnicas, etc.) y *soft skills* (colaboración, trabajo en equipo, flexibilidad, etc.), tal que puedan insertarse rápidamente en el mercado laboral. Se pretende potenciar la formación de los jóvenes en herramientas

específicas requeridas por la industria, promoviendo el desarrollo de ciertas habilidades y características personales, además de incentivar su preparación en idioma inglés [7].

Existen también plataformas orientadas a la capacitación en programación, entre las que se destacan:

Coderhouse: Plataforma de educación digital, creada en la Argentina en 2014, que ofrece capacitación en tecnologías mediante cursos virtuales.

Bootcamps: Son emprendimientos para capacitación en tecnologías. Son un fenómeno reciente, surgieron en 2011 en EEUU, pero con un gran crecimiento. Según estimaciones de LinkedIn en 2018 ya habían más de 300 bootcamps de programación alrededor del mundo, aunque en América Latina es, todavía, un fenómeno incipiente.

Platzi: Es una plataforma de educación online, que abarca diferentes áreas de aprendizaje: Desarrollo e ingeniería, Diseño y UX, Marketing, Negocios y emprendimiento, Producción audiovisual.

La carrera Licenciatura en Sistemas de Información de la Universidad Nacional del Nordeste (UNNE), con una trayectoria de más de 30 años, tiene como objetivo la formación de profesionales para el desarrollo de sistemas. Sin embargo, esta carrera no es ajena a las situaciones de retraso o abandono de los alumnos en el primer año de la carrera, compatible con el bajo nivel de graduación y la prolongación de la duración de los estudios de grado por encima del tiempo esperable, detectado en el sistema universitario argentino [8]. Preocupados por esta situación, el equipo docente lleva adelante distintas estrategias para mejorar la iniciación en la programación, siendo el aprendizaje de este campo de estudio una de las razones del abandono y/o retraso de los estudiantes.

Cabe señalar el notorio incremento de interesados en realizar esta carrera,

producto quizás de las acciones mencionadas de promoción de vocaciones TIC o por la demanda de programadores con salarios tentadores. Como demostración de este incremento, en la tabla 1 se muestra la evolución de los nuevos inscriptos de la carrera Licenciatura en Sistemas de Información. Se aclara que para el año 2023 este valor no está publicado, pero se considera la cantidad de estudiantes que han realizado el curso de ambientación que organiza la Facultad de Ciencias Exactas, previo al inicio de las actividades. Se puede apreciar que desde el 2018 al 2020, un promedio de 300 nuevos inscriptos, en tanto que, en los últimos 3 años, del 2021 al 2023, el promedio es de 600 estudiantes. También se puede observar que se ha incrementado la participación de las mujeres, siendo este un objetivo que ha llevado a la implementación de distintas estrategias, a nivel mundial y nacional, para acercar a las jóvenes a la formación en tecnologías.

Tabla 1: Nuevos inscriptos Licenciatura en Sistemas de Información

Año	Nuevos Insc.	Varones	Mujeres	%Mujeres
2018	286	226	60	21%
2019	257	202	55	21%
2020	397	309	88	22%
2021	520	388	132	25%
2022	696	533	163	23%
2023	779	563	217	27%

Fuente: UNNE en cifras. Link: <https://rb.gy/cx1mre>

Estos estudiantes poseen un perfil diferente al de años anteriores porque, gracias a las políticas y la difusión de las demandas, muchos de ellos han realizado o están realizando capacitaciones en programación. El desafío de la enseñanza actual es mantener la motivación de los estudiantes y lograr que comprendan la importancia de los fundamentos de la disciplina, del conocimiento de métodos y el desarrollo de habilidades que se requieren para ser un profesional de informática con sólidos conocimientos

que le permitirán crecer y evolucionar en las organizaciones. Más allá del código y de las herramientas específicas de programación, lo importante es que obtengan la formación necesaria para “resolver problemas”, utilizando los saberes propios de la disciplina.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación de este proyecto están enfocadas en:

- Estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática.
- Evaluación de métodos y herramientas para la enseñanza de programación y robótica para docentes de los niveles educativos no universitarios.

3. RESULTADOS OBTENIDOS/ESPERADOS

En la línea de estrategias educativas para la enseñanza de programación en carreras de Informática, en [9] se presentan los resultados de la evaluación de habilidades de pensamiento computacional, tomadas de un test diagnóstico realizado al inicio de una asignatura de introducción a la programación. Los estudiantes que participaron de esta experiencia son ingresantes de una carrera de Informática. El dictado de esta asignatura ha regresado a la presencialidad plena y al inicio del curso se les tomó un test diagnóstico con actividades de resolución de problemas que los alumnos tenían que realizar en forma intuitiva, sin ninguna indicación adicional, más que el planteo de la actividad. Adicionalmente se recabó información acerca de la experiencia previa que tenían en programación. Los resultados obtenidos permitirán consolidar estrategias de enseñanza que se enfoquen en la adquisición de habilidades de pensamiento computacional y enfatizar la idea de que la programación no es una actividad

meramente técnica, sino que requiere un conjunto de habilidades combinadas de resolución de problemas.

En la línea de formación docente en [10] se describe la experiencia del dictado del curso “La Programación y su Didáctica 2” en FaCENA (UNNE), que se ofrece como continuación del curso La Programación y su Didáctica 1, el cual aborda con mayor profundidad los conceptos y técnicas de programación, tales como, estructuras de datos, algoritmos de ordenamiento y de búsqueda. Se presentan los contenidos, la metodología de dictado y las dificultades que aparecieron en el desarrollo del curso. Se destaca que este dictado se implementó en modalidad híbrida, accediendo a la solicitud expresa de los docentes participantes dado que esta modalidad les permitió gestionar mejor su tiempo, evitando los traslados. En cuanto a los contenidos y las estrategias de enseñanza se ofrecieron diversas actividades y herramientas que permiten a los docentes contar con recursos para transmitir estos conceptos a sus estudiantes. Se señalaron algunos aspectos que deberían ser revisados, como la importancia de la característica de conexión off-line que poseen los entornos utilizados en el curso, que favorece una mayor aceptación por parte de los docentes participantes. Así también, se observa como aspecto que requiere una revisión, la alta carga horaria del curso que se prolonga a lo largo de tres meses, lo que requiere de un gran esfuerzo por parte de los cursantes que suman tiempo extra a las actividades normales que tienen durante la semana.

En cuanto a Resultados Esperados, el proyecto continuará con las siguientes actividades:

En la línea sobre estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática

- a) Relevar y analizar las ofertas de capacitación en programación que están disponibles para los interesados.
- b) Realizar propuestas educativas que articulen con aquellos aspectos que se valoran en el ámbito laboral, por ejemplo, el trabajo colaborativo, la noción de proyecto, las habilidades de autoaprendizaje, entre otras, que contribuyan a lograr el perfil profesional que la carrera define.

Algunos avances realizados en esta línea son los resultados obtenidos en una encuesta realizada a los ingresantes 2023 de la LSI.

Respondieron la encuesta 779 estudiantes, de los cuales 72,3% son varones, 27,2% mujeres y 0,5% prefiere no decirlo.

En cuanto al rango de edades, la Fig. 1 muestra que el 49% tiene 19 o menos años.

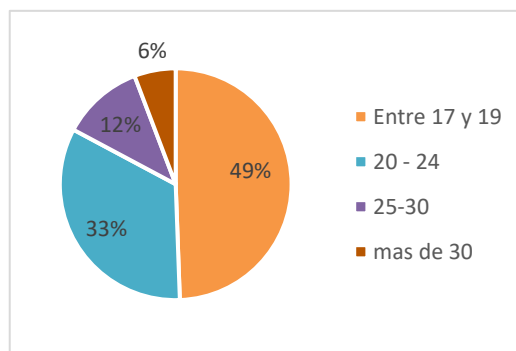


Fig.1: Rango de edades de los ingresantes LSI

Otras consultas fueron:

- Si Trabaja: El 30% contestó que SI.
- Si tiene experiencia en programación: El 26% contestó que SI.
- Si cursó o está cursando otra carrera universitaria: El 42% contestó que SI. Esto evidencia la reconversión que están haciendo algunos profesionales.
- Si está realizando alguna actividad de capacitación en programación. Si la respuesta fue positiva, que indique que capacitación: El 17% contestó que SI.

En la Fig. 2 se muestran plataformas y/o tecnologías que los alumnos indicaron.

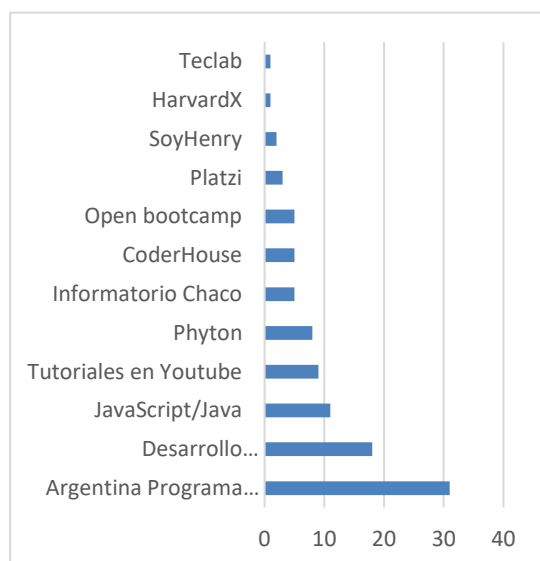


Fig.2. Plataformas y/o tecnologías de capacitación

De estos resultados, se destaca la capacitación por su cuenta a través de tutoriales en YouTube.

Esta información será un insumo inicial para avanzar en el conocimiento de esta problemática y sustentar propuestas superadoras.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan cinco docentes investigadoras, entre ellas dos tesis de posgrado que desarrollan su Trabajo Final de la Maestría en Tecnologías de la Información de la UNNE.

5. REFERENCIAS

[1] Ochoa, L. A., Valenzuela, A., Estela, D., Márquez, F. (2018). "La indagación como estrategia para la educación STEAM". Organización de Estados Americanos. Disponible en: <https://tinyurl.com/y9ptbgs1>

[2] The Royal Society. (12 de 01 de 2012). "Shut down or restart? The way forward for computing in UK schools". Disponible en: <https://royalsociety.org/-/media/education/computing-in-schools/2012-01-12-computing-in-schools.pdf>

[3] EDUC.AR. "Núcleos de Aprendizajes Prioritarios (NAP) de Educación Digital, Programación y Robótica". Disponible en: <https://www.educ.ar/recursos/150123/nap-de-educacion-digital-programacion-y-robotica/download/inline>

[4] Dapozo, G., Greiner, C., Petris, R. (2021). "Acciones en la UNNE para promover el pensamiento computacional mediante la enseñanza de la programación"- En "Pensamiento Computacional, Programación Creativa y Ciencias de la Computación para la Educación: Reflexiones y experiencias desde América Latina". Quito. Ecuador. ISBN: 978-9978-55-203-2. Ediciones CIESPAL, 2021. PP 121-141. Disponible en: <https://ediciones.ciespal.org/index.php/ediciones/catalog/view/30/34/230-1>

[5] Cámara de la Industria Argentina del Software (CESSI). "Evolución del empleo y la cantidad de empresas - Software - 2015 a 2021". Disponible en: <https://cessi.org.ar/wp-content/uploads/2022/06/Software-principales-indicadores-a-2021-1.pdf>

[6] Ministerio de Economía. "Argentina Programa 4.0". Disponible en: <https://www.argentina.gob.ar/economia/conocimiento/argentina-programa>

[7] Polo IT-Chaco. Informatorio. Disponible en: <https://poloitchocho.ar/informatorio/>

[8] García, A. "Políticas institucionales para mejorar la retención y la graduación en las universidades nacionales argentinas". Debate Universitario, 4(7), 7-24. 2015

[9] Gladys N. Dapozo, Cristina Greiner, Raquel H. Petris, Emanuel Irrazábal, Ana María Company, María Cecilia Espíndola, Yanina Medina. "Evaluación de habilidades de pensamiento computacional al inicio de una asignatura de programación en una carrera de Informática". Libro de actas XVII Congreso de Tecnología en Educación & Educación en Tecnología. UNER. ISBN 978-950-698-522-6. Pag. 138-147

[10] Ana M. Company, María C. Espíndola, Raquel Petris, Numa Badaracco, Gladys Dapozo. "Dictado del curso "La Programación y su Didáctica 2" en FaCENA – UNNE". Libro de actas de las Segundas Jornadas de Didáctica de las Ciencias de la Computación. Pag. 71-75. Disponible en: <https://repositorio.unne.edu.ar/handle/123456789/50765>

EYE TRACKING PARA LA ENSEÑANZA DE LA PROGRAMACIÓN

Guillermo Guerrero Jorge P. Rodríguez Laura A. Cecchi

email: guillermo.guerrero@est.fi.uncoma.edu.ar, {j.rodrig,lcecchi}@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

RESUMEN

Actualmente, existe un gran interés en alentar mediante el aprendizaje de la programación, el desarrollo de la capacidad de resolución de problemas, en estudiantes de todas las modalidades de la educación obligatoria. La interacción hombre-máquina mediante los medios convencionales (teclado y mouse), suele ser un obstáculo para una población de personas con discapacidades motrices al momento de aprender sobre programación. Eye tracking es un conjunto de técnicas que ofrecen una forma alternativa de interacción con la computadora a partir del seguimiento ocular. El objetivo de esta línea de investigación, desarrollo y transferencia es diseñar e implementar un entorno de programación open source, que utilice estas técnicas, para producir código, empleando la vista como medio de interacción hombre-máquina. En este sentido, se pretende producir conocimiento en relación a las posibilidades del eye tracking, como dispositivo tecnológico, destinado a favorecer el aprendizaje de la programación en personas con disminución de la movilidad. Se espera involucrar tanto a estudiantes con discapacidades motrices, como también a docentes especialistas en Educación Especial, en los procesos de enseñanza y aprendizaje de la programación a partir de este nuevo entorno.

Palabras Clave: Enseñanza de la Programación - Educación en Ciencias de la

Computación - Eye Tracking - Interacción hombre-máquina - Plataforma para la Enseñanza de la Programación

CONTEXTO

Esta línea de investigación se desarrolla en el contexto de los temas de interés que promueve el Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA), de la Facultad de Informática. En particular, se enmarca en el ámbito de dos proyectos de investigación miembros del GILIA, ambos financiados por la Universidad Nacional del Comahue: *Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación (04/F022)* y *Tecnologías semánticas para el desarrollo de agentes inteligentes (04/F020)*.

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el marco de los proyectos antes mencionados y por el Consejo Interuniversitario Nacional (CIN) con una Beca de Estímulo a las Vocaciones Científicas para Estudiantes 2021. Los proyectos de investigación tienen una duración de cuatro años, a partir de enero 2022, y la beca CIN una duración de 1 año, finalizando esta última en agosto de 2023.

1. INTRODUCCIÓN

En la actualidad, son ampliamente conocidos los beneficios de aprender sobre Ciencias de la Computación en general y sobre programación en particular. Estos beneficios están orientados

principalmente al desarrollo de las capacidades para la resolución de problemas, particularmente, la construcción de estos aprendizajes son alentados durante las etapas iniciales de la educación obligatoria[1, 2]. A raíz de esto, es que actualmente existe una tendencia a nivel mundial de incorporar, a los diseños curriculares, asignaturas que se relacionan directamente con la enseñanza de la programación[3, 4, 5, 6, 7, 8].

La principal herramienta con la cual un estudiante es capaz de interactuar con una computadora, y por ende con un lenguaje de programación, es mediante el uso de teclado y mouse. Estos medios físicos plantean, en términos de coordinación y control de la motricidad fina, un obstáculo a los estudiantes con capacidad motriz completa. Por lo que, para los estudiantes que poseen capacidad motriz reducida o un impedimento de movimiento en sus extremidades, resulta aún más compleja la tarea de interactuar con una computadora.

Eye tracking, seguimiento ocular, es un conjunto de técnicas y métodos que permiten medir, analizar e interpretar datos sobre la posición y movimiento del globo ocular, el punto en que se encuentra la vista de la persona en un momento dado, cuánto tiempo se concentra la vista en un punto determinado, qué trayectoria sigue y el tamaño de la pupila[9]. Esta técnica de interacción hombre-máquina es ampliamente usada para ayudar a personas con discapacidades motrices a comunicarse[10].

Existen distintas propuestas de desarrollo open source que ofrecen medios de entrada alternativos orientados a favorecer y agilizar la navegación por el código de programas que se apoyan en cámaras para realizar ediciones simples de código, como por ejemplo selección y desplazamiento de fragmentos pequeños de texto. Éstas tienen la intención de proporcionar una interacción más fluida,

basada principalmente en atajos de teclado y eye tracking del usuario. Ejemplos de estas herramientas son Javardeye[11] y Eyenav[12]. Este tipo de productos no busca ubicarse como entorno de desarrollo basado en eye tracking, sino que están enfocados sobre el proceso de revisión y mantenimiento del código. En consecuencia, estas herramientas están destinadas a personas que saben programar y no están orientadas a la enseñanza y el aprendizaje de la programación.

Existe una población estudiantil que presenta discapacidad motriz que podría resultar beneficiada por entornos destinados al aprendizaje de la programación, que cuenten con medios de interacción basados en el eye tracking, para facilitar la codificación y ejecución de programas.

Con esta problemática como objeto de estudio, es que en esta línea de investigación y desarrollo se propone diseñar e implementar una herramienta open source que incorpore un entorno de programación que haga uso del eye tracking como método de interacción principal.

Este trabajo está estructurado como sigue. En la siguiente sección, se describe la línea de investigación y desarrollo actual. En la Sección 3, se indican algunos resultados obtenidos y esperados. Finalmente, en la última sección se comentan aspectos referentes a la formación de recursos humanos en esta temática.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación, desarrollo y transferencia propone definir un lenguaje de programación simple, a fin de diseñar e implementar un prototipo para un entorno de programación que utilice la técnica eye tracking para el desarrollo de programas, es decir un ambiente que permita producir código utilizando como medio de interacción hombre-

máquina el movimiento ocular. En este sentido, el lenguaje de programación sobre el que se está trabajando se circunscribe en una cantidad acotada de instrucciones, las que están representadas por formas gráficas y, en relación a las estructuras de control, están delimitadas a operaciones de la estructura secuencial.

Este trabajo se desarrolla en el marco de las actividades de los proyectos de investigación *Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación y Tecnologías Semánticas para el desarrollo de agentes inteligentes*.

Por un lado el proyecto *Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación*, busca desarrollar modelos teóricos que contribuyan a la producción de un marco conceptual de referencia, a fin de asistir la inclusión de la computación en la educación. En este sentido se plantea como objetivo el diseño, implementación y evaluación de herramientas que los soporten. Por otra parte, el diseño de lenguajes caracterizados por gramáticas formales y el desarrollo de entornos que cumplan los principios de la Web Semántica son objetivos específicos del proyecto de investigación *Tecnologías semánticas para el desarrollo de agentes inteligentes*.

La definición y caracterización formal de un lenguaje de programación adecuado para este contexto, como el desarrollo de un entorno de programación que contribuya a hacer posible la construcción de saberes del área de conocimiento algoritmos y programación a personas con discapacidades motrices, es de interés para ambos proyectos, colaborando en su diseño e implementación.

Así, un usuario que no cuente con la capacidad de interactuar por medio del teclado y mouse, podrá desarrollar, modificar y ejecutar

programas, usando la vista como principal medio de interacción con la computadora.

El propósito es contribuir a la reducción de las brechas producidas por los métodos de entrada convencionales (por ejemplo, mouse y teclado) al buscar incluir a estudiantes con discapacidades motrices a los ambientes destinados a aprender prácticas y conceptos sobre el área de conocimiento de algoritmos y la programación.

Desarrollar una herramienta como la que se propone en esta línea de investigación y desarrollo, permitirá ampliar las posibilidades didácticas y la población destino.

3. RESULTADOS OBTENIDOS y ESPERADOS

Inicialmente se realizó un relevamiento, sobre distintas tecnologías y librerías existentes, que permitieran obtener información ocular de un usuario; entre ellas WebGazer, OpenGazer, PyGazer y MediaPipe.

La librería seleccionada es MediaPipe porque es la que otorga mayor rendimiento, mayor tasa de frames al capturar dispositivos de entrada de video y una de las mejores estimaciones del modelo de Face Mesh. Asimismo, ofrece soporte para diferentes lenguajes de programación y otorga un mejor control sobre las valoraciones otorgadas por los modelos de estimación. Además proporciona una documentación actualizada, detallada para su uso y alcance, lo que se traduce en una mejora importante durante el proceso de implementación.

Uno de los resultados obtenidos, es un entorno interactivo basado en la librería antes mencionada que permite detectar, mediante el uso de una cámara, la estimación de la posición de su iris. Este se utiliza para determinar una de las cuatro direcciones posibles que el usuario puede mirar.

Por otra parte, se definió un lenguaje de programación con cuatro instrucciones, que se asociaron a las cuatro direcciones de mirada del usuario. Con esta información es posible posicionar un selector sobre los elementos que componen el lenguaje de programación a enseñar. Además permite obtener un input del usuario para confirmar la selección de elementos del lenguaje. De esta manera, es factible identificar puntos claves o landmarks en tiempo real sobre éste y utilizarlos para crear indicadores de distancias, que son usados en conjunto con la posición estimada de iris para medir y conocer la dirección de la mirada. Actualmente, el prototipo permite responder a los inputs del eye tracking del usuario para seleccionar elementos del lenguaje que le posibilite la construcción de un programa.

Como trabajo futuro, se pretende completar el proceso de codificación, ejecución y visualización de resultados. En este sentido, se proyecta desarrollar la interfaz que permita la ejecución del programa y la presentación de resultados en un ambiente gráfico.

Se espera lograr un entorno de programación sencillo, lúdico e interactivo enriquecido con la técnica eye tracking, que ofrezca nuevas oportunidades en la construcción de programas y que permita ampliar la disponibilidad de recursos didáctico disciplinares, asociadas a los procesos de desarrollo de código, que esté destinado a los diferentes niveles del sistema educativo.

Asimismo, se espera poder transferir este desarrollo a la comunidad, involucrando a docentes especialistas en Educación Especial, en los procesos de enseñanza y aprendizaje de la programación, lo que contribuirá a mejorar la herramienta.

4. FORMACION DE RECURSOS HUMANOS

En la Universidad Nacional del Comahue, Facultad de Informática, se otorgaron Becas CIN para estimular la vocación científica. Una de esas becas fue otorgada a uno de los autores de este trabajo, que está desarrollando su tesis de grado de la Licenciatura en Ciencias de la Computación en esta temática. Por otra parte, otro de los autores de este trabajo está desarrollando su trabajo final integrador de la Maestría en Enseñanza en Escenarios Digitales, posgrado que dicta la AUSA (Asociación de Universidades Sur Andina), de la que la Universidad Nacional del Comahue forma parte.

5. BIBLIOGRAFÍA

- [1] Seymour Papert. Children, computers and powerful ideas, 1990.
- [2] Jeannette Wing. Computational thinking. Communications of the ACM ,49:33–35, 03 2006.
- [3] The K–12 Computer Science Framework. The Computer Science Teachers Association. New York, 2016.
- [4] María Bonello and Hernán Czemerinski. Program.ar: una propuesta para incorporar ciencias de la computación a la escuela argentina. 2015.
- [5] Megan Smith. Computer science for all. The White House, 2016.
- [6] Katrina Falkner, Keith Quille, Sue Sentance, Rebecca Vivian, Sarah Barksdale, Leonard Busuttil, Elizabeth Cole, Christine Liebe, Francesco Maiorana, and Monica McGill. An international comparison of k-12 computer science education intended and enacted curricula. pages 1–10, 11 2019.
- [7] Royal Society. After the reboot: Computing education in UK schools. Policy Report, 2017.

[8] Consejo Provincial de Educación de la Provincia de Neuquen. Diseño Curricular Jurisdiccional de los tres primeros años de la Escuela Secundaria Neuquina. Resolución N°1463/18, 2018.

[9] Sylwester Bialowas and Adrianna Szyszka. Eye-tracking research , pages 40–60. 2021.

[10] M. Maboe, Mm Eloff, and Marthie Schoeman. The role of accessibility and usability in bridging the digital divide for students with disabilities in an e-learning environment. pages 222–228, 09 2018.

[11] André L. Santos. Javardeye: Gaze input for cursor control in a structured editor. Programming '21, page 31–35, New York, NY, USA, 2021. Association for Computing Machinery.

[12] Stevche Radevski, Hideaki Hata, and Kenichi Matsumoto. Eyenav: Gaze-based code navigation. pages 1–4, 10 2016.

DEMANDA LABORAL Y COMPETENCIAS EN INTELIGENCIA ARTIFICIAL DENTRO DE CARRERAS GRADO EN CIENCIAS APLICADAS

Silvia Lanza-Castelli ⁽¹⁾ María Paula González ^(1,2)

(1) Licenciatura en Informática
Decanato de Ciencias Aplicadas
Universidad Siglo 21, Córdoba, Argentina
silvia.lanzacastelli@ues21.edu.ar, paula.gonzalez@ues21.edu.ar

(2) Dpto. de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur, Bahía Blanca, Argentina
mpg@cs.uns.edu.ar

RESUMEN

La llegada de la Inteligencia Artificial cambia nuestra realidad atravesándola de manera disruptiva. En particular, las experiencias de enseñanza-aprendizaje universitarias enfrentan un cambio radical de paradigma. En este escenario, abordar la formación de nuevas competencias resulta fundamental para poder asegurar un alto nivel de formación integral que potencie una efectiva inserción laboral y una tasa de desempeño profesional exponencial.

En concordancia con lo anterior, este trabajo presenta una nueva línea de investigación que plantea definir un conjunto minimal de competencias procedimentales en uso de herramientas y tecnologías propias de la Inteligencia Artificial fuertemente requeridas por la Industria 4.0 a corto y mediano plazo. Dicho conjunto minimal de competencias debería incluirse dentro de todas las carreras de grado de formación en ciencias aplicadas en nuestro país. Se parte de la utilización de herramientas de medición propias del denominada Información del Mercado Laboral (Labor Market Information) como metodología de diagnóstico situacional, las cuales son aplicadas a un conjunto de empresas relevadas en diferentes proyectos de alto impacto internacional.

En base a los resultados obtenidos, se establece el mencionado conjunto minimal para avanzar luego hacia la observación y análisis de su cubrimiento dentro de la oferta académica actualmente disponible. Debido al grado de

avance actual respecto de la concreción de la línea de investigación presentada, solo las carreras del Decanato de Ciencias Aplicadas de la Universidad Siglo 21 son consideradas. Como objetivo final se busca aportar conocimiento y datos reales que ayuden a plantear ofertas académicas adecuadas de fuerte impacto socio-económico y alta calidad.

Palabras clave: Inteligencia Artificial, Carreras de Grado en Ciencias Aplicadas, Competencias/Habilidades 2030, Demanda Laboral

ENCUADRE CONTEXTUAL Y OBJETIVOS

La sociedad 4.0 plantea nuevos desafíos en todos los ámbitos de la realidad. Un mundo laboral en donde la norma es la contingencia despliega nuevos desafíos, en particular los asociados al uso de herramientas basadas en Inteligencia Artificial (IA) (Popkova, E. G., & Gulzat, K., 2020). En consecuencia, las experiencias de enseñanza-aprendizaje universitarias enfrentan un cambio radical de paradigma (Ocaña-Fernández et al., 2019), en donde abordar la formación de nuevas competencias resulta fundamental para poder asegurar un alto nivel de formación integral que potencie una efectiva inserción laboral y una tasa de desempeño profesional exponencial (Guile D., 2018)

Tal como se plantea en el Proyecto OCDE (Felder et al, 2018), dos preguntas claves deben

ser consideradas: ¿Qué conocimientos, habilidades, actitudes y valores necesitarán los estudiantes de hoy? ¿Cómo pueden los sistemas de formación desarrollar estos conocimientos, habilidades, actitudes y valores de manera efectiva? La necesidad de capacitación es urgente. Muchas economías están experimentando un desempleo masivo. Gran cantidad de calificaciones tradicionales o programas de capacitación, como los aprendizajes, operan dentro de un marco de "tiempo cumplido" (un mínimo de dos años) (Alfredo, M. A., Granovsky, P., & Verchelli, V., 2021).

En el contexto anterior, un gran número de las habilidades que se vuelven redundantes o dan paso a la IA no simbólica son habilidades de orden inferior, mientras que las que se crean o surgen dentro del nuevo escenario 4.0 son de orden superior y en su mayoría conceptuales (diseño, creatividad, filosofía, ética), derivando en la necesidad de nuevas competencias. A medida que el empleo se vuelva más conceptual, las experiencias educativas formales de grado se fusionarán cada vez más con experiencias del mundo real; y probablemente las propuestas didácticas migrarán hacia modelos basados en las sinergias e interacción entre la academia y el mundo laboral (Guitérrez-Portela, F et al., 2019).

Este trabajo presenta una nueva línea de investigación que plantea como objetivo general la definición de un conjunto minimal de competencias procedimentales en uso de herramientas y tecnologías propias de la Inteligencia Artificial fuertemente requeridas por la Industria 4.0 a corto y mediano plazo. Dicho conjunto minimal de competencias debería incluirse dentro de todas las carreras de grado de formación en ciencias aplicadas en nuestro país.

Los objetivos específicos incluyen profundizar el conocimiento sobre el marco de aprendizaje conceptual de demandas propuestas por la OECD para el Futuro de la Educación y las Habilidades 2030 (Felder, F. et al, 2018); analizar marcos de orientaciones actuales tanto sobre las competencias básicas en IA como sobre

recomendaciones éticas asociadas (Board, D, 2019) identificar el conjunto minimal de competencias asociadas al uso de herramientas basadas en IA que cubra las habilidades del mundo laboral antes identificadas; obtener métricas (indicadores) asociados; y comparar los resultados obtenidos en función de las tecnologías evaluadas. Además, se busca favorecer la formación de recursos humanos en las áreas involucrada.

La línea de investigación, comenzada en el 2022, se desarrolla en el marco del proyecto Herramientas Basadas en Inteligencia Artificial para Analizar el Impacto en la Demanda de Habilidades del Mercado Laboral de la Secretaría de Investigación y Transferencia Científica de la Universidad Siglo 21,¹ dentro de la línea temática relacionada con la Innovación Educativa en el área de las Ciencias Aplicadas, cuyo enfoque se orienta a la investigación del diseño o gestión de cambios significativos en una o más de las dimensiones que impacten en los procesos educacionales: pedagogía, currículo, didáctica, tecnología u organización.

ENCUADRE METODOLÓGICO

En primer lugar, una serie de actividades introductorias relacionadas con el relevamiento del estado del arte, la selección de herramientas de consulta y búsqueda automática y su configuración de parámetros iniciales fueron realizadas. Para el relevamiento del estado del arte se desarrolló un instrumento mediante un formulario para la revisión sistemática de la literatura SLR (ó Systematic Literature Review) (García-Peñalvo, 2022) existente en buscadores Google Scholar, Scopus y distintos repositorios internacionales. Se realizó una entrevista al Executive Head Digital del Milton College & Grainger de la University College London² del Reino Unido (Langford, 2022) junto a expertos en educación digital.

Luego de esas actividades iniciales, el encuadre metodológico utilizado propone un abordaje incremental en ciclos experimentales, que

¹ <https://21.edu.ar/investigacion>

² <https://www.ucl.ac.uk/>

incluye en cada ciclo experimental el análisis de un contexto real cada vez más globalizado no solo geográficamente sino también en relación a las carreras de grado consideradas. Dentro de cada ciclo, se parte de la utilización de herramientas de medición propias del denominada Información del Mercado Laboral (Labor Market Information o LMI) (Maurya, M., & Shah, N., 2014) como metodología de diagnóstico situacional. Se aplican estas herramientas a diferentes fuentes de datos provenientes de empresas buscando identificar que competencias y/o herramientas basadas en IA se requieren en los perfiles profesionales y en las prácticas y procesos productivos de cada una.

Se utiliza un enfoque de ingeniería reversa que abarca tanto dimensiones cuantitativas como cualitativas. En este sentido, aplicando triangulación cualitativa y técnicas de interpretación de fenómenos orientados a minimizar el sesgo subjetivo en los resultados, se busca responder interrogantes como ¿Los programas de formación de grado en ciencias aplicadas están en condiciones de incluir en su curricula el uso transversal de herramientas basadas en IA para anticipar y emparejar los desajustes de habilidades demandadas del mundo laboral?, ¿Podrán las herramientas de IA y el LMI bosquejar el impacto en la demanda de habilidades del mercado laboral y disminuir los desajustes?, entre otras.

En base a los resultados obtenidos, se establece el conjunto minimal mencionado en la sección anterior, para avanzar luego hacia la observación y análisis de su cubrimiento dentro de la oferta académica actualmente disponible. Para este paso se introduce el relevamiento documental y la revisión bibliográfica de Programas Analíticos y Proyectos Curriculares propios de carreras de grado de formación en ciencias aplicadas. Seguidamente, se incorporan técnicas del enfoque sistémico y un diseño de investigación explicativo para, de manera contextualizada, derivar relaciones causa-efecto que resalten en las premisas la innovación educativa provocada

por la incorporación de herramientas IA en la curricula antes mencionada; y en las conclusiones los ítems candidatos al conjunto minimal de competencias demandadas por el mercado laboral (rescatadas en el trabajo de ingeniería reversa previamente realizado).

Finalmente, un proceso de validación es propuesto. Todo lo obtenido se sujeta a la opinión de expertos y especialistas en educación en IA y en Didáctica de las Ciencias. Además, se incluye el relevamiento de la opinión de estudiantes avanzados y jóvenes profesionales en actividad. Los resultados obtenidos son utilizados como como punto de partida del próximo ciclo experimental de tal manera que el conjunto minimal derivado dentro del ciclo n mantenga sus elementos y eventualmente los amplíe en el ciclo $(n+1)$.

TRABAJO REALIZADO Y EN MARCHA

A seis meses del comienzo del proyecto mencionado en la primera sección del presente trabajo, las actividades introductorias fueron ejecutadas; y las dos primeras partes del primer ciclo de la metodología reportada en la sección anterior están siendo llevadas a cabo.

En particular, para esta primera aproximación se incluyó a un grupo de empresas que utilizan IA dentro de sus prácticas y que son sugeridas por la Ejecutiva Digital Mrs Penny Langford de MK College.³ Además, se utilizaron datos disponibles en el Observatorio de Tendencias Sociales y Empresariales de la Universidad Siglo 21,⁴ enfatizándose aquellas enfocadas en la transformación digital. También se consideraron las empresas mencionadas en un informe de UNESCO (Unesco, 2021). En cuanto a las propuestas académicas, solo las carreras del Decanato de Ciencias Aplicadas de la Universidad Siglo 21 son consideradas.

Como primeros resultados cuantitativos del proceso de ingeniería reversa efectuado, y teniendo en cuenta los interrogantes establecidos

³ <https://mkcollege.ac.uk/>

⁴ <https://21.edu.ar/identidad21/siglo-21-investiga-habra-mayor-empleo-humano-en-empresas-con-la-incorporacion-de-tecnologias-digitales>

(ver sección Encuadre Metodológico), se concluye que el 19% de las empresas consultadas implementaron activamente la IA, mientras que el 51% informan que están explorando su incorporación dentro de sus procesos productivos. Estos resultados están por debajo de la realidad internacional, en donde según consultoras de primera línea como IBM y Morning Consult⁵ el 35% de las empresas a nivel mundial adoptaron ya a la IA como una estrategia fundamental; y el 42% se halla en el proceso de análisis de su incorporación.

En cuanto a los resultados cuantitativos, se destaca tanto el uso de técnicas inteligentes para migrar del acercamiento predictivo al análisis prescriptivo de datos como el uso de sistemas de recomendación inteligentes asociado a la estrategia de Inteligencia de Negocios (Business Intelligence); la incorporación de herramientas IA propias de las Ciencias de Datos para el procesamiento de Big Data; el procesamiento de lenguaje natural mediante técnicas inteligentes de reconocimiento de voz complementadas con técnicas de visión artificial para el manejo de la Inteligencia de Clientes (Customer Intelligence); el uso de chatbots para la atención de quejas, la presencia de Deep y Machine Learning para la segmentación de usuarios, personalización y definición de las campañas de Marketing; el etiquetado inteligente y los almacenes inteligentes para la administración y gestión del proceso logístico; el soporte inteligente a las primeras etapas del proceso de contratación de personal; y el monitoreo del desempeño laboral mediante software inteligente.

Actualmente se halla en fase de observación el grado de cubrimiento de competencias presentes en los programas de grado de las carreras en ciencias de la Universidad Siglo 21 tendientes a satisfacer la demanda observada. Todas las carreras del Decanato de Ciencias Aplicadas están siendo consideradas a través del relevamiento documental y la revisión bibliográfica de sus propuestas disciplinares. De forma paralela, expertos en educación en IA y en Didáctica de las Ciencias están siendo consultados para definir a partir de lo ya

realizado una primera aproximación al conjunto minimal de competencias que se busca establecer. En este sentido se están llevando a cabo entrevistas abiertas a profesores y autoridades de la casa de estudios mencionada. Además, anticipándose a próximos ciclos, la Ejecutiva Digital Mrs Penny Langford de MK College group. [7] y [8]

En cuanto a la formación de recursos humanos, se llevo a cabo en noviembre del año pasado una capacitación dentro del Master en Business de la Universidad Siglo 21; y se avanza en la participación activa de todos los miembros del equipo del proyecto que enmarca lo aquí presentado. Dicho equipo está integrado por dos investigadoras formadas docentes titulares disciplinares de la carrera de Licenciatura en Informática de la Universidad Siglo 21 y dos alumnos de grado avanzados. Se espera una reestructuración del equipo de trabajo en los próximos meses.

CONCLUSIONES Y TRABAJOS FUTUROS

Esta propuesta presenta a una línea de investigación orientada a establecer correlaciones entre las demandas del mundo laboral respecto a la incorporación de la IA como motor fundamental de las prácticas tendientes a lograr modelos empresariales de calidad total; y la capacitación en competencias y saberes provista por núcleos de formación de grado en carreras de ciencias aplicadas. Los objetivos incluyen la definición de un conjunto minimal de habilidades y competencias y la posterior observación de su grado de existencia e implementación en propuestas curriculares vigentes enmarcadas en lo antes mencionado. Se reporta el trabajo realizado enfatizándose que se trata de una primera fase de establecimiento del estado del arte, añadiéndose una parte del primer ciclo del proceso de investigación en marcha. Actualmente, entrevistas a expertos y docentes de esta casa de estudios se están llevando a cabo. Se espera concluir con el resto de las fases del primer ciclo para mediados del 2023. Como

⁵ <https://www.ibm.com/downloads/cas/GVAGA3JP>

trabajo futuro a corto plazo, los resultados obtenidos en el primer ciclo aun no cerrado deben ser validados. A mediano plazo, otros Programas Analíticos y otras Propuestas Curriculares de otros centros de formación en grado de ciencias aplicadas necesitan ser incorporados. Asimismo, otros bancos de referencias de empresas con alta penetración de estrategias de IA dentro de sus prácticas y procesos laborales deben considerarse. La formación en recursos humanos continuará afianzándose a través del trabajo colaborativo del equipo de investigación involucrado.

REFERENCIAS

Alfredo, M. A., Granovsky, P., & Verchelli, V. (2021). Cambio tecnológico, empleo y Formación Profesional: el papel de los actores sectoriales en Argentina.

Board, D. I. (2019). AI Principles: Recommendations on the ethical use of artificial intelligence by the Department of Defense. Supporting document, Defense Innovation Board, 2, 3.

Felder, F., Grayling, A., Steinberg, L. & Bentley, T. (2018) Future of Education and Skills 2030: Conceptual Learning Framework Ed del Proyecto OECD Disponible en <https://www.oecd.org/education/2030-project/> (último acceso marzo,2023).

García-Peñalvo, F. J. (2022). Developing robust state-of-the-art reports: Systematic Literature Reviews. Guile D., Spours K. &

Grainger P. (2018) Post-14 Centre for Education and Work. 4th Industrial Revolution, The Future of Work in association with the G20

Guitérrez-Portela, F., Hernández-Aros, L., Moreno-Hernández, J. J., & Ruiz-Correa, M. A. (2019). Análisis de la Incidencia que tiene en el desarrollo de un país la articulación de los procesos investigativos gestados al interior de las universidades con base tecnológica y las empresas intensivas en Tecnología. Revista Sinergia, (5), 117-131.

Langford, P (2022) Executive Head Digital, Milton College & Grainger, P., Honorary Senior Research Associate, UCL,UK

Maurya, M., & Shah, N. (2014). Labour market information system. International Journal of Application or Innovation in Engineering & Management (IJAIEEM), 3(3), 534-541.

Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., & Garro-Aburto, L. L. (2019). Inteligencia artificial y sus implicaciones en la educación superior. Propósitos y representaciones, 7(2), 536-568

Popkova, E. G., & Gulzat, K. (2020). Technological revolution in the 21 st century: digital society vs. artificial intelligence. In The 21st century from the positions of modern science: Intellectual, digital and innovative aspects (pp. 339-345). Springer International Publishing.

Unesco (2021) Inteligencia Artificial y Educación. Guía para las personas a cargo de formular políticas. Ed Unesco, ISBN 978-92-3-300165

Proyecto Esport UNLP

Diplomatura en Deportes Electrónicos

Armando De Giusti, Patricia Pesado, Cesar Estrebou, Luciano Iglesias, Santiago Medina

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 - La Plata, Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 - La Plata, Buenos Aires
(adegiusti, ppesado, cesarest, smedina) @lidi.info.unlp.edu.ar, li@info.unlp.edu.ar

RESUMEN

El Proyecto Esport que se desarrolla en el III-LIDI de la Facultad de Informática de la UNLP, propicia la conformación, entrenamiento y participación de equipos de la Universidad Nacional de La Plata en distintas disciplinas de los deportes electrónicos. Se trata de un proyecto multidisciplinario en el que participan actores con diferentes conocimientos de base. El Centro de Entrenamiento montado en el Centro de Innovación y Transferencia Tecnológica de la Facultad (CIYTT), permite tanto la práctica de los distintos equipos como la realización de ligas interfacultades e interuniversidades.

En el marco del proyecto, la Facultad definió y dictó los cursos para la primer cohorte de una Diplomatura en Deportes Electrónicos durante el segundo semestre de 2022. Su plan de estudios permite formalizar estudios que abarcan las diferentes áreas de los Esports, de modo de abrir un marco de posibilidades para alumnos universitarios y pre-universitarios interesados en el tema y que motiva el desarrollo de vocaciones en las disciplinas que conforman el ecosistema.

Palabras clave: deportes electrónicos, videojuegos, competencias, diplomatura.

CONTEXTO

Los “deportes electrónicos” son básicamente competiciones de videojuegos, que se realizan en forma individual o por equipos. Su crecimiento ha sido explosivo en los últimos años y ya se ha discutido su inclusión en los Juegos Olímpicos futuros.

En este contexto hay mucho interés de los jóvenes por participar en todas las actividades vinculadas con los Esports (desde el desarrollo de los videojuegos mismos hasta la gestión de torneos o la conformación y entrenamiento de equipos que compiten).

El desarrollo de la actividad (en Argentina y en el mundo) indica posibilidades laborales en diferentes áreas disciplinarias asociadas con los Esports.

En este contexto, la Universidad Nacional de La Plata ha generado un Proyecto (que se desarrolla en la Facultad de Informática, en el III-LIDI) desde 2021 fomentando la práctica de deportes electrónicos por sus alumnos, docentes y no docentes. En este Proyecto colaboran alumnos/especialistas de diferentes disciplinas (Informática, Ciencias Económicas, Comunicación, Psicología, etc).

Desde la Facultad de Informática se entiende que también es un elemento motivador para desarrollar estudios universitarios en diferentes disciplinas (en particular Informática) asociadas con los Deportes Electrónicos.

1. INTRODUCCIÓN

Desde hace unos años la UNLP cuenta con un espacio destinado a la práctica de deportes electrónicos. La iniciativa permite contar con planteles oficiales de la Universidad, para competir en torneos nacionales e internacionales de videojuegos que fomenten la participación de la comunidad universitaria.

En un breve recorrido por los esports, los primeros torneos profesionales de esports se pueden rastrear a partir de la década del 80, aunque recién en los últimos años cobraron gran popularidad, en parte gracias a las plataformas de streaming.

Desde un punto de vista deportivo, en 2018 se incluyeron a los esports como un deporte de demostración en los Asian Games, y en la edición del 2022 pasarán a ser deporte oficial. Incluso Japón pretendió que se los incluya en los Juegos Olímpicos 2021. Además, en Estados Unidos y China se ofrecen becas para universidades que cuentan con sus propias ligas nacionales universitarias.

Por otra parte, también se llevan a cabo campeonatos nacionales, como es el caso del torneo organizado por la Asociación de Fútbol Argentino (AFA), la Bundesliga, la Liga de España y la Premier League. Todas cuentan con sus versiones de torneos de esports, donde participan la mayoría de sus equipos.

El interés de la Facultad de Informática, desde el III-LIDI ha sido no sólo “utilizar” juegos y plataformas para videojuegos y competencias, sino también desarrollar capacidades para generar los mismos, disponer de una infraestructura propia para torneos y gestión de equipos así como investigar en temas asociados tales como Realidad Extendida o Inteligencia Artificial aplicada a las estrategias de juegos.

En este marco, la propuesta en la Universidad, surgió a partir de la inquietud de un equipo interdisciplinario integrado por investigadores de la Facultad de Informática y de otras Unidades Académicas de la UNLP, estudiantes y graduados de diferentes disciplinas y el apoyo del el Centro de Producción Multimedial de la Universidad (CEPROM). A partir de allí se desarrolló la estructura y se generó el marco institucional para que estudiantes, docentes, no docentes y egresados puedan participar de los denominados Esports.

En el Centro de Innovación y Transferencia Tecnológica de la Facultad de Informática se ha instalado el Centro de Entrenamiento en Esports, para los equipos de la UNLP. Allí se cuenta con una importante infraestructura que está operativa desde Septiembre 2022.

En ese contexto la Facultad de Informática implementó una Diplomatura en Deportes Electrónicos, que está destinada a quienes hayan finalizado la escuela secundaria y de

forma particular a estudiantes de la Universidad Nacional de La Plata, tomando como base la Ordenanza 290/16 de la UNLP referida a las Diplomaturas y la Ordenanza 334/22 de la Facultad de Informática de la UNLP “REGLAMENTO PARA LAS DIPLOMATURAS EN LA FACULTAD DE INFORMÁTICA”.

La Diplomatura adoptó la modalidad “a distancia con instancias presenciales”, tratándose de un programa que se pueda hacer en modalidad “a distancia sincrónica” combinado con actividades presenciales, de modo de utilizar el equipamiento y casos prácticos de Esports en el ClyTT de la Facultad de Informática. Para su desarrollo se utiliza la legislación y recursos de EAD de la Facultad de Informática. Las actividades se realizan en el entorno de enseñanza aprendizaje IDEAS de la Facultad, con tutores para atención de los alumnos y posibilidad de consultas presenciales atendidas por el equipo del proyecto Esports.

Esta formación busca que el diplomado tenga una visión del ecosistema de los deportes electrónicos, es decir, sobre los jugadores, equipos, organizadores, comunicadores, entrenadores, publicidad, entre otros aspectos. Además, que conozca los fundamentos tecnológicos de soporte para los Esports.

Se combinan contenidos tecnológicos de impacto en los Esports y que son conocidos y tratados en la Facultad y en particular en el III-LIDI con contenidos relacionados con la gestión de proyectos de Esports (Gestión/ Psicología/ Marketing/ Comunicación/ Dirección de proyectos de Esports) los cuales caracterizan varias de las carreras existentes en el país.

La iniciativa contempla que se realicen actividades concretas en el ámbito del proyecto de Esports que tiene la UNLP.

2. LÍNEAS DE INVESTIGACIÓN / DESARROLLO

Las líneas de investigación y desarrollo del proyecto están relacionadas con los espacios curriculares que se desarrollan en la Diplomatura de Deportes Electrónicos y los temas de las PPS.

Plan de Estudios de la Diplomatura en Deportes Electrónicos

Espacios Curriculares	
Introducción. Aspectos generales de los Esports	20hs.
Tecnología asociada con Esports	18hs.
Proyectos de Esports	18hs.
Dirección de Proyectos de Esports	20hs.
Aspectos tecnológicos en Proyectos de Esports	20hs.
Informática y Esports	20hs.
Aspectos de Psicología y Comunicación	18hs.
Aspectos Complementarios de Informática y Tecnología	19hs.
Práctica Profesional Supervisada	80hs.
TOTAL CARGA HORARIA	232hs.

1- Introducción. Aspectos generales de los Esports.

- Evolución y estructura deportiva del videojuego competitivo.
- Categorización genérica de los videojuegos competitivos.
- Contexto Socioeconómico de la industria esports.
- Eventos. Estructura deportiva. Tipos de Competiciones. Plataformas y Formatos.
- Organización de un evento de Esport. Roles.
- Panorama general de la industria esports. Ecosistema actual y funcionamiento.
- Estructura organizacional. Recursos humanos. Gestión de talentos. Retos presentes y futuros del sector.
- Modelos de gestión y negocio. Aspectos legales y marco regulatorio.

2- Tecnología asociada con Esports

- Arquitectura física de soporte para Esports. Servidores.
- Servidores locales y en la nube.
- Comunicaciones.
- Streaming. Tecnología de transmisión de eventos de Esport.

- Escalabilidad con el número de usuarios/jugadores y la resolución en tiempo real.
- Equipamiento especial para Esports.
- Evolución de la tecnología de soporte desde los videojuegos a los Esports.

3- Proyectos de Esports.

- Diseño de productos y gestión de proyectos.
- Diferencia entre producto y proyecto.
- Definición/especificación de un proyecto. Definición de producto. Organizando el proyecto.
- Costos y presupuesto. Contenidos e innovación.
- Importancia del B2B en esports. Tipos de acciones B2B que se realizan en esports: eventos, competiciones, marca blanca. Agentes del sector importantes.
- Contabilidad y análisis financiero. Modelización financiera y valorización.
- Capital riesgo y gestión emprendedora.

4- Dirección de Proyectos de Esports

- Definición e identificación del proyecto de negocio.
- Modelo y actividad del proyecto.
- Plan de Marketing. Plan de operaciones.
- El branding y la construcción de marca. Identidad visual. "Vender" el producto.
- La estrategia de marketing. Objetivos y procesos de elaboración. Ejecución táctica de la estrategia. Campañas y patrocinios.
- Plan de RRHH. Plan de inversiones y económico.
- Estructura legal. Calendario y evaluación del proyecto.
- Vinculación universitaria.

5- Aspectos tecnológicos en Proyectos de Esports

- Realidad Virtual.
- Realidad Aumentada.
- Realidad Extendida.
- Ambientes inmersivos.
- Visualización 2D y 3D. Aplicaciones a videojuegos.
- Bases de datos orientadas a Esports. Gestión de los datos.
- Seguridad de los Datos.

6- Informática y Esports

- Conceptos básicos de programación de videojuegos.
- Videojuegos y Juegos Serios.
- Juegos serios y videojuegos enfocados a Educación.
- Casos de estudio.

7- Aspectos de Psicología y Comunicación.

- Psicología organizacional.
- Gestión de jugadores. Roles. Propuesta de valor y comunicación interna. Cultura organizacional.
- Entrenamientos y aspectos deportivos. Metas y objetivos. Fortalezas. Debilidades.
- Amenazas. Oportunidades.
- Psicología deportiva: Motivación y manejo de emociones. Preparación mental. Ansiedad precompetitiva. Expectativa. Frustración. Rutinas de pensamiento y relajación.
- Rendimiento. Herramientas y dispositivos de concentración y comunicación efectiva. Visualización. Liderazgo. Compromiso.
- Comunicación. Comunicación formal. Plan de Comunicación. Aspectos de periodismo.
- Proyección y tipos de alcance.
- Comunicación informal. Gestión de contenidos y comunidades. Herramientas y canales de distribución de contenidos. Twitch. Redes Sociales y el Plan de Redes Sociales.

8- Aspectos Complementarios de Informática y Tecnología

- Conceptos de Big Data. Aplicaciones en Esports.
- Análisis de datos de Esports.
- Estudio de casos.
- Conceptos de Inteligencia Artificial y sus aplicaciones en Esports.
- Inteligencia Artificial y estrategias de juego.
- Casos de estudio.
- Análisis inteligente de datos "post-juego". Y "post torneo".

9- Práctica Profesional Supervisada

La lista de temas propuestos desde la coordinación de la diplomatura para realizar la PPS han sido:

- Análisis informático de videojuegos.
- Análisis de Juegos Educativos.
- Aplicaciones móviles orientadas a juegos educativos.
- Estrategias inteligentes orientadas a deportes electrónicos.
- Servidores y comunicaciones en sistemas de deportes electrónicos.
- Limitaciones de la tecnología en Esports.
- Análisis de la estructura deportiva de un equipo competitivo.
- Análisis comparativo de juegos para Esports.
- Planificación de Torneos (Financiera / Marketing).
- Herramientas de gestión organizacional.
- Estrategias de comunicación externa.
- Análisis del desarrollo de Esports en Argentina y el mundo.

La aprobación de la PPS contará con una breve síntesis escrita y un coloquio técnico.

Las PPS pueden ser realizadas en forma individual o en grupo de hasta 3 alumnos, para realizar un estudio complementario / estudio de tecnologías / análisis comparativo de alternativas / investigación y/o desarrollo dentro del proyecto de Esports UNLP, etc.

La idea principal es vincular al alumno de la Diplomatura con el ámbito de desarrollo del proyecto que es un modo de integrarlos a la realidad de los Esports, tanto desde el punto de vista de las competencias como de la tecnología o la gestión de un proyecto de Esport.

3. RESULTADOS ESPERADOS/OBTENIDOS

Los objetivos para 2023 son:

- Consolidar a la UNLP como una referencia universitaria en Deportes Electrónicos.
- Capacitar en los conocimientos básicos de Deportes Electrónicos.
- Organizar actividades en la UNLP para incrementar la participación en el Proyecto Esports.
- Participar como UNLP en torneos de Deportes Electrónicos.

➤ Equipos UNLP. Centro de Entrenamiento.

En 2022 se ha puesto en marcha el Centro de Entrenamiento para los equipos que representan a la UNLP. La idea es fortalecer el funcionamiento del Centro (en el CiyTT Facultad de Informática) y alentar la participación de alumnos de las Facultades de la UNLP en los equipos. Actualmente se preparan equipos en juegos de estrategia (League of Legends, Age of Empires) y de disparos (Counter Strike, Valorant). Al mismo tiempo se ha presentado el primer equipo femenino UNLP en una disciplina de Deportes Electrónicos.

➤ Diplomatura en Deportes Electrónicos

En la actualidad está finalizando la primer cohorte de la Diplomatura en Deportes Electrónicos, donde los alumnos (que cursan carreras en distintas facultades de la UNLP) han logrado:

- Tener una visión clara del ecosistema de deportes electrónicos: jugadores, equipos, organizadores, comunicadores, entrenadores, publicidad, etc.

- Conocer los fundamentos tecnológicos de soporte para la actividad de Esports, en particular la tecnología para el desarrollo de videojuegos y su gestión y perfeccionamiento.
- Incorporar herramientas para desenvolverse en el entorno de los deportes electrónicos de diferentes áreas del conocimiento aplicadas a los Esports.
- Realizar una práctica concreta en el ámbito del proyecto de Esports de la UNLP, en los temas de su interés.

➤ **Participación de la UNLP en torneos.**

Es un eje importante la visibilidad de la UNLP como participante en torneos de diferentes disciplinas en el país (tanto virtuales como presenciales), seleccionando los torneos y perfeccionando las disciplinas en las que se entrenen equipos competitivos en el Centro de Entrenamiento.

➤ **Organización de torneos internos a la UNLP**

La participación de alumnos / docentes y no docentes de la UNLP en torneos específicos organizados desde el proyecto es importante para consolidar el tema en todas las Unidades Académicas y eventualmente en los Colegios de la UNLP, con un calendario anual y propiciando la participación de (en lo posible) al menos UN equipo por Facultad.

Todo esto sobre el soporte WEB y de gestión de torneos propio, desarrollado en la Facultad de Informática.

➤ **Investigación y Desarrollos asociados con los deportes electrónicos.**

Independientemente de otras actividades de difusión/información/competencias, interesa potenciar el crecimiento del tema de los deportes electrónicos desde la Investigación y Desarrollo en videojuegos / soporte de infraestructura / soporte de software/ Inteligencia artificial aplicada a los videojuegos / juegos educativos, etc, desde el Instituto.

➤ **Difusión en Paneles/ Charlas/ Conferencias/ Jornadas.**

y mantenimiento del sitio web del proyecto, la Diplomatura y la Organización de torneos.

4. FORMACION DE RECURSOS HUMANOS

Todos los resultados obtenidos/esperados del proyecto, se enmarcan en actividades de investigación, desarrollo y docencia realizadas desde el LIDI.

Podemos citar el desarrollo de un sitio web para la organización de torneos que se está realizando en el marco de una tesina de licenciatura o proyectos de innovación con alumnos desarrollados en la temática de los esports.

Respecto de la diplomatura, esta primera cohorte contará con unos 40 egresados, que han formalizado estudios que abarcan las diferentes áreas de los Esports, de modo de abrir un marco de posibilidades para alumnos interesados en el tema y que decidan profundizar sus estudios en alguno de los aspectos relacionados al mundo de los deportes electrónicos.

5. BIBLIOGRAFÍA

- Ordenanza 334/22 de la Facultad de Informática de la UNLP "REGLAMENTO PARA LAS DIPLOMATURAS EN LA FACULTAD DE INFORMÁTICA".
- Antón, M. (2018). "Los deportes electrónicos: el espectáculo en las competiciones de videojuegos" [Tesis de Doctorado, Universidad Complutense de Madrid]. E-Prints Complutense. <https://eprints.ucm.es/id/eprint/54986/1/T41009.pdf>
- AEVI - Spanish Association of Videogames. (2020). "Los esports en España: situación actual y posición de la industria. Asociación Española de Videojuegos." http://www.aevi.org.es/web/wp-content/uploads/2020/12/Informe_esports_ES_P_20.pdf
- Carrillo, J. A. (2015). "La dimensión social de los videojuegos' online': de las comunidades de jugadores a los 'esports'." Index. Comunicación: Revista científica en el ámbito de la Comunicación Aplicada, 5(1), 39-51.
- Taylor, T. L. (2012). "Raising the Stakes: E-sports and the Professionalization of Computer Gaming." Cambridge: MA: MIT Press
- Gutierrez Gonzalez Ángel. "Tecnologías de la Información - Un enfoque interdisciplinario" Alfaomega Grupo Editor Argentino
- Austerb David. "La Tecnología del Streaming de Vídeo y Audio" Editorial Escuela de Cine y Vídeo de Andoaín, S.L.

- Angel Arias. "Computación en la Nube" Editorial: CreateSpace Independent Publishing Platform; 2nd edición
- Guillén Muñoz, Á. (2020). "La planificación estratégica en equipos de esports". Universidad de Valladolid. <https://uvadoc.uva.es/bitstream/handle/10324/43529/TFG-N.%201476.pdf>
- Pérez-Sagarzazu, G. (2019). "Introducción de los deportes electrónicos en la Universidad y diseño de su estructura organizativa" UNAV Esports.
- Nielsen Esports (2022). "The ROI of Esports." The Nielsen Company.
- Scholz & Barlow. (2019). "Esports is Business." Springer International Publishing.
- Cáliz Ortiz, J. S., & Paredes Solano, A. D. (2020). "Cómo facilitar el proceso de reclutamiento de aspirantes latinoamericanos a equipos de esports." Biblioteca Digital Icesi.
- Cristófol, F., Martínez-Ruiz, A., Román, I. & Rodríguez, C. (2020). "Evolución de las estrategias de patrocinio en los esports en España: 2013-2021." Ámbitos: Revista Internacional de Comunicación 48, pp. 188-204. Es
- Fanjul, C., González, C., & Peña, P. J. (2019). "La influencia de los jugadores de videojuegos online en las estrategias publicitarias de las marcas: comparativa entre paña y Corea." Comunicar, 58.
- Nallar, D. (2016). "Diseño de juegos en América Latina II: Diseño y narrativa transmedia."
- Observatorio de Industria de Videojuegos (2021). Resumen Anual 2021. UNRaf. https://www.unraf.edu.ar/images/INVESTIGACION/OBSERVATORIO_VideoJuegos/PDFInformes/Observatorio_Videojuegos_UNRaf_2021.pdf
- "Aprende a crear tus propios videojuegos con Unity 3D Volumen 1: Introducción práctica a Unity 3D" (Spanish Edition) Edición Kindle. Editorial TRINIT Asoc. Informáticos de Zaragoza. Año 2018.
- Luis Ruelas. "Desarrollo de Videojuegos y Realidad Virtual con Unity 3D." Editorial Ra-Ma. ISBN 978-84-18551-90-1
- Jose M. Martinez, Fernando Navarro, Antonio Martinez · 2018. "Realidad Virtual y Realidad Aumentada." Editorial Ra-Ma. ISBN13 9788499647395
- Cabero Almenara, J. y Jiménez García, F. "Realidad Aumentada. Tecnología para la Formación." Editorial Síntesis. 2016. ISBN: 978-84-9077-258-4
- Documentación en Videojuegos: Documento de diseño (GDD) en <https://eldocumentalistaudiovisual.com/2015/02/06/documentacion-en-videojuegos->
- documento-dedisenogdd/ RA-MA S.A. Editorial y Publicaciones
- David Ballejo Fernández , Cleto Martín Angelina. "Desarrollo de Videojuegos 1: Arquitectura del Motor de Videojuegos" "Desarrollo de Videojuegos 2: Programación Gráfica" "Desarrollo de Videojuegos 3: Técnicas Avanzadas" "Desarrollo de Videojuegos 4: Desarrollo de Componentes". Editados por la Universidad de Castilla La Mancha
- Ana Belén Gomez Sanz . "Gamificación y Juegos Serios. Curso Práctico" RA-MA S.A. Editorial y Publicaciones
- Ricardo Casañ Pitarch. "Videojuegos Serios en educación" Editor: Universitat Politècnica de València
- Antón, M; García, F (2014) "Deportes electrónicos. Una aproximación a las posibilidades comunicativas de un mercado emergente." Questiones publicitarias, 1, pp.98-115
- Mendoza, G. (2019). "El trabajo psicológico online con equipos de esports: experiencias en Vodafone Giants". Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 4(1), 1-7.
- Pedraza-Ramirez, I. (2019). "Generación LoL: entrenamiento psicológico mediante una propuesta holística con un equipo profesional de Esports." Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 1-10.
- Vaamonde, A. G. N., Toribio, M. J., Molero, B. T., & Suárez, A. (2018). "Beneficios cognitivos, psicológicos y personales del uso de los videojuegos y esports: una revisión" Revista de Psicología Aplicada al Deporte y al Ejercicio Físico, 3(2), 1-14.
- Pablo L. Gutierrez Utande – Angel Arroyo Castillo (Tutor) "Aplicación de Inteligencia Artificial en Videojuegos" Universidad Politécnica de Madrid – (2017)
- Bernard Marr - "BIG DATA: La utilización del Big Data, el análisis y los parámetros SMART para tomar mejores decisiones y aumentar el rendimiento" Editorial TEELL - 2016
- Eric Sadin - "La Inteligencia Artificial o el Desafío del siglo" Colección Futuros Próximos – 2020

Programación y Autismo: una arquitectura epistémica centrada en Ciencia Ciudadana

Federico Amigone, Jorge Rodríguez
{fede.amigone,j.rodrig}@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

La presente Línea de Investigación y Desarrollo, propone el estudio y el modelado de prácticas basadas en evidencia que permitan, desde el ámbito de la ciencia ciudadana, producir aprendizaje en programación para la población de infancias, adolescencias y juventudes en el espectro autista. Para ello, se propone como eje central un dispositivo epistemológico con formato escuela de programación para estudiantes autistas, denominada Escuela PASAYO.

La Escuela PASAYO es un entorno de diseño crítico, prueba y facilitación de experiencias de aprendizaje en programación para jóvenes en el espectro autista y sus familias.

Este dispositivo permite a familias residentes en diferentes puntos del país experimentar una metodología experimental de aprendizaje en programación basada en la transposición de narrativas y evaluar sus resultados. Estos tránsitos formativos producen datos epistemológicos que retroalimentan el diseño de las prácticas propuestas.

Durante el año 2022 se trabajó con 17 familias durante dos cuatrimestres, obteniendo aproximadamente 200 artefactos computacionales ejecutables bajo diferentes modalidades de programación: tangible, bloques y texto.

Palabras Clave: programación, autismo, ciencia ciudadana

Contexto

Esta línea de investigación y desarrollo se enmarca, por un lado, en el contexto de los temas de interés promovidos por el Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA), de la Facultad de Informática de la Universidad Nacional del Comahue (UNCo).

En particular, se enmarca en el ámbito del proyecto de investigación de la Facultad de Informática denominado *Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación*, presentado en la convocatoria 2022.

1. Introducción

La irrupción de la tecnología en la constitución de los procesos sociales, políticos, productivos y económicos plantea profundos desafíos y oportunidades para las juventudes en el espectro autista. Por un lado, comprender cómo funcionan las computadoras y el software ya no es opcional, es requisito para gozar de una ciudadanía plena en el siglo XXI. Sin una mínima capacidad de comprensión sobre las tecnologías y las Ciencias de la Computación, las personas en el espectro son relegadas a los márgenes de la actividad social, quedando como espectadores de un suceder que escapa a su comprensión e intervención. Se obtura así su derecho a la participación de los debates públicos; son situados forzosamente a una condición pasiva en la cual no se encuentran habilitados a opinar,

quedando así relegados a las opiniones de otros. Por otro lado y más allá de la comprensión ciudadana del mundo, el conocimiento en programación es una plataforma sobre la que pueden construirse, impulsarse y defenderse derechos de acceso al trabajo para las y los jóvenes en el espectro autista.

Existen experiencias aisladas en otros países, tanto en el ámbito de la formación en conceptos relacionados a la programación como iniciativas que buscan ampliar la participación de las personas en el campo laboral tecnológico [6][8]. Estos estudios e iniciativas sugieren que existe afinidad fuerte entre el autismo y la programación y que esto resulta satisfactorio al menos en tres dimensiones: mejora las posibilidades de inserción laboral, contribuye a hacer efectivo el derecho a conocer que le asiste a todas las personas y mejora los procesos de socialización y empoderamiento [9][10].

No obstante, actualmente, el acceso al conocimiento tecnológico se encuentra fuera del alcance de las infancias, adolescencias y juventudes en el espectro autista [7]. Las instituciones educativas, de todos los niveles, que ofrecen formación en este campo no cuentan con dispositivos y enfoques didácticos específicos para las personas en el espectro. Y por otro lado, los diseños curriculares para estos trayectos formativos suelen estar sujetos a rigideces normativas y conservadoras que si bien pueden funcionar para personas situadas en la media funcional, habitualmente no se acomodan a jóvenes en el espectro autista [1][2].

Desde la perspectiva epistemológica, la programación y el autismo configura un espacio de amplia vacancia. No existen suficientes iniciativas de investigación que hayan logrado consolidar conclusiones teóricas y prácticas basadas en evidencia, ni tampoco la demarcación de los contornos para

la definición de una didáctica específica posible [7]. Las iniciativas en investigación que tengan por objeto de estudio y desarrollo a los dispositivos, mecanismos y didácticas específicas para producir aprendizaje en programación son, por tanto, tan necesarias como escasas. Los dispositivos no solo deben diseñarse, sino que deben construirse y experimentarse en el campo de estudio. La construcción de plataformas didácticas, costosa en recursos económicos, humanos y temporales, requiere de aproximaciones de especialistas y del contacto efectivo con la población autista interesada.

Este trabajo se enmarca dentro de la Línea de Investigación y Desarrollo destinada a producir recursos didácticos para enseñar Ciencias de la Computación y evaluar su efectividad en el ámbito de la educación.

El resto del artículo está organizado de la siguiente manera. A continuación, describimos la línea de investigación y desarrollo. En la sección 3, presentamos los resultados obtenidos y esperados. Finalmente, en la sección 4, comentamos aspectos relacionados a la formación de recursos humanos en el marco de esta línea, así como del proyecto de investigación en su conjunto.

2. Línea de Investigación y Desarrollo

La Escuela PASAYO define su demarcación epistemológica en la investigación de enfoques didácticos específicos y el estudio de dispositivos pedagógicos sobre los cuales evaluar el desarrollo de habilidades de programación en las infancias, adolescencias y juventudes en el espectro autista.

El campo de despliegue son familias con integrantes, niños, niñas, adolescentes o jóvenes en el espectro que cuenten facultades motrices para la operatividad con una

computadora, algún grado de interés por la tecnología y la hipótesis de un potencial interés futuro en las actividades relativas a la programación.

El modelo de investigación está basado en la evaluación empírica de la efectividad para producir habilidades en programación a partir de tres ejes:

- **Escuela PASAYO**

La escuela constituye un dispositivo epistémico de campo sobre el cual se nuclea familias que participan en el proceso de investigación. Al incorporar familias a la escuela, se registra el perfil neurológico de la/el joven en el espectro y se exploran las oportunidades y fortalezas de el/la estudiante en relación a la disciplina de la programación. Se le asigna un facilitador PASAYO al núcleo familiar y se comienzan a registrar los resultados experimentales del proceso de investigación.

Por lo tanto, la Escuela constituye un dispositivo de investigación que nutre de información recuperada de la experiencia real a los procesos de definición de enfoques metodológicos destinados a orientar la enseñanza de la programación a personas con autismo.

- **Plataforma PASAYO**

La plataforma PASAYO es un entorno web experimental para la investigación ciudadana y el desarrollo colaborativo de experiencias de programación en el espectro autista. Posee tres modos de programación:

Modo Tangible: Mediante la plataforma PASAYO TANGIBLE, la programación se realiza utilizando recursos concretos y físicos [3]. Los elementos del lenguaje de programación toman la forma de flechas direccionales que se ubican sobre un área

de programación demarcada. En otro sector, un avatar personalizable ejecuta el código tangible, producido por la persona en el espectro, desplazándose sobre un escenario denominado micromundo.

Modo Bloques: La plataforma PASAYO Bloques permite un acceso de reducida complejidad a la programación con computadoras [4]. Con la plataforma PASAYO Bloques, puede realizarse una transición suave entre la programación tangible y la programación en computadora. Los bloques se arrastran y apilan de la misma forma que se hacía con los elementos del lenguaje en el modo tangible, y este código produce un efecto sobre el avatar digital en su micromundo. En el modo Bloques es además un entorno colaborativo para producir y compartir desafíos y experiencias de programación con bloques.

Modo Texto: PASAYO Texto es el entorno ofrecido por la plataforma PASAYO para la programación bajo el modo textual [5]. Es el nivel más complejo y potente de la propuesta. Se programa utilizando lenguajes industriales como Python y Javascript. Desde PASAYO Texto es posible crear y compartir desafíos así como programarlos en un entorno colaborativo en tiempo real.

- **Metodología PASAYO**

El esquema epistemológico está inspirado en el concepto de ciencia ciudadana, es decir, el trabajo científico realizado voluntariamente por miembros de la sociedad en general, a menudo en colaboración con investigadores profesionales e instituciones científicas [11].

En la dimensión didáctica, la Escuela PASAYO trabaja con un modelo de facilitación experimental para el diseño de intervenciones didácticas. Esta

facilitación se organiza bajo un esquema de ternas compuestas por las figuras del estudiante autista, el facilitador familiar y el facilitador PASAYO. Los resultados de cada experiencia son volcados en una base de datos epistémica sobre la cual es posible realizar conclusiones de la efectividad del modelo propuesto, evaluar el impacto en el proceso de aprendizaje y construir métricas de desempeño del proceso formativo.

3. Resultados obtenidos y esperados

En el contexto de las actividades desarrolladas en esta línea se logró desplegar los primeros experimentos de aplicación de enfoques y plataformas en la enseñanza de la programación a personas en el espectro autista. Estos experimentos de campo incorporan las modalidades de programación tangible, bloques y texto a lo largo de un trayecto con la participación de un niño en el espectro. En 2021, inicia el desarrollo de plataforma PASAYO que incorpora entornos propios de programación tangible, bloque y texto en lenguaje javascript. También en 2022, se amplía el equipo de trabajo integrando a docentes investigadoras con experiencia en la facilitación de grupos y ampliación de derechos, se conforma formalmente la Comunidad de Ciencia Ciudadana para la Computación - C4. Al momento, y luego de un año de trabajo con 17 familias, el enfoque propuesto verifica la producción de artefactos computacionales en distintos modos de programación, tangible, bloques y texto. Esta producción aporta indicios alentadores acerca de que el enfoque habilita la construcción de las estructuras cognitivas que permiten el dominio progresivo de la algoritmia en estudiantes dentro del espectro autista. En general, a la mayoría de las familias les tomó aproximadamente un mes con dos estímulos semanales, realizar las propuestas de

programación tangible y esto las ubicó en un estadio cognitivo satisfactorio para comenzar a programar en el paradigma de bloques.

En 2022 se concretan acuerdos de colaboración con cuatro empresas tecnológicas que aportan recursos económicos para sostener la incorporación de becarios investigadores, lo cual permitió ampliar el equipo de trabajo PASAYO. La Escuela PASAYO inicia las actividades formales en mayo de 2022.

Los resultados sugieren que el modelo de trabajo propuesto resulta efectivo. En el plano didáctico específico, no se observaron dificultades significativas en el abordaje inicial a las complejidades algorítmicas bajo el modo tangible. Se verificó que la flexibilidad de la propuesta funcionó bien singularizando las actividades a la agenda propia de los y las estudiantes. Se produjo conocimiento computacional en niñas, niños y adolescentes que no tenían ningún conocimiento algoritmo previo.

Respecto del modelo de facilitación, las ternas demostraron ser una arquitectura posible y potente. En los casos donde la familia logró regularidad en los espacios de facilitación, se observaron progresos incrementales en las soluciones construidas por los y las estudiantes. El pasaje del paradigma tangible al paradigma bloques resultó ser apropiado, habilitando la producción de soluciones sin mayores complicaciones habida cuenta de que los conceptos basales de instrucción, algoritmo y lenguaje fueron bien modelados bajo el paradigma tangible.

Por otra parte, el enfoque de diseño participativo parece una respuesta satisfactoria al problema de producción de recursos situados. La plataforma PASAYO resulta ser un medio valioso para la producción de recursos y su socialización.

En esta sección se presentan los resultados obtenidos y esperados que se ubica como una continuidad de la Línea de Investigación destinada a estudiar aspectos curriculares relacionados a la

4. Formación de Recursos Humanos

Se espera que el desarrollo de esta Línea de Investigación contribuya a la formación de recursos humanos en el campo de las Ciencias de la Computación en la Educación.

En este sentido, los integrantes de esta Línea de Investigación están finalizando maestrías en campos relacionados a las Ciencias de la Computación y Enseñanza en Escenarios Digitales.

5. Referencias

[1] Gribble, J., Hansen, A., Harlow, D., y Franklin, D. (2017). Cracking the code: the impact of computer coding on the interactions of a child with autism. En *Proceedings of the 2017 conference on interaction design and children* (pp. 445–450).

[2] Wille, S., Century, J., y Pike, M. (2017). Exploratory research to expand opportunities in computer science for students with learning differences. *Computing in Science & Engineering*, 19 (3), 40–50.

[3] Sapounidis, T., Stamelos, I., y Demetriadis, S. (2016). Tangible user interfaces for programming and education: A new field for innovation and entrepreneurship. En *Innovation and entrepreneurship in education* (Vol. 2, pp.271–295). Emerald Group Publishing Limited.

[4] Weintrop, D. (2019). Block-based programming in computer science education. *Communications of the ACM*, 62 (8), 22–25.

[5] Weintrop, D., y Wilensky, U. (2017). Comparing block-based and text-based programming in high school computer science

classrooms. *ACM Transactions on Computing Education (TOCE)*, 18 (1), 1–25.

[6] Krzeminska, A., Austin, R. D., Bruyère, S. M., y Hedley, D. (2019). The advantages and challenges of neurodiversity employment in organizations. *Journal of Management and Organization*, 25 (4), 453–463. doi: 10.1017/jmo.2019.58

[7] Munoz, R., Villarroel, R., Barcelos, T. S., Riquelme, F., Quezada, A., y Bustos-Valenzuela, P. (2018). Developing computational thinking skills in adolescents with autism spectrum disorder through digital game programming. *IEEE Access*, 6, 63880–63889.

[8] Zubair, M. S., Brown, D. J., Hughes-Roberts, T., y Bates, M. (2021). Designing accessible visual programming tools for children with autism spectrum condition. *Universal Access in the Information Society*, 1–20.

[9] Grandin, T. (1999). Choosing the right job for people with autism or asperger's syndrome. Indiana Resources Center for Autism.

[10] Eiselt, K., y Carter, P. (2018). Integrating social skills practice with computer programming for students on the autism spectrum. En *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1–5).

[11] Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., . . . Wagenknecht, K. (2021). *The science of citizen science*. Springer Nature.

Desarrollo de Recursos Educativos para la Enseñanza de las Ciencias de la Computación en la Escuela Secundaria.

Daniel Dolz Gerardo Parra Jorge Rodríguez Susana Parra
{ddolz, gparra,j.rodrig,susana.parra}@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
Departamento de Computación Aplicada - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

La presencia de contenidos sobre Ciencias de la Computación en el currículum escolar está aumentando rápidamente. Las discusiones curriculares vigentes proponen con mayor frecuencia desarrollar un recorrido amplio por las áreas de conocimiento.

En este contexto, los recursos desenchufados, sobre todo los desarrollados en el ámbito de la iniciativa *CS Unplugged*, se ubican como una pieza fundamental para las propuestas que buscan impulsar la inclusión de conocimiento computacional a la educación obligatoria.

Esta propuesta se enmarca dentro de la Línea destinada a desarrollar recursos didácticos y evaluar su efectividad. En particular se trabaja sobre la construcción y evaluación de una colección de recursos orientados a facilitar la enseñanza de las Ciencias de la Computación.

Palabras Clave: Educación en Ciencias de la Computación, Escuela Secundaria, Recursos Educativos Desenchufados, Participatory Design.

Contexto

El Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA), de la Facultad de Informática de la Universidad Nacional del Comahue (UNCo) promueve el estudio de diversos temas de interés. Esta línea de investigación y desarrollo se enmarca en el contexto de este grupo.

En particular, la línea de investigación se desarrolla en el ámbito de los proyectos de investigación de la Facultad de Informática denominados *Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación (04/F022)* y el proyecto *Computación Aplicada a las Ciencias y al Medio Dirección (04/F021)*, iniciados en Enero de 2022 y con finalización prevista para Diciembre de 2025.

1. Introducción

En los últimos años, en todo el mundo se está volviendo cada vez más evidente la presencia de contenidos de Ciencias de la Computación en el currículum escolar. Las iniciativas tienen como objetivo que la población estudiantil de todos los niveles educativos, tengan acceso a los conceptos centrales de la disciplina [8,9,10].

Las tendencias curriculares actuales promueven desarrollar un recorrido amplio por las distintas áreas de conocimiento. Si bien en la República Argentina se observa un proceso dispar e iniciativas con cierta preponderancia del área de Algoritmos y Programación, se nota que progresivamente se tiende a ampliar el recorrido.

En esta dirección, la provincia de Neuquén aprobó recientemente un nuevo diseño curricular para la escuela secundaria. Este diseño, que entra en ejecución en el presente año, incorpora, entre otros, contenidos de

Teoría de la Computación, Representación del Conocimiento y Redes de Computadoras, entre otros nudos disciplinares [11].

Estos contenidos usualmente son considerados abstractos y con cierta complejidad para estudiantes sin formación previa en la disciplina. Los recursos desenchufados, sobre todo los desarrollados en el ámbito de la iniciativa *CS Unplugged*, están ampliamente difundidos y suelen ofrecer un primer contacto satisfactorio [1, 2, 3].

Aunque inicialmente fueron pensados para actividades no escolarizadas, sobre todo para eventos de divulgación científica, las escuelas los adoptan como forma de ofrecer los primeros contactos con conceptos abstractos sobre computación.

Si bien en el marco de esta Línea de Investigación se realizaron estudios previos en relación a su efectividad en el ámbito escolar y acerca de la forma que deben adoptar para adaptarse a contextos institucionalizados que aportan indicios alentadores, es necesario realizar estudios a mayor escala [13, 14, 15].

La alta disponibilidad de recursos educativos desenchufados ofrece un recorrido amplio por las áreas de conocimiento. No obstante, la necesidad de producir nuevos recursos se mantiene vigente ya que existen tópicos con cobertura mejorable. Este es el caso de conceptos relacionados al área de Representación del Conocimiento.

Este trabajo se enmarca dentro de la Línea de Investigación y Desarrollo destinada a producir recursos didácticos para enseñar Ciencias de la Computación y evaluar su efectividad en el ámbito de la educación secundaria.

En particular, plantea trabajar sobre el desarrollo y evaluación de recursos educativos desenchufados orientados a facilitar la enseñanza de conceptos relacionados a la Representación del Conocimiento, la refactorización de la colección recursos

elaborados previamente y el estudio a mayor escala acerca de la efectividad de este tipo de recursos para contribuir a la mejora de la enseñanza en el contexto de la educación secundaria.

En este contexto se trabaja en el marco de los enfoques metodológicos basados en la investigación y el diseño participativos definidos específicamente por esta Línea de Investigación y Desarrollo [12].

El resto del artículo está organizado de la siguiente manera. A continuación, describimos la línea de investigación y desarrollo. En la sección 3, presentamos los resultados obtenidos y esperados. Finalmente, en la sección 4, comentamos aspectos relacionados a la formación de recursos humanos en el marco de esta línea, así como del proyecto de investigación en su conjunto.

2. Línea de investigación y desarrollo

El trabajo presentado en este artículo se enmarca en la Línea de Investigación y Desarrollo denominada *Recursos Educativos Desenchufados para la Enseñanza de las Ciencias de la Computación en la Escuela Secundaria*.

El trabajo a desarrollar en el contexto de esta línea se organiza y estructura a partir de los siguientes ejes:

- **Formalizar la definición de aspectos metodológicos para el diseño participativo.**

En esta línea se formuló un marco metodológico para apoyar los procesos de diseño participativo de recursos educativos desenchufados destinados a introducir conceptos básicos e intermedios de Ciencias de la Computación.

El marco metodológico está basado en el modelo Participatory Design Framing,

donde los docentes de las escuelas se involucran activamente en el proceso de elaboración de recursos educativos.

Sobre este modelo se desplegaron acciones para el diseño de recursos educativos en las áreas de Redes de Computadoras y de Teoría de la Computación.

Se busca trabajar en mejorar el grado de formalización del enfoque conceptual para avanzar en la definición de un marco de trabajo adecuado, modular y convenientemente flexible.

- **Refinar y refactorizar recursos educativos.**

Se está construyendo una colección de recursos educativos destinados a mejorar el aprendizaje de conceptos abstractos inherentes a ciertas áreas de Ciencias de la Computación.

Este desarrollo se basa, principalmente, en la iniciativa CS Unplugged y en el aprendizaje experiencial. Están diseñados para ofrecer un primer contacto con conocimiento disciplinar a estudiantes de cualquier año de estudio de la escuela secundaria sin conocimiento previo en las Ciencias de la Computación.

En general adoptan estrategias definidas en el ámbito del aprendizaje colaborativo, organizando la actividad en pequeños grupos. Se trata de actividades breves que incorporan el ciclo del aprendizaje experiencial, el juego y la distribución de la complejidad [14, 15].

Se busca refinar y refactorizar sobre la base las opiniones de docentes y estudiantes. También a partir de los resultados obtenidos en experimentos realizados en el marco de esta línea.

- **Estudiar la efectividad de la tipología particular de recursos.**

Se desplegaron trabajos de campo, con grupos de 20 a 60 estudiantes, en escuelas secundarias. Las experiencias fueron facilitadas por sus propios docentes [13, 15].

Las experiencias realizadas con este tipo de recursos educativos en una población variada de estudiantes permitieron asegurar que estas prácticas educativas contribuyen a la efectividad y facilidad en la enseñanza de conceptos fundamentales sobre Teoría de la Computación y Redes de Computadoras [13, 15]. Completada la experiencia se realizaron encuestas a la población estudiantil participante acerca de su percepción en relación a la utilización de los recursos y sobre su efectividad en términos de aprendizajes construidos permitiendo dar cuenta de lo positivo y alentador de los resultados. Las encuestas permitieron recuperar percepciones en relación a la utilidad, es decir, en qué grado los estudiantes consideran que la actividad desarrollada resultó útil para ayudar a comprender conceptos abstractos; en relación al impacto, si ellos consideraron que se trata de una experiencia placentera y sobre la organización, si consideraron que desarrollar una experiencia lúdica antes de la exposición formal es una forma adecuada de organizar actividades de aprendizaje.

Se busca realizar estudios a mayor escala con intención de otorgar mayor consistencia a los resultados. Las nuevas investigaciones buscan refinar y profundizar las conclusiones obtenidas, considerando una mayor población estudiantil, analizar la organización y diferenciación de los grupos de estudiantes.

- **Producir y evaluar nuevos recursos**

Las tendencias vigentes en el contexto de la enseñanza de la computación y los nuevos diseños curriculares reconocen que es

prioritario que los estudiantes comprendan la forma en la que trabajan las computadoras y dispositivos de uso diario. Se propone desarrollar y probar en las aulas una colección de actividades desenchufadas enchufadas para enseñar conceptos sobre Representación del Conocimiento.

La colección de recursos educativos adopta la forma de taller creativo donde se debe utilizar conocimiento disciplinar para comprender y crear nuevos sistemas de representación del conocimiento.

Los siguientes aspectos conforman la caracterización de los grupos destinatarios de la propuesta:

- Año de estudio: cualquier año de la escuela secundaria.
- Conocimientos disciplinares asumidos: Ninguno
- Tiempo destinado: aproximadamente 40 minutos.
- Tamaño del grupo recomendado: se recomienda que los grupos estén conformados por 3 a 4 integrantes. De una actividad participan varios grupos en interacción continua.

Aspectos metodológicos:

A continuación, se presentan algunas opciones metodológicas adoptadas para este trabajo.

- Promover la colaboración, la utilización de recursos físicos actúa como facilitador de la actividad grupal.
- Complejidad creciente, se parte de problemas simples. La complejidad se incrementa progresivamente y se distribuye en el grupo.
- Problemas reales, se analiza y trabaja sobre casos reales.
- Taller Creativo, se producen nuevos sistemas de representación del conocimiento.
- Aprender de la experiencia, centrado en producir conocimiento abstracto y

conceptual a partir de reflexionar sobre experiencias concretas.

3. Resultados obtenidos y esperados

En el contexto de las actividades desarrolladas en esta línea se avanzó en la definición de enfoques metodológicos basados en el Diseño Participativo e Investigación Participativa. Estos enfoques están destinados a la producción y evaluación de recursos educativos para enseñar Ciencias de la Computación.

Se realizaron trabajos de campo para abordar la enseñanza de temas relacionados a las redes de computadoras y las teorías de la computación. Los estudios desarrollados aportan resultados alentadores en relación a la efectividad de los recursos elaborados [13, 14, 15].

El diseño y estudio de efectividad fueron orientados por los modelos teóricos definidos en ésta Línea de Investigación.

Un próximo paso a desarrollar en esta línea es el de concretar trabajos de campo a mayor escala.

Por otra parte, se está trabajando en la ampliación de la colección de recursos educativos tendiendo a un cubrimiento mayor del campo disciplinar.

4. Formación de Recursos Humanos

Uno de los objetivos del desarrollo de esta línea de Investigación es contribuir a la formación de recursos humanos en el campo de Educación en Ciencias de la Computación. En este sentido, dos integrantes del proyecto de investigación están finalizando la Maestría en Ciencias de la Computación de la Facultad de Informática, UNCo. Además, dos de los autores del artículo están cursando la Maestría en Enseñanza en Escenarios Digitales, que desarrollan de manera conjunta las Universidades Nacionales de Cuyo, Comahue,

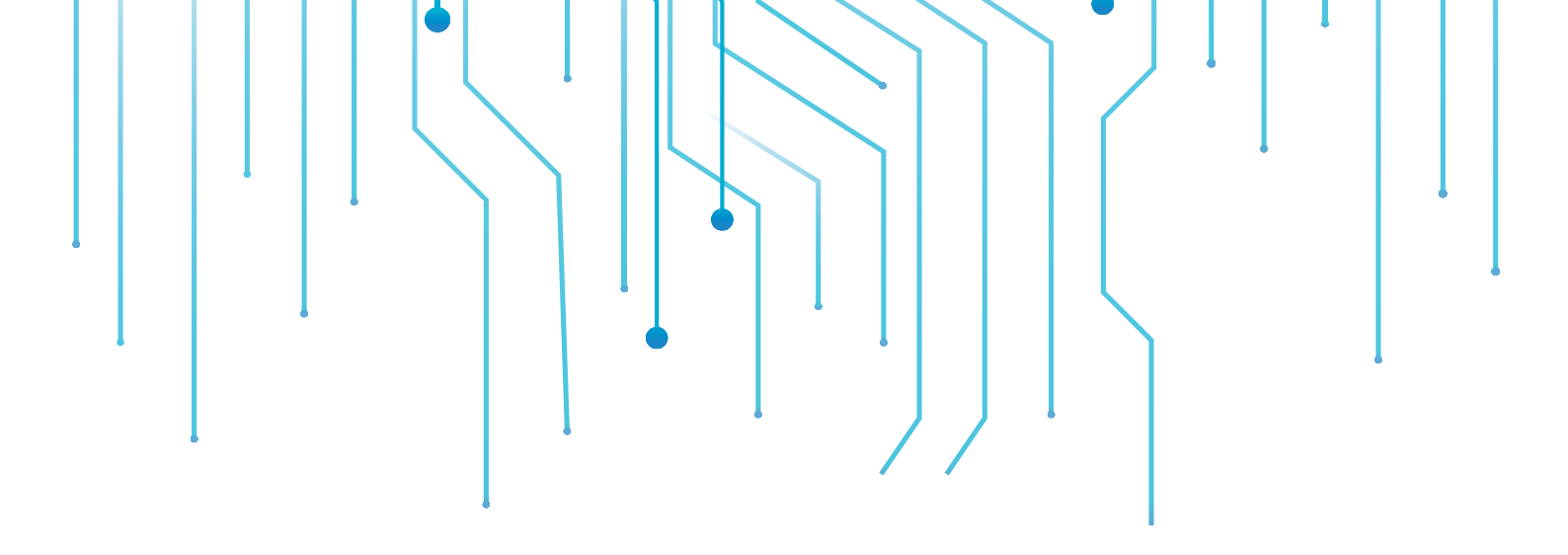
Patagonia Austral, Patagonia San Juan Bosco, San Luis, Chilecito y La Pampa.

A fines del año pasado, un integrante docente del proyecto ha defendido su tesis de Licenciatura en Ciencias de la Computación en la temática del grupo. Los integrantes alumnos del proyecto también desarrollan sus tesis en temas relacionados.

Por otra parte, actualmente se están desarrollando al menos tres tesis de Licenciatura en temas de interés del grupo de investigación.

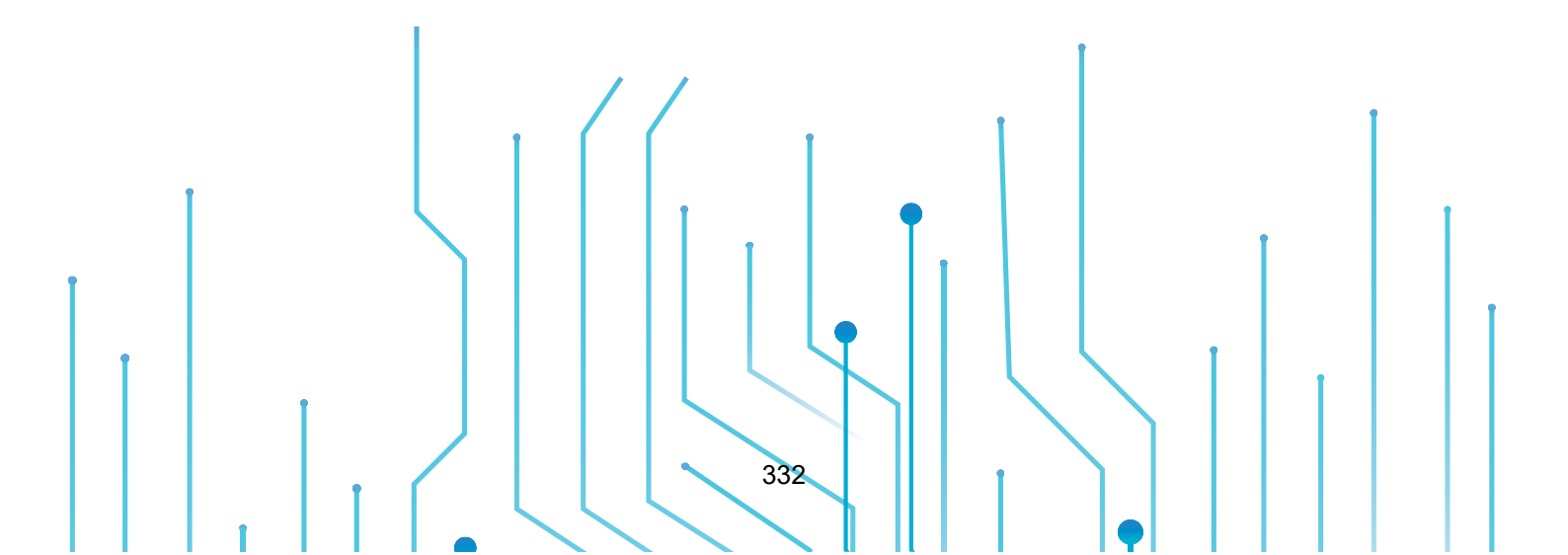
5. Referencias

- [1] T. Bell and J. Vahrenhold. Cs unplugged—how is it used, and does it work? In *Adventures Between Lower Bounds and Higher Altitudes*. Springer, 2018.
- [2] T. Nishida, S. Kanemune, Y. Idosaka, M. Namiki, T. Bell, and Y. Kuno. A cs-unplugged design pattern. *ACM SIGCSE Bulletin*, 41(1):231–235, 2009.
- [3] R. Taub, M. Armoni, and M. Ben-Ari. Cs unplugged and middle-school students’ views, attitudes, and intentions regarding cs. *ACM Transactions on Computing Education (TOCE)*, 12(2):1–29, 2012.
- [4] H. Lewis and C. Papadimitriou. *Elements of the Theory of Computation*. Second Edition. Prentice Hall, 1998.
- [5] J. Hopcroft, R. Motwani and J. Ullman. *Introduction to Automata Theory, Languages and Computation* (3rd Edition). Addison Wesley, 2006.
- [6] A. M. Colmenares E. Investigación-acción participativa: una metodología integradora del conocimiento y la acción. *Voces y Silencios. Revista Latinoamericana de Educación*, 3(1):102–115, 2012.
- [7] J. Martí. La investigación-acción participativa: estructura y fases. 2017.
- [8] K-12 Computer Science Framework Steering Committee. *The K–12 Computer Science Framework*. ACM, 2016.
- [9] F. Sadosky. CC – 2016 Una propuesta para refundar la enseñanza de la computación en las escuelas argentinas. Fundación Sadosky, Argentina, 2013.
- [10] R. Society. *After the reboot: Computing education in UK schools*. Policy Report, 2017.
- [11] C. P. E. de la Provincia de Neuquén. *Diseño Curricular Jurisdiccional de los tres primeros años de la Escuela Secundaria Neuquina*. Resolución N° 1463/18, 2018.
- [12] B. DiSalvo, J. Yip, E. Bonsignore, and D. Carl. Participatory design for learning. In *Participatory design for learning*, pages 3–6. Routledge, 2017.
- [13] Dolz, Daniel, et al. Recursos Educativos Desenchufados para la Enseñanza de las Ciencias de la Computación en la Escuela Secundaria. XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020). Neuquén, 2020.
- [14] Dolz, D., Parra, G., Rodríguez, J. Diseño Participativo para desarrollar Recursos Educativos para la Enseñanza de las Ciencias de la Computación en la Escuela Secundaria. *Una Máquina de Turing en la Escuela*. WICC 2022. U. Champagnat. Mendoza, Argentina.
- [15] Rodríguez, J, et al. *Una Máquina de Turing en la Escuela*. CACIC 2022. U. N. de La Rioja. La Rioja, Argentina.



IS

Ingeniería en Software



PRINCIPIOS ESENCIALES DE LA INGENIERÍA DE SOFTWARE EN ÁMBITO PÚBLICO.

Daniel Aguil Mallea, Horacio Pendentí, Moyano Ezequiel, Karina Manzaraz, Cintia Aguado, Norma Vecchi, Matías Mancilla, Lautaro Barba, Federico Borgna.
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: ++54-2901-432403
{daguil, hpendenti, [emoyano](mailto:emoyano@untdf.edu.ar), [amanzaraz](mailto:amanzaraz@untdf.edu.ar), [caaguado](mailto:caaguado@untdf.edu.ar), [ngvecchi](mailto:ngvecchi@untdf.edu.ar)}@untdf.edu.ar, matiassnm.1@gmail.com,
fborgna_52@hotmail.com, lau_mb94@hotmail.com

Resumen

Desde hace años el software es parte de nuestras actividades cotidianas, tanto que ya se piensa que nuestra civilización depende del mismo.

Las prácticas actuales de ingeniería de software requieren de desarrolladores con más conocimientos sobre métodos y procesos, que van más allá de lo estrictamente técnico.

En este sentido, la formación profesional de cada integrante es un aspecto relevante a la hora de armar un equipo ajustado a métodos.

En el ámbito público la dinámica de cambio en los requerimientos del software no escapa a la realidad común de otros ámbitos; sin embargo, existen otros factores que afectan a los proyectos y a los procesos del desarrollo del software que inciden en la calidad del producto final, plazos de entrega y en los costos asociados.

La investigación se centrará en la necesidad de analizar y caracterizar los procesos de desarrollo de software en el ámbito de las instituciones públicas de Tierra del Fuego, a través de estrategias de análisis de datos, que permita determinar las mejores técnicas para mejorar la productividad y la calidad del software. La experiencia y conocimientos del equipo de investigación contribuirán a desarrollar software de calidad en las instituciones del ámbito público.

Palabras claves: Ingeniería de software, Administración Pública, Proceso de desarrollo de software.

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la Universidad Nacional de Tierra del Fuego, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática vinculada a la ingeniería de software, específicamente, su impacto en el ámbito de las instituciones públicas.

Actualmente esta línea se desarrolla en el marco de un proyecto de investigación (Oct/2022-Sept/2024) El mismo ha obtenido aval UNTDF.

El Proyecto se lleva adelante con un enfoque multidisciplinario por parte de organismos de investigación y gestión en la provincia de Tierra del Fuego.

1 Introducción

El desarrollo de software ha tenido importantes avances desde sus inicios mejorando en su transcurso las actividades y procesos que se utilizan, llegando a formar parte de nuestras actividades cotidianas[1]. Las metodologías, de hecho, han logrado divergir en función del contexto con la intención de explotar las características que subyacen en el ambiente para el cual se quiere producir.

Sin embargo, el desarrollo de software continúa presentando importantes desafíos, la construcción de sistemas de software es fundamentalmente difícil y cada vez más compleja[2]; requiere de profesionales

altamente formados (o especializados) en todos los procesos que involucra.

En el caso que nos ocupa, el ámbito público, la importancia del software se ve potenciada por las necesidades de una sociedad como ésta, que requiere (o necesita) de ciertas facilidades como *e-government*, gestiones transparentes y participación ciudadana, entre otros.

El desarrollo y mantenimiento de aplicaciones de este tipo y los propios ERPs, se convierten en actividades esenciales de alta demanda y cada vez más críticas, especialmente en cuanto a tiempo. Estas restricciones, sumadas a otras relacionadas con los recursos humanos, factores políticos y sociales, a menudo atentan contra la gestión planificada de proyectos de software.

Ante la ausencia de una receta mágica que permita producir software de alta calidad, en tiempo y costo, la alternativa es usar la Ingeniería de Software.

La IEEE define la Ingeniería de software como la aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento del software; esto es, la aplicación de la ingeniería al software. Esto requiere, entonces, que los equipos de desarrollo sean capaces de trabajar en forma sistemática, disciplinada y cuantificable; consecuentemente, los integrantes de estos equipos deben estar capacitados para hacerlo de esa manera.

La Ingeniería de software, en adelante IS, es la disciplina que viene a dar respuesta a las dificultades de desarrollar software de calidad en tiempo y presupuesto acordado. Define procesos y métodos para lidiar con los puntos críticos del desarrollo.

Nuestra provincia no escapa a esa realidad donde los recursos humanos formados en la disciplina son particularmente escasos o se ven absorbidos por tareas rutinarias ante el poco personal a su cargo, sin poder atender aspectos vinculados al desarrollo; motivo por el cual es usual encontrarse con dificultades al intentar

que un equipo de desarrollo utilice enfoques metodológicos para obtener mejores resultados. Si a lo anterior se le suman los inconvenientes derivados de planificaciones de sistemas que han sido desarrollados de manera “artesanal” o alejados de los lineamientos que propone la disciplina, estos resultados son aún más pobres.

Se pueden sintetizar las actividades de la Ingeniería de Software como la realización del modelado para la adquisición de conocimiento, enfocándose solamente en los detalles relevantes, lo que además permite realizar un proceso de abstracción en pos de reducir la complejidad inherente al producto.

En términos de experiencia podemos asegurar que contar con procesos bien definidos y aplicarlos metodológicamente reduce el riesgo asociado a desarrollar sistemas complejos[3].

Cabe destacar que los procesos que se mencionan hacen referencia a los involucrados a lo largo de todo el desarrollo de software incluida su evolución, distinción cada vez más irrelevante[4].

La tarea de mejorar los procesos de desarrollo es difícil, y por lo general no sale como uno quisiera[5], aunque cuando la tarea logra resultados satisfactorios estos son profundos: los proyectos de software tienen más probabilidad de completarse con el tiempo, mejora la comunicación dentro del equipo involucrado en el desarrollo de software, el nivel de confusión y caos que con frecuencia prevalece para grandes proyectos de software se reduce de manera sustancial, el número de errores que encuentra el cliente disminuye sustancialmente, la credibilidad de la organización de software aumenta y la administración tiene un problema menos por el cual preocuparse[6].

Un aspecto a tener en cuenta es la formación del equipo de desarrollo, obtener un conocimiento acabado de las capacidades de los miembros del equipo, además de estudiar y analizar los programas de capacitación continua en el área, podría manifestar las posibles

relaciones respecto a la falta de aplicación o implementación de los procesos que propone la disciplina.

El software es “conocimiento” y los organismos deben resguardar el suyo con la mayor calidad posible. Su autonomía dependerá en gran parte de la calidad y cantidad del conocimiento que posean.

En el marco de los problemas enunciados, la presente investigación abordará el estudio de las situaciones particulares que enfrentan los organismos públicos[7] en nuestro contexto provincial, en cuanto a la producción de software.

Motiva llevar a cabo el presente proyecto realizar una contribución al conocimiento sobre la aplicación de mejores prácticas de Ingeniería de Software en las áreas de TICs pertenecientes a instituciones públicas[8] [9].

Los procesos de desarrollo de software en el ámbito del sector público, sobre todo en la provincia de Tierra del Fuego, están fuertemente vinculadas a la relación recurso-tiempo; existen factores propios y ajenos a la disciplina que atenta contra la necesidad de llevar proyectos de esta naturaleza abordando las metodologías adecuadas.

Para ello, será necesario conocer con mayor detalle las situaciones que afectan a los equipos de las áreas de desarrollo de software en la provincia, elaborar un plan para recolectar información relevante, realizar un análisis sobre esta base detectando características esenciales que merecen especial atención y realizar un aporte al uso de las mejores prácticas de la ingeniería de software para lograr los objetivos en este tipo de contextos más particular, promoviendo recomendaciones de acción en pos de producir software de mayor calidad en forma más efectiva; a través de la cual no solo se beneficiarán los empleados y directivos de esas instituciones que harán uso de los desarrollos, sino toda la comunidad que recibe los servicios de cada una.

2 Línea de Investigación y Desarrollo

Como parte de la línea de investigación y desarrollo, el equipo de investigadores lleva adelante trabajos relacionados con las temáticas vinculadas a la ingeniería de software, específicamente, su impacto en el ámbito de las instituciones públicas; se plantean los siguientes objetivos a alcanzar en el presente proyecto:

Objetivos Generales

A- Analizar los aspectos principales que intervienen en el desarrollo de software en las instituciones públicas.

B- Generar un programa esencial que permita el desarrollo de software de calidad, en tiempos y costos razonables, en el ámbito de las instituciones públicas de nuestra provincia.

Objetivos Específicos

1.A - Identificar las actividades y estrategias utilizadas en las áreas de desarrollo de software en el ámbito de las instituciones públicas.

2.A - Caracterizar las actividades y estrategias involucradas en el desarrollo de software que realizan las instituciones del ámbito público de la provincia de TDF.

3.A - Clasificar las actividades y estrategias propuestas en términos de esfuerzo, a través de la relación recurso-tiempo en las instituciones del ámbito público.

1.B - Realizar un Mapeo sistemático de la literatura que la disciplina propone en términos de esfuerzo en general y, sobre todo, experiencias en el ámbito público.

2.B - Establecer lineamientos generales y particulares para el programa que permitan alcanzar el desarrollo de calidad en el ámbito público.

3.B - Validar los lineamientos del programa diseñado.

Hipótesis

Los productos de software realizados en las instituciones públicas podrían ser mejorados sustancialmente en términos de calidad, tiempo y costo si se aplicara un conjunto reducido de principios que la ingeniería de software propone. El conjunto de principios deberá ser consistente con las dinámicas que ocurren en las instituciones públicas de esta provincia con características particulares fuertemente influenciadas por su regionalidad.

Generar un programa que permitan el desarrollo de software en el ámbito de las instituciones públicas de nuestra provincia será fundamental para lograr la calidad de desarrollo deseado; la formación de recurso humano especializado específicamente en áreas de desarrollo de software incentivar políticas capacitación continua, como la aplicación de metodologías que aseguren el desarrollo de software permitirán obtener productos de calidad en tiempo adecuado, con los recursos disponibles y bajo presupuestos establecidos.

Resultados Esperados

El proyecto se encuentra recién iniciado por lo que a la fecha no pueden enumerarse resultados obtenidos. El grupo de investigación ha comenzado a trabajar en la primera actividad del proyecto que consiste en identificar las actividades y estrategias utilizadas en las áreas de desarrollo de software en el ámbito de las instituciones públicas, que sean consideradas de interés para el proyecto

Para lo cual se espera lograr:

- Determinar cuál es la situación actual de nuestra provincia sobre la industria del software en el ámbito público visibilizando las situaciones generales o comunes que afectan a los grupos de desarrollo.
- Comprender las problemáticas que subyacen en los organismos públicos serán aportes sustanciales para la formación de profesionales de la disciplina.

- Producir un tablero de situación respecto a los tiempos y recursos que demanda cada actividad realizada (junto a sus efectos, demoras, etc.) para conocer en detalle el área de desarrollo en cada organismo estudiado.
- Elaborar documentos de trabajo que sirvan como programas o guías para la difusión e interacción con los organismos.
- Desarrollar y definir los lineamientos generales y particulares que permitan a las áreas de desarrollo de cada organismo, determinar y alcanzar el programa en pos de obtener software de calidad.

Formación de Recursos Humanos

El Equipo de Trabajo está conformado por docentes investigadores de la UNTDF, licenciados en informática y alumnos de la Carrera.

El Proyecto incluye docentes en etapa de formación de postgrado, el Director del proyecto, Daniel Aguil está realizando actualmente su tesis de Magister en Ingeniería de Software en la Universidad Nacional de La Plata.

El co-director Horacio Pendentí se encuentra elaborando su propuesta de Tesis para el magister en Ingeniería de Software de la UNLP.

El docente Ezequiel Moyano se encuentra cursando el Magister en Ingeniería de Software de la UNLP.

La docente Cintia Aguado se acaba de recibir de Magister en Políticas Públicas.

Colaborará y prestará apoyo a un grupo de alumnos en formación de grado, con el objetivo de formarlos en los aspectos generales de la investigación y que le sirva para el desarrollo de su futura tesis de grado. Los alumnos Matías Mancilla y Lautaro Barba se encuentran cursando el último año de la carrera, mientras que el alumno Federico Borgna se encuentra desarrollando el trabajo final para el título de Analista Universitario de Sistemas.

Por otra parte, a través del proyecto se espera consolidar un equipo de trabajo de la UNTDF en la temática, sumando nuevos integrantes al grupo principal (el cual ya ha ejecutado numerosos proyectos vinculados al problema).

A su vez los docentes podrán nutrirse de la experiencia de quienes manejan diariamente los problemas y conflictos vinculados a la enseñanza de Sistemas.

Referencias

1. Stroustrup B., Software Development for Infrastructure, IEEE Computer Society. 2012.
2. Booch G., EL Lenguaje de Modelado Unificado: Un manual de referencia, Addison-Wesley, 1994.
3. Caballero, Sergio; Kuna, Horacio Daniel, Análisis y gestión de riesgo en proyectos software Un nuevo modelo integrando la metodología SEI y Magerit2. WICC 2018, Universidad Nacional del Nordeste, 2018.
4. Sommerville I., Software Engineering, Pearson Education, Harlow.2016.
5. Jacobson I, El proceso unificado de desarrollo de software , Addison-Wesley, 2013.
6. Pressman R., Software Engineering: A Practitioner's approach, Mc Graw Hill, New York, 2015.
7. Grandinetti, Rita María, Nuevas Tecnologías de la Información y Gestión de Recursos Humanos en el Ámbito Público Local: El caso de la Municipalidad de Rosario Revista Venezolana de Gerencia- 2000, ISSN 1315-9984.
8. Eduardo Iberti, Pablo Pytel y Florencia Pollo-Cattaneo. Evaluación de la viabilidad para proyectos de implementación ERP en la administración pública Argentina, - WICC 2015.
9. Sánchez Albavera, Fernando, Planificación estratégica y gestión pública por objetivos,

Instituto Latinoamericano y del Caribe de Planificación Económica y Social – 2003.

IEEE. iee.org.

Information Systems for eGovernment: A Quality-of-Service Perspective. Springer 2010.

Ley Nacional N° 25.326, de Protección de Datos Personales.

Ley Nacional N° 25.856 de Producción de Software.

Ley Nacional N° 25.922 y modificatoria 26.692, de Promoción de la Industria del Software.

Ley Provincial N° 1.279, de promoción de la Industria del Software en la Provincia de TDF, Software Público.

ONTI.

<https://www.argentina.gob.ar/onti/software-publico>.

Catálogo y recomendaciones de recursos (métodos) para facilitar experiencias de Diseño Participativo usando Design Thinking.

Corso Cynthia¹, Challiol Cecilia², Paz Menville Alejandra³, Ligorria Karina³, Serra Silvio³.

¹CIDS, UTN, FRC and also Facultad de Informática UNLP
Córdoba, Argentina

²LIFIA, Facultad de Informática, UNLP and also CONICET
La Plata, Buenos Aires, Argentina

³CIDS, UTN, FRC
Córdoba, Argentina

{ccorso,amenville,kligorria,sserra}@sistemas.frc.utn.edu.ar, ceciliac@lifia.info.unlp.edu.ar

RESUMEN

Actualmente el *Diseño Participativo* (DP) ha alcanzado un gran auge ya que representa un enfoque que permite el diseño de productos y/o servicios involucrando de manera activa a las personas destinatarias de los mismos, esto puede implicar usuarios finales con más experticia que otros. Uno de los principales aportes del DP es que garantiza que el producto o servicio diseñado se adapte realmente a las necesidades y preferencias de los usuarios finales. Existen diversas maneras de conducir una experiencia de DP, una de ellas es mediante el *Design Thinking* (DT) que ha sido usado mayormente en las organizaciones para lograr innovación de productos y/o servicios. Para poner en práctica DT existen diversos frameworks, cada uno de ellos definen sus propias etapas. Para llevar a cabo cada etapa se pueden utilizar distintos recursos, los cuales mayormente no están bien documentados o la información existente está difuminada por diversas fuentes. El objetivo de esta investigación es proponer un catálogo de recursos (métodos) que pueden ser usados en una experiencia de DP desde un abordaje de DT; buscando

que las soluciones de software resultantes sean de carácter innovador. Además para cada uno de estos recursos se brindarán recomendaciones de su puesta en práctica.

Palabras claves: *Diseño Participativo, Design Thinking, recursos, catálogo, recomendaciones.*

CONTEXTO

Este trabajo hace referencia al proyecto de investigación "Catálogo y recomendaciones de recursos (métodos) para facilitar experiencias de *Diseño Participativo* usando *Design Thinking*" PID- SITCCO0008594, homologado y financiado por la Secretaría de Ciencia, Tecnología de la Universidad Tecnológica Nacional. El contexto de desarrollo de la presente investigación es el Centro de Investigación, Desarrollo y transferencia de Sistemas de Información (CIDS) radicado en la U.T.N Facultad Regional Córdoba.

1. INTRODUCCIÓN

Actualmente el desarrollo de software es una temática de gran interés, ya que tiene un gran impacto en la vida de las

personas facilitando el desarrollo de ciertas actividades en general. El *Diseño Participativo* es fundamental para comprender y más aún empatizar con las necesidades reales de las personas para las que se está diseñando software [1]; cuando esto no acontece se generan experiencias frustrantes para los usuarios [2]. Para dar solución a esto, ha emergido el *Diseño Participativo*, en el cual las personas destinatarias del software no solo pasan a ser parte del equipo que diseña sino que además toman un rol esencial en el proceso de diseño [1]; se convierten expertos en el dominio.

El *Diseño Participativo* viene siendo explorado hace varios años [1],[3] y [4] contando con una diversidad de fuentes de información y artículos de contribución. Por ejemplo, en [5] pone el foco en el estudio aspectos relacionados con la evaluación del *Diseño Participativo*; mientras que en [6] se enfatiza la importancia del *Diseño Participativo* considerando elementos que son significativos para una experiencia de este estilo; así también existen otras fuentes de información en relación con los recursos (métodos) que se usan en las experiencias de este estilo [7] y [8].

El *Diseño Participativo* no tiene una única manera de ser conducido, en realidad existen distintas maneras de ponerlo en práctica, de hecho en [9] se destaca que *Design Thinking* podría ser muy adecuado para llevar a cabo el *Diseño Participativo*; ya que tiene el potencial de innovar en las soluciones que se obtienen. Una de las definiciones más conocidas de *Design Thinking* es enunciada por Cross [10] planteando que "el término *Design Thinking* se refiere a los procesos cognitivos del trabajo de diseño, o las habilidades de pensamiento y las prácticas que los diseñadores

utilizan para crear nuevos artefactos o ideas y resolver problemas". Al igual que el *Diseño Participativo*, *Design Thinking* es una disciplina que también viene siendo explorada y estudiada. En relación a los recursos (métodos) usados en *Design Thinking* existen diversas fuentes de información publicadas en internet [11], [12], [13], [14] y [15] como así también artículos de contribución [16], [17] que los mencionan; sin embargo estos recursos están, en algunos casos pocos documentados o hay escasa información respecto a su puesta en práctica.

En [9] se destacan las ventajas que tiene el *Diseño Participativo* para desarrollar software; remarcando que se necesita considerar al menos seis aspectos críticos: enfoques adecuados, diferencias individuales, ambiente, flexibilidad, comunicación, aspectos sensoriales y desafiar los supuestos. De hecho todas las personas poseen diferentes formas o estilos de comunicación, auditiva, visual o kinestésica; o una combinación de ellas [18]. Esto refuerza lo que plantea [9] en cuanto a la necesidad de acondicionar los recursos (métodos) a usar en las experiencias de *Diseño Participativo* teniendo en cuenta las diferentes formas de comunicación que pueden tener las personas. Sin embargo, todavía esto está poco explorado tanto en *Diseño Participativo* como en *Design Thinking*.

Por lo antes mencionado, se puede concluir que no existe disponible un catálogo centralizado de recursos (métodos) que incluya una caracterización detallada de los mismos, como tampoco hay documentación existente que contemple aspectos en relación a la posibilidad de acondicionarlos teniendo en cuenta los diferentes estilos de comunicación de las

personas. Esto acontece tanto en el *Diseño Participativo* como en *Design Thinking*. Más aún, tampoco existe información de cómo co-diseñar cuando se ofrecen diferentes alternativas de participación (oral, chat, etc.) para responder cada recurso y luego hay que converger en único resultado para continuar co-diseñando desde ahí. Por ejemplo, si se tiene en una etapa de ideación respuestas en chats, algunas orales y otras en post-it; cómo todo esto se unifica para clasificarlas y filtrar cuáles ideas terminan siendo elegidas para prototipar. Todos estos aspectos que aún no están cubiertos son la principal motivación del presente proyecto de investigación.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

Este proyecto de investigación se enmarca en diversas líneas de investigación y desarrollo. A continuación se listan detallando temas específicos de cada una de ellas:

- ☒ *Ingeniería de Software*
 - o Diseño
 - o Prototipado
- ☒ *Diseño Participativo*
 - o Abordajes
 - o Recursos (métodos)
 - o Rol del facilitador
- ☒ *Design Thinking*
 - o Frameworks conceptuales
 - o Recursos (métodos)
 - o Rol del facilitador
- ☒ *Estilos de comunicación*
 - o Visual
 - o Auditiva
 - o Kinestésica

3. RESULTADOS OBTENIDOS/ESPERADOS

La finalidad de este proyecto de investigación, planificado para ser ejecutado en el periodo 2023-2026, es lograr el cumplimiento de los siguientes objetivos:

Objetivo General

Proponer un catálogo de recursos (métodos) para el abordaje de experiencias de *Diseño Participativo* usando *Design Thinking*, brindando recomendaciones para los facilitadores de este tipo de experiencias.

Objetivos Específicos

- ☒ Analizar y documentar las bases conceptuales del *Diseño Participativo* y *Design Thinking*.
- ☒ Recopilar, analizar y documentar los recursos (métodos) usados tanto en el *Diseño Participativo* como *Design Thinking*.
- ☒ Proponer recomendaciones para la puesta en práctica de los recursos (métodos) relevados teniendo en cuenta el tipo de comunicación preponderante (visual, auditiva y kinestésica) que puedan tener las personas.
- ☒ Formular recomendaciones para los facilitadores de experiencias de *Diseño Participativo* usando *Design Thinking* en situaciones en que los participantes tienen distintos tipos de comunicación preponderante.

Resultados esperados

Al momento de la publicación de este trabajo, el proyecto se encuentra en un estadio inicial, por lo cual se detallan las principales actividades planificadas para

el cumplimiento de los objetivos anteriormente mencionados:

☒ Búsqueda de bibliografía y artículos de contribución de relevancia en la temática de *Diseño Participativo* y *Design Thinking*, considerando el uso de bases de datos como (BASE, Redalyc, Scopus entre otros).

☒ Lectura de los resúmenes (abstract) de todos los artículos de contribución que forman parte del resultado de la búsqueda, de esta manera se podrá detectar y descartar los duplicados o aquellos que no correspondan a la temática bajo estudio.

☒ Registrar un listado de los nombres de los recursos (métodos) relevados, especificando además aspectos o detalles que se mencionan en cada una de las fuentes de información; esta mecánica de trabajo prevista permitirá detectar los recursos (métodos) que estén repetidos y sobre todo complementar aspectos que quizás un trabajo lo menciona mientras que otro no.

☒ Clasificación inicial de los recursos (métodos) relevados considerando como criterio los tipos o formas de comunicación preponderante en las personas, están pueden ser auditivas, visual y kinestésica [18].

☒ Elaboración de recomendaciones considerando como debería ser la puesta en práctica del recurso (método) clasificado, para el tipo de comunicación no considerado inicialmente en el recurso.

☒ Prueba piloto para determinar si las recomendaciones propuestas en relación a la puesta en práctica son viables o no.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto está conformado por docentes-investigadores pertenecientes a

la carrera de grado de Ingeniería en Sistemas de Información. Todos los integrantes docentes del PID han participado del proceso de categorizaciones en investigación dentro del Programa de Incentivos del MECyT; así como en la categorización interna que posee la U.T.N.

En el contexto de este proyecto de investigación se está desarrollando el trabajo de Tesis de Doctorado en Ciencias de la Informática de la directora del proyecto; y cuenta con la asesoría de una investigadora formada perteneciente a la Facultad de Informática de la Universidad Nacional de La Plata, doctora en Informática, especialista en la temática de *Design Thinking* y cumple el rol de directora de la tesis doctoral mencionada anteriormente.

Este proyecto contempla la participación de becarios estudiantes con el objetivo de complementar su formación académica con un acercamiento al ámbito de la investigación científica.

5. REFERENCIAS

[1] Sanders, E. B. N. (2002). From user-centered to participatory design approaches. In *Design and the Social Sciences*, pp. 18-25.

[2] Kudryashov, L. (2021) "Participatory Design of Augmentative and Alternative Communication (AAC) Technology with Autistic Adults", Thesis of Master Science Information, University of Michigan.

[3] Luck, Rachael (2018). What is it that makes participation in design participatory design? *Design Studies*, 59 pp. 1-8.

[4] Hansen, N. B., Dindler, C., Halskov, K., Iversen, O. S., Bossen, C., Basballe,

- D. A., & Schouten, B. (2019). How participatory design works: mechanisms and effects. In Proceedings of the 31st Australian Conference on Human-Computer-Interaction, pp. 30-41.
- [5] Bossen, C., Basballe, D. A., & Schouten, B. (2019). How participatory design works: mechanisms and effects. In Proceedings of the 31st Australian Conference on Human-Computer-Interaction, pp. 30-41.
- [6] Bødker, S., & Kyng, M. (2018). Participatory design that matters-Facing the big issues. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(1), 1-31.
- [7] Shubhangi Choudhary. Participatory Design -Tools and Techniques: Purpose and Context. Recuperado de: <https://medium.com/@i.shubhangich/participatory-design-tools-and-techniquespurpose-and-context-cc877790d4a6>. Fecha de último acceso: 02/06/2022.
- [8] Awais Hameed Khan. (2020). Participatory Design Tools: Exploring the Value of Design through Materiality. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended*.
- [9] Maun, R., Fabri, M. and Trevorrow, P. (2021). Adapting Participatory Design Activities for Autistic Adults: A Review. *International Conference on Human-Computer Interaction*, pp. 300-314.
- [10] Cross N., "Design thinking: Understanding how designers think and work," Berg, Oxford.New York, 2011.
- [11] Baella.com. 22 Herramientas de Design Thinking fase a fase. Recuperado de: <https://baella.com/noticias/22-herramientas-de-design-thinking-fase-a-fase>. Fecha de último acceso: 01/06/2022.
- [12] Design Thinking en España. Herramientas de Design Thinking. Recuperado de: <https://xn--designthinkingespaa-d4b.com/herramientas-de-design-thinking>. Fecha de último acceso: 30/05/2022.
- [13] Thinkernautas. 34 herramientas para el Design Thinking en un rápido vistazo. Recuperado de: <https://thinkernautas.com/34-herramientas-para-el-design-thinking-en-un-rapido-vistazo>. Fecha de último acceso: 31/05/2022.
- [14] Design Thinking Comunidad on-line. Herramientas Design Thinking. Recuperado de: <https://www.designthinking.services/descargar-herramientas-design-thinking/>. Fecha de último acceso: 02/06/2022.
- [15] IDOM.Herramientas para implementar la metodología "Design Thinking". Recuperado de: <https://idomcpi.com/herramientas-para-implementar-la-metodologia-design-thinking/>. Fecha de último acceso: 28/05/2022.
- [16] Hehn, Jennifer; Ubernickel, Falk; Herterich, Matthias. (2018). *Design Thinking Methods for Service Innovation-A Delphi Study*.
- [17] Plattner, H. (2018). Guía del proceso creativo. Mini guía: una introducción al Design Thinking+ Bootcamp bootleg.
- [18] Lakin Duane. (2007). "Vender con PNL: Una ventaja oculta." Ed. 2da edición. Barcelona(España). Editorial Sirio.

Análisis de viabilidad del despliegue de sistemas de software en PyMES

Leandro Moreno¹, Marisa Panizzi^{1,2}, Rodolfo Bertone³

¹Programa de Maestría en Ingeniería en Sistemas de Información. Escuela de Posgrado. Universidad Tecnológica Nacional. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires. Medrano 951 (C1179AAQ), C.A.B.A., Argentina.

²Departamento de Ingeniería en Sistemas de Información. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires. Medrano 951 (C1179AAQ), C.A.B.A., Argentina.

³Instituto de Investigación en Informática (III-LIDI). Facultad de Informática. Calles 50 y 120 - La Plata - Bs. As. – Argentina.

lm.moreno1991@gmail.com; marisapanizzi@outlook.com; pbertone@lidi.info.unlp.edu.ar

Resumen

La línea I+D+I presentada se enmarca en el diseño de herramientas que permitan fortalecer el proceso de despliegue de sistemas de software en Pequeñas y Medianas Empresas (PyMES) de Argentina. La creciente demanda por parte de estas organizaciones de procesos de software sistematizados y controlados en línea responde a la necesidad de desarrollarse en entornos altamente competitivos en la industria, tanto a nivel nacional como internacional. En consecuencia, estas empresas requieren métodos y herramientas aplicables a la mejora y estabilización de sus procesos, considerando los alcances y limitaciones propios de su tamaño y capacidad operativa. Esta línea de investigación y desarrollo pretende brindar soluciones a las PyMES para que anticipen los inconvenientes en su proceso de despliegue de sistemas de software. En este artículo se presentan los hallazgos significativos de un mapeo sistemático de la literatura realizado en la construcción del estado del arte sobre el tema.

Palabras clave: proceso de despliegue, sistemas de software, viabilidad, PyMES.

Contexto

La línea I+D+I que se reporta en este artículo es financiada parcialmente por un proyecto de

investigación titulado “Estudio del proceso de implantación de sistemas informáticos en el contexto industrial de la República Argentina” (Código SIUTNBA0006576) de la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional. Cuenta con el asesoramiento científico del grupo de investigación de Ingeniería de Software del Instituto de Investigación en Informática (III-LIDI) de la UNLP.

Introducción

La implementación de estándares existentes en la industria del software frecuentemente se percibe costosa a nivel económico, así como también compleja en lo que respecta a la documentación y la burocracia interna requerida [1].

En este sentido, las organizaciones más pequeñas deben enfrentar el desafío de hallar el equilibrio entre las prácticas cotidianas horizontales, propias de las PyMES, y los procedimientos formales, definidos y documentados inherentes a la aplicación de procesos de mejora de software más avanzados [2][3].

Dado que aproximadamente el 80% de las organizaciones dedicadas al desarrollo de software en la Argentina son PyMES, las contribuciones realizadas en la materia

resultan de particular interés por su potencial alcance[4].

Dentro de las actividades relacionadas con la industria del desarrollo de sistemas de software, se encuentra el despliegue, como un proceso crucial del ciclo de vida de desarrollo de software en el que el sistema de software finalmente estará operativo para que el cliente pueda beneficiarse económicamente de su uso [5].

Los despliegues de sistemas de software comprenden actividades y prácticas que pueden fallar o poner en riesgo el éxito del proceso: ausencia de componentes (externos), descargas incompletas y despliegues erróneos [6], mientras que frecuentemente se presentan inconvenientes que eventualmente se resuelven como parte de la fase de mantenimiento. En consecuencia, las empresas pueden demorarse meses o años en lograr la estabilidad de un sistema tras finalizar el despliegue. Por este motivo, un proceso eficiente de despliegue de software es aquel que permita ahorrar recursos financieros y humanos al reducir o mitigar el impacto estas incidencias [7].

En instancias anteriores de esta investigación, en la que se pretende diseñar un método que permita evaluar la factibilidad de los despliegues de sistemas de software, se proyectó y concretó la realización del estado del arte respecto al análisis de factibilidad en los procesos de despliegue de sistemas de software [8].

Este proyecto se inició integrando la óptica tecnológica al análisis organizacional, teniendo en cuenta que éste se encuentra atravesado por la complejidad de sus elementos, el cambio en la forma de trabajar de las personas y falta de experiencia y habilidades [9].

A los efectos de obtener una visión integral del tema, se consideró la categorización realizada por Petersen *et al.* en [10] en la que se reconocen cuatro elementos centrales para la evaluación de proyectos de software, adicionales a la organización y al mercado. Estos elementos son:

1) Producto: el sistema de software desarrollado. Entre los aspectos que permiten entender y evaluarlo en contexto, se destacan: la madurez, la calidad, el tamaño, el tipo de sistema, el nivel de personalización o adaptación y el lenguaje de programación.

2) Procesos: el flujo de trabajo del desarrollo. Dentro de los aspectos de análisis se encuentran las actividades involucradas, el flujo de trabajo y orden de ejecución de las actividades y los artefactos resultado de cada actividad o iteración.

3) Prácticas, herramientas y técnicas: en este grupo se comprenden los aspectos que brindan un abordaje sistematizado en la interacción con el objeto de estudio. Aquí destacan las herramientas CASE y las metodologías aplicadas.

4) Factor humano: este aspecto es también muy relevante, dado que tiene un gran impacto sobre el proceso de desarrollo. En este punto, debe atenderse a los roles asignados al proyecto de software y su experiencia.

5) La organización comprendida por la estructura, atendiendo al modelo jerárquico y flujo de comunicaciones, las certificaciones que obtiene a partir de la capacitación y la proyección (local, internacional).

6) El mercado: compuesto por los competidores y potenciales clientes, se ve afectado por el número de estos últimos, el

segmento involucrado, la estrategia de ventas en el largo plazo y restricciones.

Estos aspectos analizados serán considerados para el diseño del método de cálculo de viabilidad de despliegue de sistemas de software.

En consecuencia y a los efectos de conocer el estado del arte, se realizó un mapeo sistemático de la literatura (en inglés, *Systematic Mapping Study* o SMS), de acuerdo con las directrices propuestas por Kitchenham *et al.* [11]. Este procedimiento comprende una etapa de planificación en la que se definió la cadena de búsqueda, se identificaron los términos principales, se detallaron los criterios de inclusión y exclusión, y se realizó el esquema de clasificación de los artículos a partir de las siguientes preguntas de investigación:

PI1: ¿Qué aportes se han realizado acerca de la evaluación de factibilidad sobre los procesos de despliegue de software?

PI2: ¿Qué factores se consideran en las evaluaciones que se realizan sobre los procesos de despliegue de software?

PI3: ¿Qué tipos de investigaciones se encuentra en los artículos?

Para dar respuesta a dichas preguntas, se dio lugar a la fase de ejecución del SMS, en la que se leyeron y clasificaron los artículos primarios hallados de acuerdo con el protocolo de revisión definido en la etapa de planificación. Esto ameritó una búsqueda automática en las librerías y plataformas digitales *Scopus*, *IEEE Xplore* y *ACM Digital Library*, por tratarse de las bibliotecas más utilizadas en investigación en Ingeniería del software, seleccionando artículos de congresos y artículos de revistas. La búsqueda se realizó en el período comprendido entre enero del año 2016 hasta julio del año 2022. Finalmente, de un total de

177 artículos encontrados, se analizaron en total, 25 estudios primarios.

Luego en la etapa de análisis de los estudios primarios, se organizaron los resultados de acuerdo a la clasificación definida en el protocolo de revisión para responder a dichas preguntas.

Respecto a la primera pregunta de investigación (PI1), los resultados del muestreo estudiado arrojan que mayoritariamente, los aportes en la materia son guías, modelos o herramientas (20% cada categoría), aplicables en entornos on-cloud y en contextos de despliegue continuo. Por otra parte, se destacan otros tipos de contribuciones, en los que son presentados casos de uso y pruebas de aplicación diseñadas para algún sistema en particular.

Por otra parte, la segunda pregunta de investigación (PI2), respecto a los factores tenidos en cuenta sobre las evaluaciones de los procesos de despliegue de software, se destaca el foco realizado sobre el factor proceso (92% del muestreo). En menor medida, se encuentran analizados otros factores como la organización (36%), las prácticas, herramientas y técnicas (26%) y el producto (28%). Finalmente, solo el 16% recogen al factor humano dentro de su trabajo, mientras que el mercado es observado en el 8%.

Este resultado permite concluir inicialmente que el foco principalmente se encuentra sobre los aspectos técnicos y procedimentales de los despliegues, y que existe una relativa vacancia de trabajos que analicen las variables del personal involucrado y el mercado en que se desarrolla la actividad.

Finalmente, respecto a la última pregunta de investigación (PI3), se identifica que en proporción similar se hallan trabajos correspondientes a propuestas de soluciones

(36%), validaciones (36%) y evaluaciones (28%) según la clasificación propuesta por Wieringa *et al.* [11].

En líneas generales, el trabajo realizado permitió comprobar la relevancia y actualidad del tema presentado, así como también el carácter prospectivo con foco en los procesos que tienen la mayoría de los artículos encontrados.

Este hallazgo tiene particular sentido al analizar el contexto y la dificultad en la implementación de procesos estandarizados que tienen las PYMES, en línea con lo expuesto por Mishra *et al.* [3]. Adicionalmente, resulta relevante la ausencia del análisis del factor mercado en diálogo con la variable organizacional, tomando en cuenta que las actividades de despliegue de software dentro de la industria frecuentemente están asociadas a entidades lucrativas.

En función de los resultados mencionados anteriormente y observando los objetivos del proyecto en curso, se encuentra en fase de análisis la composición de la dimensión “organización” como parte del método de cálculo de viabilidad de despliegue de sistemas de software y luego realizar una prueba de concepto en una PyME de Argentina.

Líneas de Investigación, Desarrollo e Innovación

La línea de I+D+I presentada consiste en el diseño de un método que permita calcular la viabilidad del proceso despliegue de sistemas de software. Una vez diseñado el método se

validará en la industria mediante un conjunto de estudios de caso.

Resultados y Objetivos

En esta línea de I+D+I en progreso respecto al fortalecimiento del proceso de despliegue de sistemas de software, se han logrado una serie de resultados que se detallan a continuación:

- a) Académicos, se han concluido dos trabajos de especialidad en Ingeniería en Sistemas de Información, una tesis de Maestría en Ingeniería en Sistemas de Información defendida y otra depositada y próxima a defender, y un trabajo de especialidad en Ingeniería en Sistemas de Información en desarrollo. Además, se ha defendido una tesis Doctoral en Ciencias Informáticas.
- b) Producción Científica: se ha presentado la línea de I+D+I en eventos científicos de alcance nacional (WICC¹ 2017, WICC 2019, WICC 2020, WICC 2021, WICC 2022, CACIC² 2016, CACIC 2018, CACIC 2019, CACIC 2020, CACIC 2021, CACIC 2022, CONAISI 2022³, y en el ámbito internacional, CIACA⁴ 2017, SEKE⁵ 2017, CIBSE⁶ 2019, CIBSE 2020 CIBSE 2021 y ICAETT⁷ 2019. Además, se lograron tres publicaciones en Springer, una publicación en la revista *Brazilian Journal of Development* y otra en la revista *Journal of Computer Science and Technology* (JCS&T).
- c) Proyectos de Investigación: los proyectos finalizados UTNBA4347 titulado: “Impacto del factor peopleware en el proceso de implantación de sistemas informáticos” y UTNBA6576 titulado: “Estudio del proceso de implantación de sistemas informáticos

¹ Workshop de Investigadores en Ciencias de la Computación

² Congreso Argentino de Ciencias de la Computación.

³ Congreso Nacional de Ingeniería Informática / Sistemas de Información

⁴ Conferencia Iberoamericana de Computación Aplicada.

⁵ International Conference on Software Engineering and Knowledge Engineering.

⁶ Congreso Iberoamericano en Ingeniería de Software.

⁷ Congreso Internacional sobre Avances en Nuevas Tendencias y Tecnologías.

en el contexto industrial de la República Argentina". El proyecto a ejecutarse a partir del 1 de abril PID SIECBA0008576 titulado "Herramientas para fortalecer el proceso de despliegue de sistemas de software".

- d) Formación en investigación: proceso de aprendizaje constante de métodos de investigación de ingeniería de software experimental [12].

Formación de Recursos Humanos

El grupo se encuentra conformado por un director, dos tesis de Maestría, de los cuales una de las tesis ha sido depositada, y un asesor científico-tecnológico. Se estima la formación de un Magister en Ingeniería en Sistemas de Información, que actualmente se encuentra desarrollando como primera instancia su trabajo de especialidad.

Referencias

[1] Laporte, C.Y., Alexandre, S., O'Connor, R.V. (2008). A Software Engineering Lifecycle Standard for Very Small Enterprises. Euro SPI 2008, CCIS, pp. 129-141.

[2] Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5, pp. 14-37.

[3] Mishra, Deepti & Mishra, Alok. (2009). Software Process Improvement in SMEs: A Comparative View. *Computer Science and Information Systems*. 6, pp. 111-140. 10.2298/CSIS0901111M.

[4] OPPSI. (2018). Reporte anual sobre el Sector de Software y Servicios Informáticos de la República Argentina. https://cessi.org.ar/wp-content/uploads/2022/02/oppsi_informe_coyuntur_a_2017.pdf

[5] Panizzi, Marisa Daniela (2022). Tesis doctoral. *DepProMod: Modelo de Proceso de Despliegue de Sistemas de Software*. Universidad Nacional de La Plata. <http://sedici.unlp.edu.ar/handle/10915/139509>

[6] Subramanian N. (2017). The software deployment process and automation. *CrossTalk*, 30(2), pp. 28-34

[7] Jansen S., Brinkkemper S, (2006). Definition and validation of the key process of release, delivery and deployment for product software vendors: Turning the ugly duckling into a swan IEEE International Conference on Software Maintenance, ICSM, art. No. 4021334, pp. 166-175.

[8] Moreno Leandro, Panizzi Marisa, Bertone Rodolfo (2022). Revisión de la Literatura sobre el Análisis de Factibilidad del Despliegue de Sistemas de Software. En las actas del 10mo. Congreso Nacional de Ingeniería en Informática y Sistemas de Información (CoNaIISI '23), pp. 371-378. <https://doi.org/10.33414/ajea.1146.2022>

[9] Reascos I., Carvalho J., Bossano S. (2019). Implanting IT Applications in Government Institutions: A Process Model Emerging from a Case Study in a Medium-Sized Municipality. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance*, pp. 80-85.

[10] Petersen, K. Wohlin, C. (2009). Context in industrial software engineering research. *Third International Symposium on Empirical Software Engineering and Measurement*. pp.402-404

[11] Wieringa R., Maiden N., Mead N., Rolland C. (2005). Requirements engineering paper classification and evaluation criteria: A proposal and a discussion. *Requirements Engineering*, 11, pp. 102-107.

[12] Genero, M., Piattini, M., & Cruz Lemus, J. A. (2014). *Métodos de investigación en Ingeniería del Software*. Madrid: Ra-Ma Editorial y Publicaciones.

Métodos, técnicas y herramientas para mejorar y evaluar la calidad de requisitos en proyectos ágiles de software

Gabriela Tomaselli; Noelia Pinto; César Acuña; Nicolás Tortosa
Centro de Investigación Aplicada a TIC (CInApTIC)
Universidad Tecnológica Nacional, Facultad Regional Resistencia (UTN-FRRe)
French 414, Resistencia, Chaco
{gabriela.tomaselli;ns.pinto;csr.acn;nicotortosa}@gmail.com

RESUMEN

A fin de contribuir a la calidad de los procesos de desarrollo de software guiados por prácticas ágiles, se ha presentado Agile Quality Framework (AQF) [1], integrando un modelo de calidad de software denominado Quality Agile Model (QuAM) [2][3], y una herramienta de software denominada QuAGI [4] que automatiza la gestión del modelo y permite obtener el nivel de calidad del proceso en cuestión.

El objetivo del Componente N° 4 de QuAM es mejorar la aplicación de la práctica ágil de definir necesidades como historias de usuario, promoviendo la importancia de contar con requerimientos bien definidos y completos, que contribuyan a la realización del proyecto.

Diversas experiencias de validación mostraron que una incorrecta obtención, análisis, especificación, y validación de requerimientos, aumenta los riesgos y podría volver inmanejable al proyecto. Sin embargo, la comunidad de desarrollo de software aún presenta desconocimiento acerca de cómo las prácticas de ingeniería de requerimientos pueden resolver problemas comunes asociados a esto y colaborar en la mejora de la calidad.

Se presenta aquí la línea de investigación orientada a ampliar AQF, contribuyendo a la mejora de la gestión de requerimientos y obteniendo un impacto positivo en la evaluación de calidad de proyectos ágiles de software.

Palabras clave: Calidad de software, Procesos ágiles de desarrollo de software, Gestión de requerimientos; Ingeniería de requerimientos

CONTEXTO

El trabajo que aquí se presenta está enmarcado en el proyecto “Métodos, técnicas y herramientas para mejorar y evaluar la calidad de requisitos en proyectos ágiles de software”, que es financiado por la Universidad Tecnológica Nacional (UTN) y ejecutado en el Centro de Investigación Aplicada a TIC (CInApTIC) de la Facultad Regional Resistencia, con el código SIECRE0008643.

Asimismo, algunas actividades son compartidas con el proyecto de investigación y desarrollo “Métodos, técnicas y herramientas para mejorar y evaluar la calidad de la capacidad de producción de entregables dentro de un proceso ágil de desarrollo de software” (SIECRE0008608TC), también financiado por UTN y ejecutado en el CInApTIC.

1. INTRODUCCIÓN

En la actualidad, resulta indiscutible la amplia adopción de prácticas ágiles en la industria del software; de acuerdo al Informe sobre el Estado de la Agilidad 2021, más de la mitad de los encuestados (52%) afirma que la mayoría o todos los equipos de su empresa han adoptado agilidad [5].

Sin embargo, a pesar de ello y de las ventajas obtenidas a partir de su aplicación, principalmente en lo relativo a la adaptación a los cambios, continúan siendo insuficientes los aspectos relacionados a modelos de calidad tanto de proceso como de producto destinados específicamente a enfoques ágiles de desarrollo de software.

En este sentido, como resultado de proyectos anteriores, se desarrolló AQF, integrando, por un lado, un modelo de calidad de software para procesos ágiles de desarrollo, denominado QuAM que favorece la medición y seguimiento de diversos aspectos. Y, por otro lado, una plataforma tecnológica que, a través de una herramienta de software denominada QuAGI automatiza la gestión de dicho modelo y permite la obtención, en simultáneo, de los niveles de calidad asociados al proceso de desarrollo de software guiado por prácticas ágiles.

Actualmente, el modelo QuAM, está formado por un conjunto de cuatro componentes que influyen en la calidad de proyectos ágiles junto a un procedimiento de evaluación de calidad que los integra y permite asociar el proceso evaluado con un nivel de calidad alcanzado. El Componente N° 4 “Gestión de Requerimientos”, tiene por objetivo mejorar la aplicación de la práctica ágil de definir necesidades como historias de usuario, promoviendo la importancia de contar con requerimientos bien definidos y completos, que contribuyan a la realización del proyecto.

El Componente N° 4 de QUAM se enfoca en la evaluación de calidad respecto a la gestión de requerimientos y requisitos del proyecto ágil, respondiendo al desafío implícito de procesos que involucren la participación del cliente y la rápida reacción al dinamismo de requerimientos y entregas continuas del producto [6].

El Manifiesto Ágil establece como prioridad la satisfacción del cliente mediante la entrega temprana y continua de software con valor, considerando que los requerimientos pueden cambiar en cualquier momento del desarrollo. Esto es posible, según expresan Serna y Suaza Jiménez [7], con una correcta documentación elaborada a partir de una gestión adecuada de requerimientos y requisitos, favoreciendo la entrega temprana del software requerido con un nivel de calidad específico.

Sin embargo, resulta necesario establecer claramente la diferencia entre las Historias de Usuario (User Stories – US) y las tradicionales especificaciones de requisitos, sin que ello implique una contraposición entre ambas; la

realidad de los proyectos y de los negocios conlleva ajustes entre los postulados teóricos y la práctica, esto es, que la teoría ágil debe adaptarse a cada caso en particular, obteniendo la verdadera riqueza y productividad de las múltiples soluciones que ofrece la ingeniería del software [8].

La aplicación de técnicas de ingeniería de requerimientos (IR) a procesos ágiles de desarrollo de software es un campo relativamente reciente, que aún no está completamente explorado y comprendido; un mapeo sistemático sobre esta área temática identificó la necesidad de realizar investigaciones sobre la elicitación de requisitos, la gestión del cambio, la medición de requisitos, las herramientas de gestión de requisitos y los estudios comparativos entre requisitos tradicionales y ágiles [9].

Las actividades de la IR tradicional se realizan durante la fase de análisis del ciclo de vida de desarrollo de software; su adaptación al enfoque ágil no es directo. La IR en el desarrollo ágil es informal y se basa en las habilidades y conocimientos de los individuos [10]; su caracterización no es clara, no sólo en ámbitos de investigación sino también para la comunidad de desarrollo de software, que no reconoce aún la aplicación de prácticas de IR en procesos ágiles [6].

Así pues, si bien existen propuestas que relacionan cuestiones asociadas a calidad y agilidad [11][12], son pocos los estudios que se enfocan en la evaluación del proceso de obtención de requerimientos [13] y ninguno de ellos se integra en un framework para la evaluación de la calidad de procesos ágiles, como aquí se propone. No se halló evidencia de iniciativas orientadas a enriquecer el proceso de desarrollo de software guiado por prácticas ágiles mediante la inclusión de métodos y técnicas de gestión de requerimientos, con el fin de mejorar su calidad.

Además, resulta necesario que la evaluación se realice desde el punto de vista de las prácticas ágiles, de modo tal de recomendar mejoras tanto a nivel del modelo de calidad definido (QuAM) como del proceso de desarrollo de software en sí

mismo, contribuyendo a incrementar la calidad del software final.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este Proyecto surge a partir de la necesidad detectada en experiencias de validación realizadas en proyectos anteriores, donde los estudios de casos reales permitieron identificar que la gestión de requerimientos y requisitos resulta ser un factor cuyo impacto sobre la calidad de software debe evaluarse.

A través de este Proyecto se pretende analizar y evaluar la aplicabilidad de técnicas y prácticas de la Ingeniería de Requisitos a proyectos ágiles de Software, a fin de enriquecer el proceso de desarrollo, aportando valor al producto construido, y en definitiva mejorar la calidad tanto del proceso como del producto de software. Para iniciar se ha realizado una revisión somera de la literatura existente, que relacione cuestiones específicas de la ingeniería de requisitos con prácticas ágiles, poniendo énfasis en la evaluación y mejora de la calidad, tanto desde un punto de vista teórico como su aplicación en casos reales; como resultado de este relevamiento no se encontró evidencia de la existencia de métodos, técnicas o herramientas que permitan evaluar la calidad de requisitos en proceso ágiles de desarrollo.

En consecuencia, resulta de importancia desarrollar tales métodos y herramientas, a fin de colaborar con la mejora continua del proceso ágil de desarrollo, y finalmente ampliar el modelo QuAM, que si bien ya contempla la Gestión de Requerimientos en su Componente N°4, resulta susceptible de optimización en este aspecto.

En vista de lo expuesto, se ha definido como objetivos de esta investigación contribuir a la mejora de la calidad de procesos ágiles de desarrollo de software, y de los productos de software que estos generan, mediante el desarrollo de métodos, técnicas y herramientas para evaluar y mejorar la calidad de requisitos en proyectos ágiles de software.

A fin de cumplimentar tal objetivo se llevarán a cabo las siguientes actividades en el marco de este proyecto:

- ☒ Ejecución de una revisión sistemática de bibliografía sobre métodos, técnicas y herramientas existentes para evaluar la calidad de requisitos de procesos ágiles de desarrollo.
- ☒ Ampliación del modelo QuAM, desarrollando métodos y técnicas para la gestión y evaluación de calidad de requisitos de software.
- ☒ Implementación de una herramienta de software que dé soporte a la nueva configuración del modelo QuAM.
- ☒ Validación de la propuesta

3. RESULTADOS OBTENIDOS/ ESPERADOS

Este proyecto de investigación se centra en la mejora de la calidad del proceso ágil de desarrollo de software, con particular énfasis en la gestión de requerimientos, buscando nutrirse de técnicas clásicas y altamente probadas en la IR clásica, extrapolándolas y adaptándolas a entornos ágiles. Se espera obtener como resultado final, la ampliación del modelo QuAM, desarrollando métodos y técnicas para la gestión de requerimientos y la evaluación de su calidad, e implementar una herramienta de software que dé soporte a la nueva configuración del modelo QuAM.

Se pretende que las técnicas y herramientas desarrolladas contribuyan al mejoramiento de la calidad de procesos de software en la región nordeste de Argentina (NEA), generando experiencias de vinculación entre la Facultad Regional Resistencia de la UTN y empresas de la industria del software del NEA, nutriendo la calidad de los resultados que se obtengan al mismo tiempo que se amplían las posibilidades de transferencia, logrando un impacto positivo en el medio.

En este sentido, existe una larga historia de vinculación y colaboración con empresas

pertenecientes al Polo IT Chaco y Polo IT Corrientes, que en forma permanente muestran su interés y predisposición tanto para la provisión de casos de estudio como para la validación de las propuestas y resultados surgidos del proyecto.

Para la evaluación de la mejora en la calidad del proceso resultante de la utilización del nuevo método y/o técnica de gestión de requerimientos desarrollado será necesaria la participación activa de capital humano y conocimiento de las empresas del sector de software y servicios informáticos, como así también la vinculación con otras organizaciones tales como el Instituto Chaqueño de Ciencia, Tecnología e Innovación (ICCTI).

A nivel académico, este proyecto supone la introducción de una nueva perspectiva para el estudio combinado entre enfoques ágiles, calidad de software e IR, impulsando la aparición de nuevos proyectos de investigación en los que docentes, estudiantes y profesionales colaboren en el análisis de casos de estudio tendientes a mejorar la competitividad de las empresas de la industria del software en nuestra región.

4. FORMACIÓN DE RECURSOS HUMANOS

Es posible distinguir aportes del proyecto a la formación de recursos humanos en dos sentidos:

1) *Formación de estudiantes y transferencia de resultados de investigación al aula.* Las actividades y resultados de investigación se vinculan en forma directa con la carrera Ingeniería en Sistemas de Información (ISI) de la UTN-FRRe, y su interacción con estudiantes en el aula contribuyen a su formación e incentivan las vocaciones científicas. Se destaca la relación con las asignaturas electivas del 4to nivel de la carrera: (i) Calidad del Producto y Proceso de Software y (ii) Técnicas de Desarrollo de Software Ágiles. Los resultados obtenidos durante el proyecto serán también compartidos con la asignatura Ingeniería de Software, curricular del 4to nivel, donde los

estudiantes podrán aplicar de forma empírica los conocimientos aprendidos durante el curso. Por otra parte, resulta destacable que el nuevo Diseño curricular de la carrera contempla el cambio en la denominación de esta asignatura a “Ingeniería y Calidad de Software”, teniendo en cuenta nuevas Competencias Específicas establecidas para profesionales de Ingeniería en Sistemas de Información: “Establecer métricas y normas de calidad de software”. En este sentido, el aporte del proyecto aumenta su relevancia en relación con posibilidades de inserción de alumnos y jóvenes graduados en el equipo, fomentando la postulación a becas de grado y posgrado de la UTN.

Cabe aclarar que todas estas asignaturas están a cargo de docentes-investigadores pertenecientes al CInApTIC, La participación de estudiantes becarios y becarias de la carrera de ISI posibilitará la formación de Recursos Humanos en esta área de vacancia regional, tanto a nivel científico como profesional.

2) *Formación de recursos humanos en investigación.* El proyecto cuenta actualmente con dos vacantes para estudiantes con Becas de Investigación y Servicios financiada por la Secretaría de Asuntos Universitarios (UTN), una plaza para Becario Alumno de Rectorado (BAR) y una plaza de Beca de Iniciación a la Investigación y Desarrollo (BINID) para graduados financiadas por la Secretaría de Ciencia y Tecnología (UTN). Todas ellas actualmente en proceso de selección de postulantes, con el objetivo de incorporar y formar jóvenes investigadores (estudiantes, egresados y/o egresadas de ISI) en la temática que el proyecto descripto aborda.

En cuanto a la formación de Posgrado de los docentes investigadores que participan en el proyecto, el equipo de trabajo de esta línea de investigación del CInApTIC está integrado por su Directora, Doctora en Ciencias Informáticas por la Universidad Nacional de La Plata, y Co Directora, ambas con categorías en el Programa de Incentivos y como Docentes Investigadores

de UTN. Además, forman parte del equipo dos Docentes Investigadores, uno de ellos Máster en Tecnologías de la Información y Sistemas Informáticos y de Doctor en Informática y Modelización Matemática (2007), ambos por la Escuela Técnica Superior de Ingeniería Informática de la Universidad Rey Juan Carlos (Madrid - España), y un Ingeniero en Sistemas de Información, quien juntamente con la Co Directora actualmente son tesis de Doctorado en Informática dictado en conjunto entre la UTN-FRRe, la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones, y la Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional del Nordeste. En particular, se encuentra en sus inicios la realización de la tesis doctoral de la Esp. Gabriela Tomaselli, cuyo tema está estrechamente relacionado con la temática del Proyecto y, a partir de actividades compartidas, la tesis del Ing. Nicolás Tortosa, la cual se enmarca en un proyecto relacionado.

5. BIBLIOGRAFÍA

- [1] Pinto, N., Tortosa, N., Geat, B. C., Ibáñez, L., & Bollati, V. (2018). Quality evaluation of agile processes: Measurement of requirements management using AQF v2, *11th International Conference on the Quality of Information and Communications Technology (QUATIC 2018)* (pp. 15-20). IEEE.
- [2] Pinto, N., Acuña, C., & Cuenca Pletsch, L. R. (2016). Quality Evaluation in Agile Process: A First Approach, *XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)* (pp. 525-534).
- [3] Pinto, N., Tomaselli, G., Acuña, C., & Cuenca Pletsch, L. (2017). QuAGI: Una propuesta para el seguimiento y evaluación de proyectos de Software Ágiles, *V Seminário Argentina-Brasil de Tecnologias da Informação e da Comunicação (SABTIC 2017)* (pp. 127-136).
- [4] Pinto, N. S., Tortosa, N., Cabas Geat, B., Ibáñez, L., & Acuña, C. J. (2019). Validación de la reingeniería aplicada sobre la primera versión de Agile Quality Framework. *Electronic Journal of SADIO*, 18(1), 93-109.
- [5] digital.ai. (2021). *15th Annual State of Agile Report*. <https://digital.ai/resource-center/analyst-reports/state-of-agile-report>
- [6] Inayat, I., Salim, S. S., Marczak, S., Daneva, M., & Shamshirband, S. (2015). A systematic literature review on agile requirements engineering practices and challenges. *Computers in Human Behavior*, 51(B), 915-929.
- [7] Serna M., E., & Suaza Jiménez, J. H. (2016). Documentar la elicitación de requisitos: Una revisión sistemática. *Ingeniare. Revista chilena de ingeniería*, 24(4), 703-714.
- [8] Garzas, J. (2011). *La historia de usuario no es el "requisito" de las metodologías ágiles*. <https://www.javiergarzas.com/2011/12/historia-de-usuario-diferente-de-requisito.html>
- [9] Curcio, K., Navarro, T., Malucelli, A., & Reinehr, S. (2018). Requirements engineering: A systematic mapping study in agile software development. *Journal of Systems and Software*, 139, 32-50.
- [10] Dingsøyr, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software*, 85(6), 1213-1221.
- [11] Matalonga, S., & Rivedieu, G. (2015). AGIS: hacia una herramienta basada en ISO9001 para la medición de procesos ágiles. *Computación y Sistemas*, 19(1), 163-175.
- [12] Pasini, A. C., Esponda, S., Boracchia, M., & Pesado, P. M. (2013). Q-Scrum: una fusión de Scrum y el estándar ISO/IEC 29110, *XVIII Congreso Argentino de Ciencias de la Computación (CACIC 2012)* (pp. 898-909).
- [13] Lucassen, G., Dalpiaz, F., van der Werf, J. M. E., & Brinkkemper, S. (2016). Improving agile requirements: the quality user story framework and tool. *Requirements Engineering*, 21(3), 383-403.

Métodos, técnicas y herramientas para mejorar y evaluar la calidad de la capacidad de producción de entregables dentro de un proceso ágil de desarrollo de software

César J Acuña; Nicolás Tortosa; Gabriela Tomaselli; Noelia Pinto
Centro de Investigación Aplicada a TIC (CInApTIC)
Universidad Tecnológica Nacional, Facultad Regional Resistencia
French 414, Resistencia, Chaco
{csr.acn;nicotortosa;gabriela.tomaselli;ns.pnto}@gmail.com

RESUMEN

Para mejorar los procesos de desarrollo de software en las organizaciones, y facilitar la adopción de prácticas ágiles que aseguren la calidad de estos, se ha desarrollado en proyectos anteriores, Agile Quality Framework (AQF), una propuesta que integra un modelo de calidad (QuAM) junto a una aplicación web (QuAGI) que permite la automatización de dicho modelo. Actualmente, el modelo QuAM, está formado por 4 componentes que influyen en la calidad de proyectos ágiles junto a un procedimiento de evaluación de calidad que los integra.

Luego de diversas experiencias de validación llevadas a cabo, en este trabajo se presenta la línea de investigación que busca continuar mejorando y ampliando el framework AQF, de forma tal de profundizar en la evaluación de calidad de procesos ágiles explorando aspectos específicos de sus componentes.

De hecho, esta línea de investigación buscará optimizar el framework AQF a partir de nuevos métodos, técnicas y herramientas que contribuyan a mejorar el proceso de estimación, aspecto que forma parte del Componente N° 1 de AQF, "Capacidad de Producción de Entregables", cuya función es la evaluación de factores que influyen en la obtención de entregables de valor para el cliente.

Palabras clave: Calidad del Software, Procesos ágiles de Software, Estimación ágil

CONTEXTO

El trabajo que aquí se presenta está enmarcado en el proyecto "Métodos, técnicas y herramientas para mejorar y evaluar la calidad de la capacidad de producción de entregables dentro de un proceso ágil de desarrollo de software", que es financiado por la UTN y ejecutado en el Centro de Investigación Aplicada a TIC (CInApTIC) de la Facultad Regional Resistencia, con el código SIECRE0008608TC.

Asimismo, algunas actividades son compartidas con el proyecto de investigación y desarrollo "Métodos, técnicas y herramientas para mejorar y evaluar la calidad de requisitos en proyectos ágiles de Software." (SIECRE0008643), también financiado por UTN y ejecutado en el CInApTIC.

1. INTRODUCCIÓN

La mejora e innovación de los procesos de software, con el objetivo de incrementar la calidad de sus productos y servicios, se ha convertido, en los últimos años, en el elemento diferenciador que las empresas necesitan para mejorar sus

niveles de competitividad en la Industria del Software.

Sin embargo, diversos estudios coinciden en la dificultad de las empresas de desarrollo de la región [1][2][3] para implementar programas de Mejoras de Proceso de Software (Software Process Improvement - SPI), fundamentalmente porque la aplicación de estos modelos resulta costosa en términos económicos y de esfuerzo, pues requieren una gran inversión en dinero, tiempo y recursos, sus recomendaciones son complejas de aplicar y el retorno de la inversión se produce a muy largo plazo [4]. De esta forma, los parámetros de tiempos de desarrollo y costo de soluciones afectarán directamente al trabajo que se realice, siendo la calidad la primera variable de ajuste disponible.

Debido a esto, en la búsqueda por lograr mejores procesos de desarrollo para incrementar la calidad de sus productos e incrementar su nivel de competitividad en escenarios actuales, durante los últimos años la Industria del Software ha iniciado la adopción de prácticas ágiles en sus proyectos enfocadas en ciclos de desarrollos cortos, iterativos e incrementales, con equipos pequeños y auto-organizados, diseños simples con refactorización de código, donde el desarrollo conducido por la prueba es práctica habitual y la participación de los clientes en forma frecuente permite presentar la evolución del producto en cada ciclo de desarrollo [5].

Por ello, surge la necesidad de elaborar estrategias que permitan no solo guiar a las empresas en la adopción de las prácticas ágiles sino también en cuestiones relacionadas a la gestión de la calidad que contribuyan en la generación de valor en la producción de software. En este sentido, y como primera alternativa, con anterioridad, el equipo que integra la línea de investigación que aquí se presenta ha desarrollado un framework que permite

evaluar la calidad cuando se opta por trabajar con procesos ágiles de desarrollo de software. Dicho framework se denomina AQF (Agile Quality Framework) y su versión actual está compuesta por un modelo, QuAM (Quality Agile Model) y por una herramienta de software que brinda soporte a dicho modelo, QuAGI (Quality AGile) [6][7].

El framework ha sido validado en diversas experiencias de vinculación con empresas y áreas de Sistemas pertenecientes a la Industria del Software en el NEA, y en base a los resultados obtenidos se planteó la necesidad de adecuar los componentes actuales de AQF. Uno de ellos es el N° 1 “Capacidad de producción de entregables”, el cual, basándose en los principios del Manifiesto Ágil y atendiendo la necesidad de las empresas de dar respuesta al problema de la “no entrega continua”, busca mejorar la calidad del proceso a través de la estimación óptima de los requerimientos, enfatizando prácticas que permitan realizar entregas continuas, funcionales y de valor para el cliente.

De acuerdo a los resultados de validación, sobre proyectos reales de software, en muchos casos los equipos participantes han planteado que la estimación resulta ser uno de los obstáculos más recurrentes para lograr la continuidad de la obtención de entregables y, por tanto, también constituye, un obstáculo para el éxito en la implementación del producto final. Una mala gestión respecto a: la entrega del producto en el tiempo estipulado, el uso eficiente del presupuesto asignado y el cumplimiento de las funcionalidades requeridas, inciden sobre el proyecto impactando de manera negativa sobre los niveles de calidad final asociados al ciclo de desarrollo del software y derivan, en

nuevos problemas que el equipo de trabajo debe enfrentar.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A partir de diversas experiencias de validación llevadas a cabo, con este proyecto se busca continuar mejorando y ampliando el framework AQF, de forma tal de profundizar en la evaluación de calidad de procesos ágiles explorando sus componentes de forma más detallada y a su vez, proponer nuevas técnicas y métodos que mejoren aspectos específicos del proceso ágil de desarrollo. Surge, entonces, la propuesta de esta primera aproximación de optimización de AQF que propone enfocarse en el Componente N° 1 de AQF, "Capacidad de Producción de Entregables", cuya función es la evaluación de factores que influyen en la obtención de entregables de valor para el cliente. Como, es sabido, el desarrollo de software está sujeto a restricciones técnicas y económicas, aspectos que forman parte del plan del proyecto asociado. Dicho plan abarca la distribución de actividades y tareas entre integrantes del equipo, lo cual surge en función de los requerimientos definidos. Estas actividades y tareas tienen un esfuerzo requerido para su realización, estimado a priori con algún método conocido. Y, si bien cada vez es mayor el número de aportes en la creación de modelos para la estimación del esfuerzo de desarrollo de software, lo cual demuestra un interés creciente de la comunidad científica en estos temas, aún no existen propuestas que se integren dentro de un framework completo como es el caso de AQF. Por ello, el proyecto que aquí se presenta buscará optimizar el framework AQF a partir de nuevos métodos y técnicas que contribuyan a mejorar la capacidad de producción de entregables y que, a su vez, permita evaluar dicho componente como

parte de la evaluación de la calidad del proceso ágil.

Las actividades que se llevarán a cabo en el marco de esta línea de investigación son:

- ☒ Revisión sistemática que permita definir el marco teórico, identificar trabajos relacionados y delimitar el alcance
- ☒ Estudio comparativo entre las técnicas que utilizan las organizaciones para la estimación de esfuerzo, tamaño funcional y costo de proyectos ágiles.
- ☒ Definición de métodos y técnicas para la gestión y evaluación de calidad en la estimación de esfuerzo, tamaño funcional y costo de proyectos ágiles.
- ☒ Desarrollo de aplicaciones de software para gestionar los elementos de las técnicas y métodos diseñados.
- ☒ Validación de la propuesta mediante estudio de casos en escenarios reales de producción.

3. RESULTADOS OBTENIDOS/ ESPERADOS

El trabajo que aquí se expone, surge a partir de la necesidad detectada respecto a la optimización en el proceso de estimación, lo cual resulta ser un factor cuyo impacto sobre la calidad de software impacta significativamente. Particularmente, en el caso del Componente N°1 de QUAM "Capacidad de Producción de Entregables", originalmente consideraba cuestiones técnicas, tales como el uso de herramientas de gestión de cambios y testing automatizado, como factores que favorecen la producción de entregables. Sin embargo, y luego de experiencias de validación, se modificó su estructura en base a la observación de resultados

obtenidos y propuesta de diversos equipos de proyectos ágiles evaluados considerando cuestiones asociadas a los procesos de estimación de esfuerzo.

En su configuración actual, AQF considera positivo implementar estrategias que mejoren la calidad de la estimación en los proyectos ágiles, pues esto minimizará desfases en la obtención de entregables. Las acciones tendientes a lograr más precisión en la estimación de la entrega de artefactos en cada iteración, en detrimento de una visibilidad a largo plazo, aumentarán el nivel de calidad de la Producción de Entregables en proyectos ágiles.

Por tanto, y para iniciar con esta propuesta, el equipo de trabajo se encuentra realizando el relevamiento de la situación actual respecto a métodos y modelos de estimación utilizados en proyectos de software desarrollados bajo enfoques ágiles. El relevamiento ha abarcado no solo procesos de estimación de esfuerzo sino también de costo y de tamaño funcional, ampliando el enfoque cuando se trata de procesos ágiles de desarrollo de software.

Esta información preliminar, permitirá iniciar el diseño y desarrollo de nuevos métodos, técnicas y herramientas que se integren al framework AQF, favoreciendo los aspectos que ayuden a incrementar la calidad del proceso ágil de desarrollo asociado a cada proyecto de software, optimizando el proceso de estimación ágil.

Este aporte beneficiará directamente a las instituciones y/o empresas, ya sean públicas o privadas que adopten AQF como Framework en sus procesos ágiles de desarrollo de software, ofreciendo nuevas alternativas que mejoren la capacidad de producción de entregables, impactando de forma positiva en los niveles de calidad asociados al proyecto en su conjunto. producción de

entregables, impactando de forma positiva en los niveles de calidad asociados al proyecto en su conjunto.

Asimismo, el afianzamiento de la línea de investigación en la temática del proyecto, cuyos integrantes son docentes de grado y posgrado, permitirá que los conocimientos desarrollados sean transferidos en diversas cátedras de Ingeniería en Sistemas de Información y carreras de posgrado afines. En particular, desde lo disciplinar, el proyecto permitirá generar nuevo conocimiento sobre métodos y técnicas que optimicen procesos de estimación en proyectos ágiles de software, integrando todos los componentes en un framework existente que contribuye a la mejora de la calidad de procesos de desarrollo de software.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación del CInApTIC está integrado por su Director, con título de Doctor con categoría en el Programa de Incentivos, y Co Director, tesista doctoral, ambos Docentes Investigadores de UTN. Además, forman parte del equipo dos Docentes Investigadores (Profesionales de Ingeniería en Sistemas de Información- ISI), una de ellas Doctora en Ciencias Informáticas y la otra tesista doctoral. Resulta importante destacar que el equipo no sólo incluye investigadores formados, sino que pretende incorporar y formar jóvenes investigadores (estudiantes, egresados y/o egresadas de ISI) en la temática que el proyecto descripto aborda.

La participación de personas con becas, tanto estudiantes como profesionales de la carrera de ISI, posibilitará la formación de Recursos Humanos en esta área de vacancia regional, tanto a nivel científico como profesional. Además, incentivará el intercambio con otras líneas de investigación a nivel local y con otros

centros de investigación del país y del exterior, permitiendo una mayor difusión de los conocimientos producidos y el fortalecimiento en la formación de los recursos humanos intervinientes. Sumado a esto, la participación de investigadores en este tipo de proyectos redundará en un fortalecimiento de las capacidades institucionales en investigación y desarrollo. El proyecto prevé la participación de dos estudiantes de grado (con Becas de Investigación financiadas por UTN) y un profesional con becas destinadas a graduados y graduadas de ISI.

Con respecto a la formación de doctores, en el marco de este proyecto se desarrollará la tesis del Ing. Nicolas Tortosa (Codirector de este proyecto), y, con actividades compartidas, la tesis de la Esp. Gabriela Tomaselli (Investigadora de este proyecto) que se enmarca en un proyecto relacionado. Además, el proyecto admitiría la realización de, al menos, una tesis doctoral adicional por parte de estudiantes del programa de Doctorado en Informática realizado en conjunto por la UNaM, UNNE y UTN, y cuya cohorte 2020 se encuentra en etapa de definición de tema de tesis.

REFERENCIAS

[1] Mas A., Amengual E. (2005). "Las mejoras de los procesos de Software en las pequeñas y medianas empresas (pymes). Un nuevo modelo y su aplicación a un caso real". Revista Española de Innovación, Calidad e Ingeniería del Software, Vol.1, No. 2

[2] Pasini, A. C., Esponda, S., Bertone, R. A., & Pesado, P. (2008). "Aseguramiento de Calidad en PYMES que desarrollan software." XIV Congreso Argentino de Ciencias de la Computación.

[3] Pflegger, S. (2002) "Ingeniería de Software. Teoría y Práctica." Pearson Education.

[4] Javier Garzás, Carlos Manuel Fernández & Mario Piattini (2009). "Una aplicación de la Norma ISO/IEC 15504 para la evaluación por nivel de madurez de Pymes y pequeños equipos de desarrollo". Revista Española de Innovación, Calidad e Ingeniería del Software, Vol.5, No. 2.

[5] Swebok, [Online] Available: <http://www.computer.org/portal/web/swebok>

[6] Pinto, N. S. (2020). Framework para la evaluación de calidad de proyectos ágiles de software (Doctoral dissertation, Universidad Nacional de La Plata).

[7] Pinto, N. S., Tortosa, N., Cabas Geat, B., Ibáñez, L., & Acuña, C. J. (2019). Validación de la reingeniería aplicada sobre la primera versión de Agile Quality Framework. Electronic Journal of SADIO, 18.

CREACIÓN DE UN OBSERVATORIO TECNOLÓGICO DE LA GESTIÓN DEL CONOCIMIENTO EN LA CIUDAD DE BUENOS AIRES

Straccia, Luciano; Zanitti, Ma. Ayelén; Pollo-Cattaneo, María F.

Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
Departamento de Sistemas de Información. Grupo GEMIS.

lstraccia@frba.utn.edu.ar, maria.ayelen.zanitti@gmail.com, flo.pollo@gmail.com

RESUMEN

La gestión del conocimiento (GC) aborda la creación y el uso del conocimiento en las organizaciones desde un enfoque multidisciplinar. Un observatorio tecnológico busca operar adecuadamente la información masiva y se basa en la colaboración y el uso de fuentes de información de diversa índole, se alimenta habitualmente con información de una comunidad afín y después de su procesamiento y análisis, los resultados se ponen a disposición de la propia comunidad. A partir de la experiencia del grupo de investigación en GC y considerando que en la Ciudad de Buenos Aires no existen observatorios conocidos para la temática se inició el proyecto "Creación de un observatorio tecnológico de la gestión del conocimiento en la Ciudad de Buenos Aires", cuyo objetivo es describir, analizar y presentar la situación actual de la GC en la ciudad de referencia, a través de la construcción de un observatorio tecnológico.

Palabras clave: conocimiento, gestión del conocimiento, observatorio tecnológico.

CONTEXTO

La Universidad Tecnológica Nacional (UTN) posee un Programa de Investigación, Desarrollo e Innovación, de Sistemas de Información e Informática. En la Facultad Regional Buenos Aires (UTN.BA), en el marco del

Departamento de Ingeniería en Sistemas de Información, se ha conformado en el año 2009 el Grupo GEMIS, que lleva adelante proyectos en el campo de la Ingeniería en Sistemas de Información, Ingeniería de Software e Ingeniería del Conocimiento, sus aplicaciones y sus abordajes metodológicos.

Desde el año 2018 el grupo lleva adelante proyectos de I+D vinculados a la GC, iniciando con el proyecto "La gestión del conocimiento en pequeñas y medianas fábricas de software en el Área Metropolitana de Buenos Aires" (2018-2020) y dándole continuidad con el proyecto "Arquitectura tecnológica para la gestión del conocimiento" (2021-2023), contando con trabajos finales de especialización y tesis de maestría. A partir de las experiencias y resultados obtenidos en los proyectos mencionados surgen las preguntas que dan lugar al proyecto objeto de este trabajo.

1. INTRODUCCIÓN

La gestión del conocimiento (GC) aborda la creación y uso del conocimiento en las organizaciones. Es un campo integrado que se alimenta de múltiples disciplinas que permiten desarrollar iniciativas en diversos ámbitos y en diferentes niveles dentro de las organizaciones [1] con un enfoque multidisciplinario orientado a una visión completa y sistemática [2].

El conocimiento, desde un enfoque constructivista, forma parte de la jerarquía DIKW [3], constituida por los datos en el nivel

inferior, la información en un siguiente nivel y el conocimiento en el tercer nivel; finalmente, la sabiduría constituye el nivel superior. El conocimiento es "la mezcla de creencias cognitivas (..), perspectivas, juicios, metodologías, (...) experiencias y expectativas realizadas sobre un objeto, que son adaptadas y potenciadas por la mente de un individuo (conocedor)" [4].

La información (I) puede definirse como una función de los datos [5;6] por contener tanto los datos (d) como su contexto (C_d), como se especifica en la ecuación: $I = f(d) = d + C_d$

Considerando además que se utiliza "insight" (i) para representar las implicaciones tácitas vinculadas a la información, el conocimiento (C) puede verse en la ecuación: $C = p(I) = I + C_I + i$.

Vistas de la gestión del conocimiento

Una vista describe los conceptos, elementos y características de un sistema integrado desde la perspectiva de un conjunto de preocupaciones relacionadas. Existen cinco vistas de la GC: aspectos organizativos (estructura, cultura, etc.), personas (roles, responsabilidades, etc.), proceso (y actividades), tecnologías y medición [7;8].

Actividades en la gestión del conocimiento

Existen diversos modelos de GC, entre otros, el espiral del conocimiento de Nonaka & Takeuchi [9], modelo de infraestructura de Kerschberg [10], Integrado situacional de Riesco [11], de gestión tecnológica de Paniagua & Lopez [12] y el modelo holístico de Angulo & Negrón [13].

Para poder llevar adelante un proceso integral de GC es necesario: a) identificar los procesos claves; b) identificar los conocimientos claves para esos procesos; c) identificar fuentes internas y externas que pueden hacer aportes a la transmisión de dicho conocimiento; d) establecer mecanismos de transferencia; e) establecer las actividades y roles que aseguren

el cumplimiento de esos mecanismos; y f) establecer hitos que permitan observar periódicamente la aplicación de los aspectos anteriores [14].

Entre las actividades para la ejecución de un sistema de GC se encuentran: identificación de las necesidades de conocimiento, creación del conocimiento, adquisición, compartición, aplicación, almacenamiento, refinamiento, validación [15].

Tecnologías y representación del conocimiento

En lo referido a las tecnologías para la GC se considera relevante identificar cuáles pueden contribuir al flujo de conocimiento, apoyando su conversión desde explícito a tácito y desde tácito a explícito, qué tipos de tecnologías pueden apoyar al conocimiento explícito que la una organización posee y cómo gestionar, a través de las TIC, el volumen de conocimiento explícito contenido en las colecciones de documentos de una organización. Tyndale [16] distingue entre tecnologías de la información basadas en herramientas tomadas desde otras disciplinas que han entrado en el campo de la GC y tecnologías de la información basadas en herramientas específicas del campo disciplinar. El conocimiento debe ser representado de una manera que permita a los sistemas de información procesarlo activamente [17] y relacionar los elementos de conocimiento entre sí [18]. Las tecnologías y mecanismos de representación pueden ser incluidos en alguna de las siguientes categorías: técnicas de socialización, técnicas o modelos para la explicitación y representación del conocimiento, campos disciplinares, procesos lógicos y analíticos, herramientas tecnológicas y prácticas organizacionales [19].

Evaluación

En tanto, los modelos de evaluación de la GC son abordados a partir de dos visiones diferentes: aquellas que definen indicadores y aquellas que precisan modelos de madurez. La

primera se basa en la definición de la gestión del conocimiento como un activo intangible y con capacidad de ser definido con un valor cuantificable incluido en el valor de mercado junto a otros activos; sus modelos no abordan de forma holística la evaluación y miden mayormente el impacto de uso del conocimiento y no su proceso integral. A partir de ello se diseñaron modelos de madurez, destacándose los de Pee y Kankanhalli [20], Chu [21] y De Freitas [22].

Observatorios tecnológicos

El término observatorio en el campo social proviene de la década de 1960 cuando Wood propuso que las políticas urbanas sean tratadas como un fenómeno científico y sometidas a la observación [23]. Un observatorio es un organismo creado por un colectivo, con el fin de seguir la evolución de un fenómeno [24], normalmente de carácter social y que tiene como propósito "investigar (...) los contenidos que aparecen en el espacio de observación pertinente a su área de interés; e informar a la comunidad los hallazgos" [23].

Los productos de un observatorio pueden servir, al menos, para: caracterizar una situación o momento; apoyar la toma de decisiones coyunturales y formular escenarios a futuro [25]. Un observatorio aspira "a convertirse en un centro de seguimiento y en un punto de referencia para la investigación social, científica, técnica y de acción" [26]

El concepto de observatorio tecnológico [27] busca dar respuesta al problema de operar adecuadamente la información masiva y se basa en la colaboración en línea y el uso de fuentes de información de diversa índole, que se alimenta habitualmente con información de una comunidad afin; y después de su procesamiento y análisis, los resultados se ponen a disposición de la propia comunidad [28]. Estos observatorios "deben ser lugares de encuentro activos e inclusivos para generar reflexiones. Esto refuerza la capacidad de acción mediante la difusión de estudios e informes y conlleva a

la creación de un centro de documentación abierto (...) y se crean plataformas interactivas que recogen y difunden información" [29].

Existen definiciones diversas sobre el concepto de observatorio tecnológico, desde la que afirma que es un "espacio multidimensional constituido por redes temáticas colaborativas que, de forma periódica y sistemática" [25] hasta las más acotadas como un espacio que "captura informaciones externas con el propósito de transformarlas en conocimientos específicos que conducen a sus usuarios a tomar decisiones" [30] o "lugar (físico o virtual) que permite una visión privilegiada de un campo u objeto de interés" [31].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A partir de los proyectos llevados a cabo por el grupo de investigación en los últimos años y considerando que en la Ciudad de Buenos Aires, República Argentina, no existen observatorios conocidos para la gestión del conocimiento surgen las siguientes preguntas:

a) ¿cuál es el nivel de implementación de la GC en las organizaciones en la ciudad?, b) ¿cuáles son los abordajes de la GC en las organizaciones?; y c) ¿el nivel de implementación de GC está asociado a variables que caracterizan a la organización, la industria o la región de dicha organización?

Estas preguntas dan lugar al proyecto "Creación de un observatorio tecnológico de la gestión del conocimiento en la Ciudad de Buenos Aires" con un plazo de trabajo de 48 meses.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo general del proyecto es describir, analizar y presentar la situación actual de la gestión del conocimiento en la Ciudad de Buenos Aires, República Argentina, a través de la construcción de un observatorio tecnológico.

Los objetivos específicos son: a) identificar organizaciones que poseen sistemas de GC o elementos asociados a un potencial sistema de GC; b) identificar las vistas abordadas en la GC en las empresas; c) identificar las tecnologías utilizadas en la GC en las empresas; d) identificar las actividades utilizadas; e) identificar los mecanismos de evaluación; f) obtener una caracterización general de la situación general de aplicación de GC en la Ciudad de Buenos Aires; y g) presentar los datos a través de un observatorio tecnológico.

En este proyecto se espera obtener como producto resultado una aplicación (web u con otra tecnología a determinar en el diseño del producto) como observatorio tecnológico de las organizaciones que poseen aplicación de gestión del conocimiento en la Ciudad de Buenos Aires, que permita navegar las diferentes comunas y que permita ver datos asociados a las características de dichas organizaciones.

Para la primera parte del proyecto se seguirá el proceso de diseño conceptual, diseño de la captación y diseño de la muestra [33], mientras que para la producción del software se realizará su diseño y desarrollo y pruebas con el desarrollo de un prototipado evolutivo experimental [34;35] que consiste en desarrollar una solución inicial para un determinado problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo se encuentra conformado por investigadores formados, investigadores de apoyo, y alumnos de grado. Esta línea de trabajo busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para su desarrollo en la carrera de investigadores, además de fomentar la aplicación de las temáticas en su propia actividad profesional. Se prevé también la

incorporación de becarios graduados a través de las becas BINID y se propone la integración del proyecto de investigación con tesis doctorales.

5. BIBLIOGRAFÍA

- [1] Wiig, K. (2007). Enterprise Knowledge Management.
- [2] Geisler E. y Wickramasinghe, N. (2015). Principles of Knowledge Management Theory, Practice, and Cases. New York, USA: Routledge
- [3] Ackoff, R. (1989). From data to wisdom. Journal of Applied Systems Analysis.
- [4] Díaz, M.; Millán, J. (2013). Gestión del Conocimiento y Capital Intelectual, a través de modelos universitarios. Económicas CUC.
- [5] Kaipa, P. 2000. Knowledge architecture for the twenty-first century. En Behaviour & Information Technology 19(3), 153–161.
- [6] Li, Z. (2020) On a Factorial Knowledge Architecture for Data Science-powered Software Engineering. En International Conference on Software and e-Business, Osaka, Japan.
- [7] Straccia, L., Ramacciotti, C., Pollo-Cattáneo, M.F. (2020) Una visión de la tecnología para la Gestión del Conocimiento. Resultados en la literatura latinoamericana. En Serna, E., Desarrollo e Innovación en Ingeniería, Instituto Antioqueño de Investigación.
- [8] Milton, N. (2015). The 4 legs on the Knowledge Management table.
- [9] Nonaka, I., y Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford University Press
- [10] Kerschberg, L. (2001). Knowledge management in heterogeneous data warehouse environments. En International Conference on Data Warehousing and Knowledge Discovery, Springer Berlin Heidelberg.
- [11] Riesco, M. (2010) El negocio es el conocimiento. Ediciones: Díaz de Santos. Madrid, España. 2010
- [12] Paniagua, E.; Lopez, B. (2007) La gestión tecnológica del conocimiento. España, Universidad de Murcia, 2007
- [13] Angulo, E. y Negrón, M. (2008). Modelo holístico para la gestión del conocimiento. Revista Negotium 4(11), 38 – 51

- [14] Straccia, L. (2022). La importancia de socializar el conocimiento en las empresas. *Diario La Gran Capital*, La Plata.
- [15] Straccia, L.; Maulini Buño, A.; Ramacciotti, C.; Pollo-Cattáneo, M.F. (2021). Fases propuestas para el diseño y construcción de un modelo de Gestión del Conocimiento. En Serna, E., *Desarrollo e Innovación en Ingeniería*, Instituto Antioqueño de Investigación.
- [16] Tyndale, P. (2002). A taxonomy of knowledge management tools: origins and applications. *Evaluation and Program Planning* 25 (183-190).
- [17] Portmann, E.; Kaltenrieder, P.; Pedrycz, W. (2015). Knowledge Representation through Graphs. *Procedia Computer Science*, 62, 245–248.
- [18] SUNY (2004) The Gartner Glossary of Information Technology Acronyms and Terms. SUNY Center for Professional Development.
- [19] Straccia, L.; Maulini, A.; Bongiorno, M.G.; Giorda, M.; Pollo-Cattaneo, M.F. (2022) Knowledge Representation and Technologies in the Latin American Academic Literature. En *Workshops at the International Conference on Applied Informatics*, Arequipa, Perú, October 27.
- [20] Pee, L.G. y Kankanhalli, A. (2009). A Model of Organizational Knowledge Management Maturity Based on People, Process and Technology. *Journal of Information & Knowledge Management*, 8(2), 79-99.
- [21] Chu, S.K.; Wu, W.; Chan, K.H. y Fu, O. (2011). The Relationship Between Knowledge Management and Intellectual Capital in Listed Companies of Mainland China. *Proceedings of the 8th International Conference on Intellectual Capital, Knowledge Management & Organisational Learning*, Bangkok, Thailand.
- [22] De Freitas, V. (2018). Modelo de madurez en sistemas de gestión del conocimiento desde un enfoque holístico. En *Negotium Revista Científica Electrónica de Ciencias Gerenciales*, 39(13), Fundación Unamuno, Venezuela.
- [23] Frausto, O.; Martínez, T.; González Matú, B. (2008). Observatorios e indicadores de violencia social y de género, *Revista Digital Universitaria*, 9(7), Universidad Nacional Autónoma de México.
- [24] Enjuto, N. (2008). Razón de ser de los observatorios. En *Jornada Observando observatorios: ¿nuevos agentes en el tercer sector?*
- [25] Angulo Marcial, N. (2009). ¿Qué son los observatorios y cuáles son sus funciones? En *Innovación Educativa* 9(47), abril-junio, 2009, pp. 5-17. Instituto Politécnico Nacional, Distrito Federal, México.
- [26] Álvarez, A. (2012). Formulación de un observatorio local de democracia participativa en San Cristóbal de la Laguna, Tenerife. *Cuadernos Geográficos* 50, 217-238.
- [27] Hernandez Cervantes, L.; Santillan, G.; Gonzalez-Ponce, A. (2009). Observatorios virtuales astrofísicos. En *Revista Digital Universitaria*, 10(10), Universidad Nacional Autónoma de México.
- [28] Castañeda de León, L. (2006). Observatorio virtual, más que un portal. En *Revista Enterate*, Universidad Autónoma de México, febrero 2006.
- [29] Moreno-Soler, G.; Mantilla Castellanos, J. (2016). Una revisión del concepto observatorio social: hacia una comprensión de sus objetivos, alcances, métodos y finalidades. En *Psicogente*, 19(36), Barranquilla.
- [30] Bouza Betancourt, O. (2010). Desarrollo del ámbito informacional desde la perspectiva de la sistematización de la Vigilancia Científica y Tecnológica (VCT) en organizaciones empresariales. Tesis de doctorado. Universidad de la Granada, Universidad de la Habana, 2-269.
- [31] Lazo Gonzalez, O. (2008). Propuesta teórico-metodológica de observatorios de políticas públicas en salud, USAID, Perú.
- [32] Morcela, O. (comp) (2016). Gestión del conocimiento en empresas de base tecnológica. Mar del Plata: Universidad Nacional de Mar del Plata.
- [33] INEGI (2021). Guía de diseño de la muestra para encuestas. Fase de diseño de la norma técnica del proceso de producción de información estadísticas y geográfica. Instituto Nacional de Estadística y Geografía, México.
- [34] Basili, V. (1993). The Experimental Paradigm in Software Engineering. En (Ed. Rombach, H., Basili, V., Selby, R., *Experimental Software Engineering Issues: Critical Assessment and Future Directions*. Lecture Notes in Computer Science, Vol. 706.
- [35] Ko, A. J., LaToza, T. D., & Burnett, M. M. (2015). A practical guide to controlled experiments of software engineering tools with human participants. *Empirical Software Engineering*, 20(1), 110-141.

Aprovisionamiento automático en la nube mediante la Ingeniería Dirigida por Modelos

Alberto Cortez^{1,2}, Carlos Martinez^{1,2}, Raúl Varela^{1,2},
Pablo Peña¹, Valentín Fernández¹, Claudia Naveda^{1,2}
Raúl Moralejo¹, Alejandro Vazquez¹.

UTN Facultad Regional Mendoza, Ingeniería en Sistemas de Información,

¹Laboratorio de Auditoría y Seguridad de TIC,

²Universidad de Mendoza, Instituto de Informática

RESUMEN

El despliegue y control de una aplicación compleja sobre un conjunto heterogéneo de proveedores es un problema al que los clientes de las plataformas de cloud se enfrentan diariamente. Los proveedores exponen sus servicios utilizando especificaciones independientes, incurriendo en una falta de portabilidad e interoperabilidad. Esto trae como consecuencia el uso restringido o propietario de una tecnología, solución o servicio. Por otro lado, la integración continua y el aprovisionamiento de la infraestructura en la nube, provoca tareas repetitivas.

Para resolver el problema planteado se propone en la presente investigación proveer soporte para la gestión de herramientas DevOps. Se formulan dos propuestas, definir un Lenguaje Específico de Dominio basado en el concepto de Infraestructura como Código. Y la creación de una herramienta que apoya este lenguaje que permite modelar gráficamente.

Los aspectos que podrán modelar los ingenieros de software son, el estado final de un aprovisionamiento de infraestructura en la nube y los microservicios orquestados. De este modo se genera el aprovisionamiento de scripts multiplataforma como solución

guiada por la infraestructura como código. La herramienta propuesta optimiza y facilita el proceso de desarrollo en el marco de la ingeniería de software dirigida por modelos.

Palabras Claves: Ingeniería de Software Dirigida por Modelos, DevOps, IaaS, IaC, Microservicios, CTS.

CONTEXTO

Este proyecto de investigación de reciente creación se encuentra en desarrollo en el marco del grupo AuSegTIC (Grupo de Auditoría y Seguridad de TIC) y el GRUPO GRID TICs, ambos grupos pertenecientes a la Facultad Regional Mendoza de la Universidad Tecnológica Nacional. Es un proyecto interinstitucional entre la UTN FRM y el Instituto de Informática de la Universidad de Mendoza.

La misma es parte de un proceso de incentivación para el desarrollo de actividades I/D, que se originan en el marco de este acuerdo.

INTRODUCCIÓN

La Ingeniería Dirigida por Modelos (MDE, por sus siglas en inglés)[1], es un paradigma de desarrollo de software que utiliza modelos para

la generación de código y otros artefactos [2]. Estos modelos pueden crearse con lenguajes de modelado de propósito general, por ejemplo, UML[3] o con un DSL, en inglés (Domain Specific Language)[4] Incorpora al proceso de producción de software la abstracción y el formalismo necesario, para automatizar y optimizar tareas críticas del proceso de desarrollo. De esta forma, MDE permite mejorar la productividad, la portabilidad, la interoperabilidad y el mantenimiento de los sistemas .

La Ingeniería de Software Dirigida por Modelos se puede utilizar para proveer soporte al proceso de aprovisionamiento de infraestructura como código en la nube y microservicios aplicando las mejores prácticas de DevOps.

El área de DevOps [5], está conformada por desarrolladores y personal de operaciones, que necesitan definir, actualizar y ejecutar los recursos de infraestructura en diferentes proveedores de servicios en la nube. La Infraestructura como servicio (IaaS, por sus siglas en inglés) [6], es un tipo de servicio de informática en la nube que ofrece recursos esenciales de proceso, almacenamiento y redes a petición que son de pago por uso . La Infraestructura como Código (IaC, por sus siglas en inglés) [7], es el proceso de automatizar los cambios en la infraestructura a través de código en lugar de procesos manuales, para lograr escalabilidad, fiabilidad y seguridad.

El problema que los clientes de las plataformas de cloud deben enfrentar es el despliegue y control de una aplicación compleja sobre un conjunto heterogéneo de proveedores. Los proveedores de servicios IaaS ofrecen diferentes tipos de infraestructura con diversidad de lenguajes de scripting para: definir, actualizar y ejecutar la infraestructura en la nube.

Los proveedores exponen sus servicios de acuerdo con especificaciones heterogéneas incurriendo en una falta de portabilidad e interoperabilidad convergiendo en la problemática que implica el uso restringido o propietario de una tecnología, solución o servicio desarrollado por un proveedor.

La integración continua y el aprovisionamiento de la infraestructura en la nube, provoca tareas repetitivas. Además se necesita personal con un alto grado de capacitación para configurar todo el entorno.

En la actualidad, existen varias herramientas para gestionar el aprovisionamiento de infraestructura. Estas utilizan scripts para definir el estado final de la infraestructura en la nube. Sin embargo, la gestión de lenguajes de scripts de diferentes herramientas de la comunidad DevOps para el aprovisionamiento de infraestructura, es una tarea que requiere mucho tiempo y es propensa a errores a la que se enfrentan los profesionales. Además, la integración infraestructura-microservicio se convierte en una ardua tarea cuya complejidad crece a la par de la arquitectura del software que se pretende desplegar.

Existen trabajos de investigación relacionados que abordan los problemas de la Infraestructura como Código mediante el desarrollo dirigido por Modelos (MDD, por sus siglas en inglés). Podemos destacar a: [8], [10] y [11].

En [8], (Borisova) se analiza la aplicabilidad del estándar TOSCA [9] para unificar el despliegue y la orquestación no sólo de los recursos provistos por la virtualización basada en la nube, sino también usando contenedores.

En [10], (Sandobalín) se presenta una herramienta para modelar el estado final de aprovisionamiento de la infraestructura en la nube a través de un DSL con una notación

gráfica. En [11], (Brabra) propone un enfoque de orquestación dirigido por modelos, que utiliza TOSCA para describir artefactos de recursos en la nube. Utilizan una técnica de transformación en el marco de la ingeniería dirigida por modelos (MDE) .

La diferencia fundamental de las propuestas [8], [10] y [11] con nuestra propuesta, está en el metamodelo propuesto y la notación gráfica con Sirius[12] que representa la sintaxis concreta. Los elementos contemplados en el metamodelo permiten el diseño de la notación gráfica que contempla el aprovisionamiento con contenedores utilizando Kubernetes Engine y otros proveedores de servicios como: Google y Amazon. Esto le permite al DevOps contar con una herramienta fácil de usar y flexible. Además se brinda una solución que abstrae la complejidad de especificar los recursos de infraestructura de diversos proveedores de servicios, así como la automatización de la generación de scripts para el aprovisionamiento y configuración de la infraestructura en la nube. Nuestra propuesta permite el modelado de la infraestructura teniendo en cuenta las aplicaciones que se van a desplegar, como las clases y las rutas de cada microservicio. Se integra no sólo la infraestructura, sino también los microservicios y el código, permitiendo un despliegue ya integrado.

Este proyecto de investigación tiene como objetivo optimizar el proceso de aprovisionamiento de infraestructura en la nube mediante el diseño e implementación de un artefacto que permita reducir riesgos y mejorar la calidad en el proceso productivo del software.

Para resolver el problema planteado se propone proveer soporte para la gestión de herramientas DevOps, a través de la definición de un Lenguaje Específico de Dominio basado

en el concepto de Infraestructura como Código, y una herramienta que apoya este lenguaje que permite modelar el estado final de un aprovisionamiento infraestructura en la nube como así también el modelado de los microservicios orquestados, generando el aprovisionamiento de scripts multiplataforma como solución guiada por la infraestructura como código. La herramienta propuesta optimiza y facilita el proceso de desarrollo en un marco de calidad en base al desarrollo de software dirigido por modelos.

En resumen, el aporte de esta propuesta es Accelerate, una herramienta de modelado de infraestructura para el aprovisionamiento de la nube y microservicios orquestados, que pretende abstraer la complejidad de trabajar con diferentes herramientas DevOps a través de un dominio específico y proveer una configuración inicial parametrizable, que pueda ser utilizada como un marco de trabajo que guíe tanto el desarrollo como el despliegue del software.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación propuesta se enfoca en utilizar el enfoque de la Ingeniería de Software Dirigida por Modelos. Para proveer soporte al proceso de aprovisionamiento de infraestructura como código en la nube y microservicios. Aplicando las mejores prácticas de DevOps, para la mejora de la calidad del Software, desde una perspectiva CTS (Ciencia, Tecnología y Sociedad).

El objeto de estudio es crear un nuevo artefacto denominado Accelerate para proveer soporte a la Infraestructura como Código en el proceso de aprovisionamiento de infraestructura en la nube y la orquestación de Microservicios.

RESULTADOS ESPERADOS

Se definen los siguientes objetivos:

Objetivo General

Optimizar el proceso de aprovisionamiento de infraestructura en la nube mediante el diseño e implementación de un artefacto Accelerate, a los efectos de reducir riesgos y mejorar la calidad en el proceso productivo del software con un enfoque de la ciencia tecnología y sociedad (CTS).

Objetivos Específicos:

- Diseñar una aproximación de MDE (Ingeniería dirigida por modelos) para el aprovisionamiento en la nube y modelado de microservicios orquestados. En este caso, el artefacto Accelerate plantea mejorar el proceso de aprovisionamiento de infraestructura en la nube utilizando IaC y MDE.
- Diseñar un metamodelo que permita abstraer los conceptos fundamentales de la infraestructura en la nube y que establezca la sintaxis abstracta.
- Diseñar un lenguaje específico de dominio y sus correspondientes editores proporcionando la sintaxis concreta, para el modelado y generación de prototipos.
- Definir las plantillas que generan los script de la IaC como ingreso a las transformaciones de modelo a texto (M2T).
- Definir los pasos estratégicos y sistemáticos para la incorporación del marco referencial que acompañen al proceso de desarrollo de software.
- Evaluar y validar el marco referencial propuesto para el desarrollo de software a través del enfoque de desarrollo de software de ingeniería experimental.

-Implementar la seguridad en el diseño, desarrollo y funcionamiento de Accelerate.

-Incorporar la evaluación e innovación de tecnología necesaria para el proyecto desde una perspectiva CTS.

FORMACIÓN DE RECURSOS HUMANOS

En este proyecto se ha previsto la siguiente contribución a la formación de recursos humanos:

Formación en investigación de docentes, graduados y estudiantes de las Universidades participantes.

Soporte y apoyo didáctico a las cátedras participantes (Proyecto Final, Habilitación Profesional, Desarrollo de Software Dirigido por Modelos (Electiva),y Evaluación e Innovación de Tecnología desde la perspectiva CTS (Ciencia, Tecnología y Sociedad) (Electiva) del Departamento de Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional, Facultad Regional Mendoza, como así también al Instituto de Informática y las cátedras Auditoría de Sistemas y Diseño de Sistemas de la Universidad de Mendoza,

Por tratarse de un proyecto de investigación interinstitucional, contribuye a la vinculación de las universidades participantes cooperando en la creación de conocimiento (Universidad Tecnológica Nacional Facultad Regional Mendoza y la Universidad de Mendoza).

Se formarán ayudantes de primera y segunda en las materias participantes, elaborarán seminarios y talleres en los temas incluidos en esta línea de investigación, para estudiantes, graduados y docentes de las carreras involucradas y para organizaciones públicas y privadas.

BIBLIOGRAFÍA

- [1] Schmidt, D.C.: Model-driven engineering. COMPUTER-IEEE COMPUTER SOCIETY- 39(2), 25 (2006)
- [2] Patterns: Model-Driven Development Using IBM Rational Software Architect, Peter Swithinbank, Mandy Chessell, Tracy Gardner, Catherine Griffin, Jessica Man, Helen Wylie, Larry Yusuf, disponible en ibm.com/redbooks.
- [3] Unified Modeling Language(UML).version 2.4 OMG,<http://www.omg.org/spec/UML/2.4.1/>
- [4] Brambilla M. , Cabot J. , Wimmer M. (2017) “Model-Driven Software Engineering in Practice”. Second Edition. Morgan & Claypool Publisher.University of Illinois at Chicago.
- [5] C. A. Cois, J. Yankel, and A. Connell, “Modern devops: Optimizing software development through effective system interactions.” in IPCC. IEEE, 2014, pp. 1–7.
- [6] ¿Qué es IaaS?,último acceso: febrero 2023. [Online]: Available:<https://azure.microsoft.com/es-mx/resources/cloud-computing-dictionary/what-is-iaas/>
- [7] K. Morris, Infrastructure As Code: Managing Servers in the Cloud. O'Reilly & Associates Incorporated, 2016.
- [8] Borisova, A., Shvetcova, V., & Borisenko, O. (2020). Adaptation of the TOSCA standard model for the Kubernetes container environment. Proceedings - 2020 Ivannikov Memorial Workshop, IVMEM 2020, 9–14. <https://doi.org/10.1109/IVMEM51402.2020.00008>
- [9] Lauwers, C., & Tamburri, D. (n.d.). OASIS Topology and Orchestration Specification for Cloud Applications (TOSCA) TC. https://www.Oasis-Open.Org/Committees/Tc_home.Php?Wg_abbrev=tosca.
- [10] Sandobalín, J., Insfran, E., & Abrahão, S. (2017). Automatización del Aprovechamiento de Infraestructura en la Nube. 11705/JCIS/2017/018
- [11] Brabra, H., Mtibaa, A., Gaaloul, W., Benatallah, B., & Gargouri, F. (2019). Model-driven orchestration for cloud resources. IEEE International Conference on Cloud Computing, CLOUD, 2019-July. <https://doi.org/10.1109/CLOUD.2019.00074>
- [12] The easiest way to get your own Modeling Tool [Internet]. Ottawa, Ontario:Eclipse Foundation [citado 09 Ago 2020] Disponible en: <https://www.eclipse.org/sirius/>

HACIA LA EVALUACION DE LA CALIDAD DE DATOS ABIERTOS

Ana Funes⁽¹⁾, Aristides Dasso⁽¹⁾, María Alejandra Barrera⁽²⁾

⁽¹⁾Departamento de Informática, Facultad de Ciencias Físico-Matemáticas y Naturales,
Universidad Nacional de San Luis
Ejército de los Andes 950 - 5700 San Luis, Argentina
afunes@unsl.edu.ar, arisdas@unsl.edu.ar

⁽²⁾Departamento de Informática, Facultad de Tecnología y Ciencias Aplicadas,
Universidad Nacional de Catamarca,
Maximio Victoria 55, Catamarca, Argentina
mbarrera@tecno.unca.edu.ar

RESUMEN

Según la Open Knowledge Foundation¹, los datos abiertos son datos que están disponibles de forma pública a cualquier persona que desee utilizar, reutilizar y redistribuir libremente, sujetos únicamente, como máximo, al requisito de atribuir y compartir por igual.

Estos datos suelen estar disponibles en formatos digitales y se publican bajo diversos tipos de licencias. Considerando los beneficios derivados de su adopción, múltiples iniciativas han sido creadas con el objetivo de promover su uso, tales como la Alianza para el Gobierno Abierto (Open Government Partnership), el Portal de Datos Abiertos de la Unión Europea y el Portal de Datos Abiertos de los Estados Unidos, entre otras. En Argentina, la iniciativa Datos Argentina² del gobierno argentino tiene como objetivo mejorar la transparencia y la participación ciudadana mediante la publicación de datos abiertos.

Frente a este escenario, consideramos importante contar con modelos y herramientas automatizadas que permitan llevar a cabo evaluaciones de la calidad de repositorios de datos abiertos, y que sirvan de guía a los organismos gubernamentales al evaluar la calidad de los datos que publican, contribuyendo de esta forma en la

transparencia y la rendición de cuentas, así como en la resolución de problemas sociales y económicos importantes.

Palabras clave: datos abiertos, calidad, métodos de evaluación multicriterio, métricas.

CONTEXTO

El presente trabajo se encuentra enmarcado en una investigación conjunta entre investigadores del Proyecto de Ciencia y Técnica PROICO 03-2020 "Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube", de la Universidad Nacional de San Luis. (Director: Daniel Riesco. Acreditado con evaluación externa. Financiamiento: Universidad Nacional de San Luis) e investigadores del Proyecto: "Ingeniería de Software en la era de las Industrias 4.0.", que se desarrolla en el departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca, también acreditado con evaluación externa.

En este contexto, se ha venido trabajando desde hace tiempo en el ámbito del SEG (Software Engineering Group) de la Universidad Nacional de San Luis, sobre la construcción de modelos de evaluación de sistemas complejos, donde se han obtenido resultados que han sido plasmados en diversas publicaciones (ver por ejemplo [5], [6], [7], [9], [10], [11], [12]).

¹ <https://opendatahandbook.org/guide/en/what-is-open-data/>

² <https://www.datos.gob.ar/>

1. INTRODUCCIÓN

Según la Open Knowledge Foundation³ el conocimiento es abierto si cualquiera es libre de acceder a él, usarlo, modificarlo y compartirlo, sujeto, como máximo, a medidas que preserven la procedencia y la apertura. Es un concepto general que se puede aplicar a diversos elementos o partes del conocimiento transferido, como por ejemplo al código abierto, a publicaciones, textos y, en particular, a los datos.

Los datos abiertos, entendidos como una de las formas de conocimiento abierto, son datos que cualquier persona puede utilizar, reutilizar y redistribuir libremente, sujetos únicamente, como máximo, al requisito de atribuir y compartir por igual [13].

Uno de los usos más inmediatos está asociado al gobierno electrónico ya que para los organismos gubernamentales es particularmente importante, no solo por la cantidad y la centralidad de los datos que recopila, sino también por el tipo de información que brinda la mayoría de esos datos: datos públicos por ley, los cuales podrían estar disponibles públicamente para su uso.

Esto contribuye a mejorar la transparencia de este tipo de organismos ya que permite a los ciudadanos monitorearlos y evaluarlos a través de la información que se hace pública.

En algunos casos, los datos abiertos también pueden involucrar a los ciudadanos en el proceso de toma de decisiones, lo que puede aumentar la participación y la colaboración entre las partes interesadas, así como mejorar la calidad de las decisiones tomadas.

En este sentido, no solo los organismos gubernamentales pueden beneficiarse de la adopción de datos abiertos. Otros organismos tanto públicos como privados también pueden verse favorecidos con estos y otros aspectos tales como la aplicación en la investigación y análisis de datos, o en la innovación y desarrollo económico.

Considerando los beneficios antes mencionados, creemos necesario disponer de métricas que permitan evaluar aspectos de calidad de los datos abiertos ya este conocimiento no solo va a contribuir, desde el punto de vista de los proveedores de datos, en mejorar la calidad de los datos publicados, ya que ayudará a identificar errores y deficiencias, sino también, desde el punto de vista del consumidor de los datos, para ganar en confianza sobre los mismos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Considerando los beneficios que surgen del uso de datos abiertos y de la posibilidad de evaluar su calidad, nos planteamos como objetivo principal de esta línea de trabajo la creación de un modelo de atributos de calidad de datos abiertos así como la definición, sobre la base de los atributos identificados en dicho modelo, de métricas que sirvan para obtener un indicador de la calidad de los mismos.

Para lograr dicho objetivo, los ejes principales de la presente línea de investigación involucran las siguientes actividades:

- ☒ Análisis bibliográfico de propuestas alternativas para la identificación y conceptualización de los atributos de calidad de datos abiertos, analizando y definiendo las propiedades específicas que deben ser tenidas en cuenta en dicho contexto.
- ☒ Modelización, estableciendo métodos de cálculo o procedimientos en forma de métricas, para obtener un valor numérico por cada atributo a ser considerado.
- ☒ Aplicación en la evaluación de la calidad de datos abiertos, para obtener el valor de un indicador global del Sistema/Producto evaluado.

3. RESULTADOS OBTENIDOS/ ESPERADOS

La primera de todas las actividades llevadas a cabo se basó en el análisis de

³ <https://okfn.org/>

publicaciones relacionadas a la evaluación de datos abiertos, donde hemos buscado identificar los diversos atributos de calidad que han sido considerados.

Así, por ejemplo, el trabajo de A. Vetro et al. [14] presenta un marco para la evaluación de datos abiertos, el cual viene aplicado a repositorios de datos publicados por el gobierno italiano. En este marco se identifican atributos o características de calidad tales como la fiabilidad, precisión, actualidad y accesibilidad, entre un total de siete, para los cuales se definen siete métricas simples de conteo o porcentajes. No se describe ninguna métrica que integre los resultados en un indicador global.

En otro trabajo, Tim Berners-Lee publica, en 2006, un esquema de implementación para datos abiertos, basado en cinco requisitos incrementales en exigencia [4]. Según él, el conjunto de datos abierto debe cumplir con los siguientes requisitos: (1). Disponibilidad en la web, cualquier formato proporcionado con licencia abierta; (2) Disponibilidad como datos estructurados legibles por máquina (por ejemplo, Excel en lugar de escaneo de imagen); (3). Formato no propietario disponible (p. ej., CSV en lugar de Excel); (4) Hacer uso de estándares abiertos del W3C (RDF y SPARQL) y URI para identificar cosas; (5) Vinculación de los datos a los datos de otros proveedores para proporcionar contexto. Esta propuesta, considera solo un aspecto de la calidad de los datos relativa al formato o codificación de los mismos, dejando de lado aspectos de calidad tales como la integridad, la consistencia y la puntualidad.

Por otro lado, el trabajo de Arbello et al. [1] se centra en la definición de una métrica, que han dado en llamar MELODA, para la evaluación de la reutilización de datos abiertos. La misma se basa en 4 dimensiones: estándares técnicos, acceso, legal y modelo, las cuales hacen referencia al formato de los datos; al acceso o mecanismo por el cual se hace posible la descarga o conexión con la información; a las restricciones legales de

acceso y a la descripción de la estructura de los datos, respectivamente.

Los mismos autores en [2] redimensionan su modelo, proponiendo una nueva métrica, MELODA5, la cual se basa en ocho dimensiones, que son una evolución de MELODA 4.13 y donde el nombre de la dimensión "modelo" es renombrada a "estandarización" y donde se agregan dos nuevas dimensiones, una para medir la diseminación de las actividades y otra para analizar la reputación de la fuente de datos.

Por otro lado, la Open Data Barometer (ODB) en la 4ta edición de su Informe Global⁴, establece que para que los datos abiertos sean de valor y usables, deben, además, ser comprensibles, precisos y de alta calidad. Asimismo, establece que los proveedores de datos, deberían asegurar que cuentan con mecanismos que permitan a los usuarios proveer retroalimentación, haciendo posible revisiones continuas de sus datos con el objetivo de mejora.

Por otro lado, en el portal de la W3C⁵, se brinda una serie de recomendaciones a la hora de publicar datos gubernamentales abiertos a partir de las cuales se pueden identificar atributos que hacen a la calidad de tales repositorios. También, la Sunshine Foundation⁶ identifica, en su portal, una lista de diez principios que brindan una perspectiva para evaluar hasta qué punto los datos gubernamentales son abiertos y accesibles para el público.

En base a este análisis es que, en una primera etapa, hemos podido concluir que la calidad de los datos abiertos no se restringe a evaluar sólo la calidad técnica de los datos o solo aquellos aspectos propios de datos abiertos sino que, si bien consideramos importante identificar claramente aquellas dimensiones que definen a los datos como abiertos, tales como la licencia, las formas de

⁴ <https://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport-ES.pdf>

⁵ <https://www.w3.org/TR/gov-data/>

⁶ <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

acceso, la legibilidad y el formato, entre otras, ya que al asegurar que se cumplen estos requisitos, se garantiza que los datos sean accesibles y utilizables por cualquier persona o entidad interesada en ellos, consideramos también importante, tal como establece la iniciativa de datos abiertos del Gobierno de España⁷, no dejar de lado la calidad técnica de los mismos, que aunque no esté directamente contemplada en los principios de los datos abiertos, deberían también tenerse en cuenta a la hora de producir cualquier tipo de datos como, por ejemplo, algunos atributos de calidad definidos por la ISO/IEC 25012, que podrían adaptarse fácilmente:

- ☒ *Exactitud*: se encuentran dentro del rango de valores válidos definidos para el dominio de aplicación.
- ☒ *Compleitud*: se aportan los valores correspondientes a todos los atributos disponibles.
- ☒ *Credibilidad*: tanto para los datos en sí como para la fuente de información.
- ☒ *Actualidad*: proporcionados en el momento preciso para mantener su valor.
- ☒ *Accesibilidad*: facilidad de acceso en su contexto.
- ☒ *Conformidad*: con respecto a los estándares y normativas vigentes.
- ☒ *Confidencialidad*: respetando la privacidad y seguridad de los datos.
- ☒ *Eficiencia*: para que puedan ser procesados con unos recursos razonables.
- ☒ *Precisión*: respecto al contexto al que pertenecen.
- ☒ *Trazabilidad*: respecto a la fuente u origen de los datos.
- ☒ *Comprensibilidad*: con una codificación adecuado para su posterior interpretación.

Muchos de estos atributos de calidad de los datos han sido identificados también en el proyecto de datos abiertos de la Comisión Europea [8].

Finalmente, en base al análisis realizado, se pudo observar también que muchas de las propuestas analizadas consideran atributos de

calidad comunes aunque bajo diferentes nombres.

En base a estas conclusiones, se planea (a) desarrollar un modelo jerárquico de atributos de calidad de datos abiertos, que unifique e incluya las dimensiones, subdimensiones y atributos necesarios para definir una métrica de calidad que considere estos aspectos antes analizados. (b) A partir del modelo jerárquico de calidad construido, definir un conjunto de métricas que permitan medir los atributos hojas de dicha jerarquía. (c) Definir una función multicriterio que agregue las características antes identificadas y que permita la evaluación tanto integral como de subcaracterísticas del modelo jerárquico. Este modelo cuantitativo de evaluación, permitirá obtener los valores de indicadores parciales de las características a diferentes grados de abstracción del modelo jerárquico de calidad, así como un valor de un indicador global del repositorio de datos evaluado. (d) Aplicar el modelo y las métricas asociadas para evaluar la calidad de los datos a diversos portales de datos abiertos.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación sobre construcción de modelos de evaluación de sistemas complejos ha producido dos tesis de posgrado en la Maestría en Ingeniería de Software de la Universidad Nacional de San Luis. Una de ellas sobre evaluación de atributos de calidad de sitios de gobierno electrónico [1][6] y la otra sobre el atributo de calidad Accesibilidad en aplicaciones web [9][10]. Asimismo, uno de los autores de este trabajo ha realizado también su tesis de Maestría en Ingeniería de Software de la Universidad Nacional de San Luis en temas relacionados a la evaluación de la reusabilidad de datos espaciales abiertos [3].

La propuesta aquí expuesta, también, tiene como objetivo ser motivo de desarrollo de tesis en el ámbito del Departamento de Informática de la Universidad Nacional de San Luis, como así también en ámbitos del Departamento de Informática de la Universidad Nacional de Catamarca.

⁷ <https://datos.gob.es/es/noticia/datos-abiertos-y-de-calidad>

5. BIBLIOGRAFIA

- [1] Alberto Abella García, Marta Ortiz de Urbina Criado, Carmen de Pablos Heredero. El profesional de la información, ISSN 1699-2407, Vol. 23, Nº 6, 2014 , págs. 582-588.
- [2] Alberto Abella; Marta Ortiz de Urbina Criado; Carmen De Pablos Heredero. El profesional de la información, ISSN 1699-2407, Vol. 28, Nº 6, 2019.
- [3] María Alejandra Barrera. Una estrategia de evaluación de la reusabilidad de los conjuntos de datos de los portales de Infraestructuras de Datos Espaciales Tesis de Maestría en Ingeniería de Software. Universidad Nacional de San Luis, Argentina. 2022.
- [4] Berners-Lee, T. (2006) Linked data-design issues. Tech. rep., W3C, <http://www.w3.org/DesignIssues/LinkedData.html>.
- [5] M. Castro. "Análisis de las propiedades y atributos propios de sitios de gobierno electrónico", Tesis de Maestría en Ingeniería de Software, Departamento de Informática, Universidad Nacional de San Luis, 2010.
- [6] M. Castro, A. Dasso, A. Funes. "Modelo de Evaluación para Sitios de Gobierno Electrónico", Simposio de Informática en el Estado (SIE) 2009 – 38 JAIIO, Mar del Plata, Argentina, August 26-28, 2009. pp. 200-214.
- [7] A. Dasso y A. Funes, "Threat and Risk Assessment Using Continuous Logic", Encyclopedia of Organizational Knowledge, Administration, and Technologies, 1st. edition. IGI Global. 2020.
- [8] Makx Dekkers, Nikolaos Loutas, Michiel De Keyser and Stijn Goedertier. Open Data & Metadata Quality. European Commission. Open Data Support. 2014. https://joinup.ec.europa.eu/sites/default/files/document/2015-05/d2.1.2_training_module_2.2_open_data_quality_v1.00_en.pdf.
- [9] C. Gallardo, A. Funes, H. Ahumada. "Soporte para la Medición y Evaluación de la Accesibilidad al Contenido en Aplicaciones Web", Anales de ASSE 2019 (JAIIO 2019), Salta, Argentina. pp. 56-70.
- [10] C. Gallardo, A. Funes. "Un Modelo para la Evaluación de la Calidad de la Accesibilidad al Contenido Web", CONAIIISI 2015, Bs. As., Argentina.
- [11] E. Miranda et al. "NESSy: A new evaluator for software development tool". 2nd Symposium on Languages, Applications and Technologies. 2013.
- [12] M. Peralta, C. Salgado. "Un a Herramienta para la Evaluación de Sistemas", Tesis de Licenciatura en Ciencias de la Computación, Universidad Nacional de San Luis. 2004.
- [13] The Open Knowledge Foundation. The Open Data Handbook. <https://opendatahandbook.org/guide/en/>
- [14] Antonio Vetrò, Lorenzo Canova, Marco Torchiano, Camilo Orozco Minotas, Raimondo Iemma, Federico Morando, Open data quality measurement framework: Definition and application to Open Government Data, Government Information Quarterly, Volume 33, Issue 2, 2016, Pags 325-337, ISSN 0740-624X, <https://doi.org/10.1016/j.giq.2016.02.001>.

Aplicaciones Móviles, Realidad Virtual y Realidad Aumentada

Pablo Thomas , Federico Cristina , Sebastián Dapoto , Patricia Pesado 

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{pthomas, fcristina, sdapoto, ppesado}@lidi.info.unlp.edu.ar

▪ Resumen

Se presenta una línea de investigación y desarrollo, que tiene por objeto estudiar temas relacionados con aspectos de Ingeniería de Software, orientados al desarrollo de aplicaciones móviles tridimensionales (3D) sobre diversas plataformas. En particular, se pone el foco en el desarrollo de aplicaciones 3D inmersivas, a través del uso de Realidad Virtual (RV) y Realidad Aumentada (RA). También se realiza investigación en el desarrollo de aplicaciones móviles basadas en la ubicación, haciendo uso de métodos de posicionamiento en interiores.

Palabras claves: Dispositivos Móviles – Aplicaciones 3D – Aplicaciones Multiplataforma – Realidad Virtual – Realidad Aumentada – M-Learning – Posicionamiento en interiores

▪ Contexto

Esta línea de Investigación forma parte del proyecto “*Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos. Mejora de proceso.*”, en particular del subproyecto “*Ingeniería de Software para escenarios híbridos*”, del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Existe una importante cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y

Alumnos de Doctorado en el área de Informática.

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado.

▪ Introducción

En la actualidad los dispositivos móviles permiten ejecutar aplicaciones complejas y con exigentes requerimientos de hardware. Debido a esto, existen cada vez más alternativas de motores de juego que permiten desarrollar aplicaciones tridimensionales para dispositivos móviles.

La RV es una simulación interactiva por computadora en la cual se sustituye el mundo real con información sensorial que recibe el usuario. La RV permite generar entornos inmersivos donde el usuario puede interactuar con representaciones virtuales de objetos, que de otro modo sería difícil o imposible de acceder [1].

La RA es el conjunto de tecnologías que permiten que el usuario visualice el mundo real con información virtual añadida, mediante un dispositivo tecnológico. Los elementos físicos reales y tangibles se combinan con elementos virtuales. La RA es interactiva y en tiempo real [2].

El ámbito educativo debe adaptarse a los cambios y nuevas formas de aprendizaje. M-learning (mobile learning) plantea métodos modernos de apoyo al proceso de aprendizaje mediante el uso de dispositivos móviles. Las

aplicaciones móviles 3D son una herramienta ideal para acercar a los alumnos [3] [4] [5].

Los sistemas de RA pueden transformar la forma de aprender y trabajar [6]. Mediante la utilización de gafas de RA y sensores inteligentes, es posible, por ejemplo, asistir a un técnico que realiza tareas de mantenimiento y deba realizar procedimientos con las manos libres.

Por otro lado, el posicionamiento en interiores (en inglés, indoor localization) puede ser esencial para actividades como la atención médica, la seguridad, los juegos de realidad aumentada y muchos otros servicios basados en la ubicación. Existen diferentes técnicas para estimar el posicionamiento en interiores [7]; una de las más utilizadas es mediante el uso de los denominados *beacons*, unos pequeños dispositivos basados en tecnología Bluetooth de bajo consumo, que emiten una señal que identifica de forma única a cada dispositivo. Ésta señal puede ser recibida e interpretada por un dispositivo móvil, conociendo además la distancia a la que se encuentran.

▪ Líneas de Investigación y Desarrollo

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones Móviles 3D Multiplataforma
- Mobile Learning
- Frameworks para el desarrollo de Aplicaciones Móviles 3D
- Realidad Virtual en aplicaciones móviles 3D
- Realidad Aumentada en aplicaciones móviles 3D
- Aplicaciones móviles basadas en el posicionamiento en interiores

▪ Resultados esperados/obtenidos

Los resultados esperados/obtenidos se pueden resumir en:

- Avanzar en la capacitación continua de los miembros de la línea de investigación.
- Avanzar en el aprendizaje de frameworks o motores de juego que permiten desarrollar aplicaciones 3D multiplataforma, particularmente para dispositivos móviles [8] [9] [10] [11].
- Avanzar en el desarrollo de aplicaciones educativas, teniendo como finalidad enriquecer las experiencias interactivas y motivar el aprendizaje mediante su uso.
- Avanzar en el desarrollo de aplicaciones móviles 3D con Realidad Virtual y Realidad Aumentada.
- Avanzar en la investigación y desarrollo de aplicaciones móviles basadas en el posicionamiento en interiores.
- Se ha ampliado la funcionalidad de la aplicación móvil R-Info3D [3], una herramienta de aprendizaje sencilla de los conceptos básicos para la construcción de algoritmos en la Facultad de Informática. Se ha optimizado la aplicación e incorporado funcionalidad para el modo de uso con lentes de RV. Figura 1.

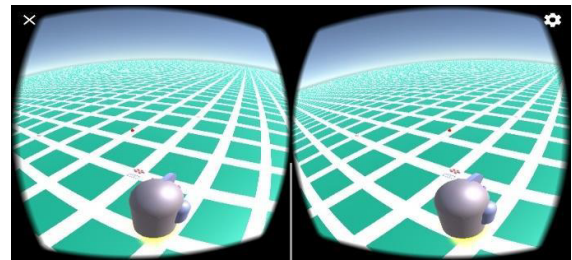


Figura 1. La aplicación móvil R-Info3D utilizado con lentes de RV.

- Se está avanzando en el desarrollo de un prototipo móvil 3D de asistencia a estudiantes de veterinaria, que permite conocer las herramientas utilizadas en las diferentes prácticas veterinarias. La definición del prototipo y su funcionalidad fue especificada con la colaboración de docentes de la Facultad de Veterinaria. La aplicación brinda información sobre las distintas herramientas de cirugía y mediante el uso de la RA hace posible su

observación en la escala original y desde cualquier punto de vista. Esto permite conocer los detalles de las herramientas y aprender el correcto uso de cada una de ellas. Figuras 2 y 3.

- Se está avanzando en el desarrollo de un prototipo de asistencia de ubicación para personas ciegas en la Facultad de Informática.

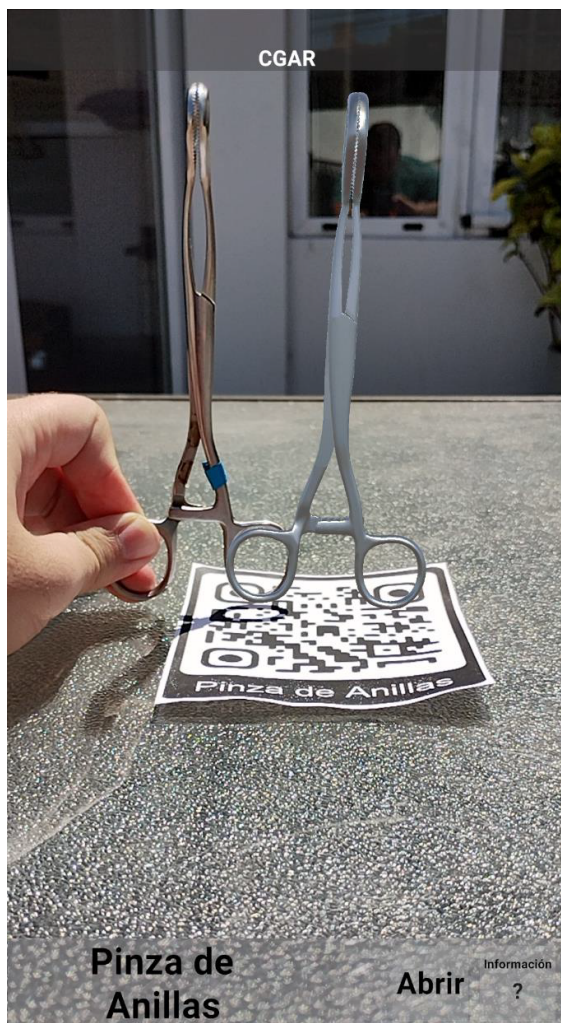


Figura 2. Prototipo móvil 3D de asistencia a estudiantes de veterinaria. Visualización de una herramienta real junto a la herramienta aumentada a escala.



Figura 3. Prototipo móvil 3D de asistencia a estudiantes de veterinaria. Visualización del set completo de herramientas.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

Referencias

1. Linowes J. "Unity Virtual Reality Projects". 2015. ISBN-13: 978-1783988556.
2. R. Silva, J.C. de Oliveira, G. Giraldi. "Introduction to augmented reality". 2003.
3. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. Capítulo de Libro: "3D Mobile Prototype for Basic Algorithms

- Learning”. Libro: “Computer Science & Technology Series - XXI Argentine Congress Of Computer Science. Selected Papers” (300 páginas). EDULP. ISBN: 978-987-4127-00-6, páginas 239-247. Año 2016.
4. Kantel E., Tovar G., Serrano A.”Diseño de un Entorno Colaborativo Móvil para Apoyo al Aprendizaje a través de Dispositivos Móviles de Tercera Generación.” IEEE-RITA 5, no. 4 (2010): 146-151.
 5. Rosa Paredes, J. Alfredo Sánchez, Liliana Rojas, Daniel Strazzulla, Ronel Martínez-Teutle. "Interacting with 3D Learning Objects". 2009 LA Web Congress. ISBN: 978-0-7695-3856-3/09.
 6. X. Pan, X. Sun, H. Wang, S. Gao, N. Wang and Z. Lin, "Application of an assistant teaching system based on mobile augmented reality (AR) for course design of mechanical manufacturing process," 2017 IEEE 9th International Conference on Engineering Education (ICEED), 2017, pp. 192-196, doi: 10.1109/ICEED.2017.8251191.
 7. Gu, Fuqiang & Hu, Xuke & Ramezani, Milad & Acharya, Debaditya & Khoshelham, Kourosh & Valaee, Shahrokh & Shang, Jianga. (2019). Indoor Localization Improved by Spatial Context - A Survey. ACM Computing Surveys. 52. 64:1-35. 10.1145/3322241.
 8. Unity 3D Homepage: <https://unity3d.com/>.
 9. Unreal Engine Homepage.
<https://www.unrealengine.com/>.
 10. CryEngine Homepage.
<https://www.cryengine.com/>
 11. Godot Engine Homepage.
<https://godotengine.org>

Evaluación de las políticas de protección de datos personales en las iniciativas de apertura de datos en el ámbito de la Universidad Nacional de Catamarca

Claudia M. Herrera¹, Carolina I. Chayle¹, Ana C. Pauletto¹; Morales Apaza Oswaldo W.N.¹
Departamento de Informática- Facultad de Tecnología y Ciencias Aplicadas - Universidad
Nacional de Catamarca
herrera.claudia.mabel@gmail.com cchayle@gmail.com

RESUMEN

El modelo de Dato Abierto se basa en el acceso y uso de la información pública por parte de terceros para entregar nuevos servicios a los ciudadanos. Trasladando estos principios al ámbito universitario la apertura de datos en los diferentes dominios que articulan la universidad (docencia, investigación, biblioteca, etc.) aporta notables beneficios tanto a la comunidad universitaria como a la propia sociedad donde se inserta. Las universidades utilizan y generan una gran cantidad de información personal, esto representa una gran fuente de recursos. A medida que se publican más datos, aumenta el riesgo de divulgar información privada, por lo que uno de los principales desafíos de estas iniciativas es garantizar un programa de datos abiertos responsable que tenga en cuenta la privacidad de los datos y el cumplimiento de la Ley de Protección de Datos Personales.

Con la presente investigación se pretende evaluar las políticas de Protección de Datos Personales en las iniciativas de apertura de datos en el ámbito de la Universidad Nacional de Catamarca, detectadas a partir de los resultados obtenidos en el proyecto “Datos abiertos en Educación, evaluación de los alcances de las iniciativas existentes. Caso de estudio Universidad Nacional de Catamarca” (2018-2019).

Palabras clave: Dato Abierto, Protección de Datos Personales, TIC, UNCA.

CONTEXTO

A partir de los resultados obtenidos en el proyecto “Datos abiertos en Educación, evaluación de los alcances de las iniciativas existentes. Caso de estudio Universidad Nacional de Catamarca” (2018-2019), se

pretende con la presente investigación evaluar las políticas de Protección de Datos Personales (PDP) en el marco de las iniciativas de apertura de datos en el ámbito de la Universidad Nacional de Catamarca (UNCA), produciendo una contribución efectiva al proyecto de investigación denominado “Las TIC al Servicio del Dato Abierto: Situación Actual, Conceptualización e Iniciativas de Apertura de Información Pública” que se ejecuta en el Departamento de Informática. Asimismo los resultados obtenidos se pondrán a disposición de la institución para su conocimiento y análisis, y se brindarán recomendaciones que permitan reforzar la transparencia y relevancia social de los datos abiertos con la debida protección de los mismos en cumplimiento de la legislación vigente.

1. INTRODUCCIÓN

En los últimos años el concepto de gobierno abierto introduce una nueva forma de relación entre la sociedad, el gobierno y la administración, en el que las Tecnologías de la Información (TI) cobran un papel central. (Aparicio y otros, 2016) Si para los gobiernos, la publicación de datos abiertos es un imperativo, para las universidades –en especial las públicas– su adopción emerge no sólo como un deber, sino como una oportunidad de extender su quehacer con la comunidad y establecer vínculos con otras instituciones de educación superior, beneficiándose también, de esta manera, de la reutilización de los datos de investigación científica. (Morales Vargas A., 2019)

Las universidades, como instituciones públicas, tienen una responsabilidad social elevada. De hecho, debido a que una de las misiones de la Universidad es contribuir, de manera general, al bienestar de la sociedad, es

preciso considerar los conceptos de transparencia y datos abiertos. Por tanto, las universidades deben desempeñar diferentes acciones a favor de la transparencia para consolidarse como organismos abiertos de cara a la ciudadanía. (Aparicio y otros, 2016)

Si se trasladan los principios de gobierno abierto al ámbito universitario, surge la idea de “Universidad Abierta”. Una universidad abierta es aquella que favorece el acceso abierto a la información, escucha a la comunidad universitaria, tomando decisiones según sus necesidades y prioridades, y genera espacios de colaboración para desarrollar servicios con todos los actores de la comunidad universitaria. Además, la apertura de datos en los diferentes dominios que articulan la universidad (docencia, investigación, biblioteca, archivo, etc.) aporta notables beneficios tanto a la comunidad universitaria y a la universidad como institución, como a la propia sociedad donde se inserta. Este carácter abierto ayudará a ampliar las fronteras del conocimiento y mejorará las bases de transparencia, colaboración y participación de la sociedad del futuro. (Aparicio y otros, 2016)

Sin embargo uno de los principales dilemas de estas iniciativas es mantener un balance entre la PDP y un enfoque basado en la apertura, transparencia y acceso a información. Este paradigma incluye el reconocimiento y consolidación del derecho a la privacidad y bajo este contexto, se ha presentado la necesidad de identificar alternativas viables que proporcionen seguridad jurídica respecto de la utilización, reutilización y redistribución de diversas categorías de datos específicos de los beneficiarios finales. (Cerna y otros, 2020)

En este sentido en el año 2000, en Argentina, se sancionó la Ley 25.326 Protección de los Datos Personales que garantiza que todas las personas puedan controlar su información personal asentada en bases de datos públicas y privadas, pudiendo a través del ejercicio del derecho de acceso, informarle sobre quién tiene sus datos, de dónde los obtuvo y con qué fin. (Senado de la Nación Argentina y otros, 2000)

El principal desafío es entonces lograr que los principios de acceso a la información que se publica y la PDP sean compatibilizados. Esto implica promover la utilidad de los datos, garantizando al mismo tiempo el adecuado cumplimiento por parte de los responsables de las bases de datos del deber de respetar los derechos que la ley reconoce a sus titulares. (Consejo para la transparencia, 2020)

A partir de los resultados obtenidos en el proyecto “Datos abiertos en Educación, evaluación de los alcances de las iniciativas existentes. Caso de estudio Universidad Nacional de Catamarca” (2018-2019), se pretende con la presente investigación evaluar las políticas de PDP en el marco de las iniciativas de apertura de datos en el ámbito de la Universidad Nacional de Catamarca, produciendo una contribución efectiva al proyecto de investigación denominado “Las TIC al Servicio del Dato Abierto: Situación Actual, Conceptualización e Iniciativas de Apertura de Información Pública” que se ejecuta en el Departamento de Informática. Asimismo los resultados obtenidos se pondrán a disposición de la institución para su conocimiento y análisis, y se brindarán recomendaciones que permitan reforzar la transparencia y relevancia social de los datos abiertos con la debida protección de los mismos en cumplimiento de la legislación vigente.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto es una investigación aplicada en la cual se realizaron estudios de carácter descriptivo y comparativo. Para la metodología de trabajo se establecieron métodos comparativos y descriptivos. La revisión bibliográfica y el análisis del estado del arte en Protección de Datos Personales (PDP), nos permitió determinar los parámetros para confeccionar la evaluación. A partir de los parámetros seleccionados por un lado se realizó un relevamiento de la utilización de ejes asociados a la PDP en las iniciativas de datos abiertos en el ámbito de la UNCA y, por el otro, se analizaron los compromisos institucionales en materia de

PDP. Esto permitió tener una imagen clara de lo que sucede con los datos personales en dichas iniciativas, cómo se almacenan y se tratan esos datos.

Con respecto a la obtención de la información se recolectaron muestras mediante la observación directa en las iniciativas de datos abiertos de la UNCA.

3. RESULTADOS OBTENIDOS/ESPERADOS

A partir de la sanción de la Ley Nacional N° 27.275 de Derecho de Acceso a la Información Pública y del Decreto 117/2016, mediante el cual el gobierno nacional impulsó el “Plan de Apertura de Datos”, los organismos encargados de desarrollar los planes para la implementación de iniciativas de datos abiertos deben centrarse además en las disposiciones establecidas en la Ley 25.326 de Protección de Datos Personales (LPDP), a través de la cual se disponen los derechos de los titulares de datos - usuarios y responsables de archivos, registros y bancos de datos, se trató entonces de responder la pregunta ¿Qué políticas de PDP se utilizan en las iniciativas de apertura de datos en el ámbito de la Universidad de Catamarca?

Con la legislación nacional vigente en materia de protección de los datos personales, se hace necesario jerarquizar como bien público la información producida por la universidad y garantizar la apertura responsable de datos lo cual requiere una adecuada gestión de la privacidad y la protección de los datos personales. Por lo que, a partir de la investigación, se brindó recomendaciones sobre la apertura responsable de datos en la UNCA, que apunten a dar cumplimiento con la normativa existente en cuanto a la PDP.

A partir de la Ley distintas provincias de la República resuelve adherir a la misma o bien redactar sus propias directrices. En el caso particular de la provincia de Catamarca, el Artículo 11° de la Constitución provincial establece: “La libertad que antecede [de expresión] comprende el libre acceso a las fuentes de información. Prohíbese el monopolio de la información gubernativa y el funcionamiento de oficinas de propaganda de

la labor oficial”. Sin embargo no existe ninguna adhesión a la Ley Nacional como así tampoco no se detectaron evidencias de algún proyecto de regulación en el tema.

Como ya se sabe, las Universidades, tanto públicas como privadas, gestionan grandes volúmenes de datos, por lo que hay un mayor riesgo de poner en peligro datos sensibles. Una de las principales fuentes de recolección de datos personales que existe en la UNCA, son las inscripciones para ingresar a las carreras que en ella se dictan. Cabe destacar que las distintas unidades académicas utilizan el sistema SIU para gestionar dichos datos. Otras de las fuentes son los formularios de inscripción a diferentes becas. Pero no tan sólo las bases de datos de alumnos son fuentes importantes de datos, también lo son los repositorios institucionales de investigación, datos curriculares del personal docente e investigador, datos de control de presencia y horario, datos económicos y financieros, entre otros.

Con respecto a la adhesión a la LPDP, en el caso particular de la UNCA, no se encontraron registros sobre adhesión a la misma como así tampoco no se visualizó en su página web y en el de las distintas unidades académicas, políticas de seguridad sobre el tratamiento de los datos personales. En el único caso en el que se pudo evidenciar alguna acción sobre la temática es en el Repositorio Institucional de Acceso Abierto (RIAA) de la Facultad de Tecnología y Ciencias Aplicadas (FTyCA), donde las producciones que se publican se lo hace bajo la Licencia Creative Commons Reconocimiento -NoComercial-CompartirIgual 4.0 Internacional. Se trata de una herramienta legal de carácter gratuito que permite a los usuarios usar obras protegidas por derecho de autor sin solicitar el permiso del autor de la obra.

Si bien las actividades realizadas en el marco del presente proyecto permitieron cumplir con los objetivos propuestos y el cronograma de actividades de acuerdo a lo previsto sin inconvenientes, se considera que es necesario dar continuidad a la investigación sobre el tema para impulsar a la UNCA hacia un

modelo de gestión transparente a través de una apertura de datos responsable dando cumplimiento al derecho que tiene todo ciudadano para acceder a información en manos del Estado o que haya sido creada con fondos públicos. Pero esta transparencia y acceso a la información, que produce la universidad, debe regirse por la Ley 25.326 de Protección de Datos Personales a fin de garantizar la protección integral de los datos personales asentados en los archivos, registros, bancos de datos, u otros medios técnicos de tratamiento de datos, sean éstos públicos, o privados destinados a dar informes, para garantizar el derecho al honor y a la intimidad de las personas, así como también el acceso a la información que sobre las mismas se registre, de conformidad a lo establecido en el artículo 43, párrafo tercero de la Constitución Nacional.

4. FORMACIÓN DE RECURSOS HUMANOS

Los conocimientos generados y experiencia adquirida mediante este proyecto permitirán luego una ampliación y aplicación de los mismos en otros procesos y dominios. Por lo tanto, es un tema que cuenta con un gran potencial de transferencia tecnológica.

Por otra parte, este proyecto aspira entre otras cosas a la formación de profesionales investigadores, y también a ser fuente de origen de trabajos de investigación de graduación y postgraduación.

El proyecto está integrado por 2 (dos) profesores adjuntos, 1 (un) docente auxiliar, y 1 (un) alumno. El mismo tiene como objetivo la formación de recursos humanos a través de:

- Participación en eventos científicos regionales, nacionales e internacionales de la especialidad, como congresos, simposios, seminarios y cursos, por ello el programa de capacitación y formación de recursos humanos, contempla las siguientes actividades:

- Participación en cursos de actualización y posgrado en el área de estudio.

- Participación de los integrantes del proyecto en la dirección, asesoramiento y evaluación de tesinas de grado de la carrera de Ingeniería

en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA.

- Desarrollo de tesis de posgrado por parte de los docentes integrantes.

En cuanto a la formación específica del equipo de docentes investigadores, se destaca que los mismos se encuentran abocados a actividades de capacitación y estudios de posgrado, para la elaboración de 2 (dos) tesis correspondiente a la carrera Maestría en Ingeniería del Software de la Universidad Nacional de San Luis.

5. BIBLIOGRAFÍA

Agencia de Acceso a la Información Pública. Dirección Nacional de Protección de Datos. CUADRO COMPARATIVO Ley 25.326 de Protección de Datos Personales - Mensaje 147/2018 Proyecto de Ley de Protección de Datos Personales. 2018.

Agencia de Acceso a la Información Pública. Protección de Datos Personales. Buscador del Registro Nacional de Bases de Datos Personales.

<https://www.argentina.gob.ar/aaip/datospersonales/reclama>

Aparicio J. M., A. Fuster, I. Garrigós, F. Maciá, J. N. Mazón, L. Vaquer, J. J. Zubcoff. Ecosistema de Datos Abiertos de la Universidad de Alicante - Universidad de Alicante. 2016.

Bernardi, Melisa Natalí. Ley 25.326 de Protección de Datos Personales. El derecho a la intimidad y el hábeas data como herramienta de protección y resguardo frente a Internet. Aplicabilidad en las provincias argentinas. Universidad Siglo 21. 2016.

Cámara De Diputados Del Honorable Congreso De La Unión. Secretaría General. Secretaría de Servicios Parlamentarios. México. Ley Federal de Protección de Datos Personales en Posesión de los Particulares. Nueva Ley DOF 05-07-2010.

Cerna M., C. F. López. Hacia una apertura de datos abiertos sobre beneficiarios finales. Banco Centroamericano de Integración Económica (BCIE) 29-09-2020. Comisión Europea. La Protección de los Datos en la Unión Europea. Reglamento General de

Protección de Datos (RGPD - UE) 2016/679. 2016.

Consejo para la transparencia. GUÍA para el resguardo de los datos personales en el desarrollo e implementación de Plataformas de Datos Abiertos por parte de los Órganos de la Administración del Estado. Junio 2020.

Cortés Víquez A. K. Entre la apertura de datos y el deber de proteger la privacidad de l Revista Sinergias #2 Un espacio para analizar, reflexionar, dialogar y divulgar. 08-03-2019.

Decreto 117/2016. Plan de Apertura de Datos. Argentina. 2016.

Gomez Morales, Gianncarlo. Docente del Diploma Internacional en Gestión de la Ciberseguridad y Privacidad en ESAN. Abril de 2019.

Morales Vargas, A. CAPÍTULO 1 “Datos abiertos y visualización de información en sitios web de universidades chilenas: una asignatura pendiente”. Visualización de la Información en el Campo del Periodismo y la Comunicación Social. Universidad de Chile. 2019.

Paullier, Valentina. Departamento de Asesoramiento Tributario y Legal KPMG Uruguay. Mayo 2020.

Puccinelli, Oscar Raúl. El Hábeas Data en las provincias argentinas y en la Ciudad Autónoma de Buenos Aires.

Presidencia de la República. Secretaría General Subjefe de Asuntos Jurídicos. LEY N° 13.709, DE 14 DE AGOSTO DE 2018.

Red Iberoamericana de Protección de Datos (RIPD).

(<https://www.redipd.org/es/la-red/organos-de-la-red-iberoamericana-de-proteccion-de-datos>)

Senado y Cámara de Diputados de la Nación Argentina. Ley N° 25.326. Protección de Datos Personales. Argentina. 2000.

Senado y Cámara de Diputados de la Nación Argentina. Ley N° 27.275. Derecho de Acceso a la Información Pública. Argentina. 2016.

Temperini, Marcelo. Datos Personales en las Universidades Nacionales Argentinas. Observatorio Iberoamericano de Protección de Datos. 2013.

Unidad Reguladora y de control de Datos Personales. Uruguay Los datos personales y su protección. 2008.

Desarrollo de Computadoras de a Bordo para Misiones Satelitales de Órbita Baja

☒

Fernando Asteasuain – Federico Collado - Fernando Pazos- Federico D’Angiolo- Manuel Dubinsky- Matías Loieuseau – Noelia Aparicio- Katherine Caceres –Leonardo Guanco
Contacto Principal: fasteasuain@undav.edu.ar
Emails: <InicialNombre.Apellido>@undav.edu.ar

Carrera de Ingeniería en Informática – Universidad Nacional de Avellaneda

☒

☒

☒

XXXXXXXXXX

La industria aeroespacial se ha convertido en los últimos años en una de las áreas de mayor potencial dentro del mundo software tecnológico [1-5]. En particular, el diseño y desarrollo de los denominados nano satélites ha crecido exponencialmente en los últimos dos años. Este tipo de satélites, con bajos costos de desarrollo y de producción, permiten realizar valiosas tareas de producción y recolección de datos en diversos dominios como agroindustria, estudios climáticos, biología especializada o estudios espaciales y astronómicos [1].

Es crucial en este contexto el cuidadoso y meticuloso análisis y desarrollo de las computadoras a bordo (*OBC, por OnBoard Computers en inglés*), las cuales están a cargo de coordinar todo el funcionamiento del nano satélite. Por ejemplo, aspectos clave a considerar son la cantidad de servicios que proveerá y cuánto consume cada uno de la batería del artefacto.

El presente proyecto está enfocado en la exploración, diseño y verificación de OBC para nano satélites de órbita baja.

Palabras claves: *Nano satélites, Computadoras de a bordo, software satelital.* ☒

XXXXXXXXXX

La presente investigación se encuentra enmarcada dentro del proyecto “Desarrollo de Computadoras de a Bordo para Misiones

Satelitales de Órbita Baja” desarrollado en la carrera de Ingeniería en Informática de la Universidad Nacional de Avellaneda.

El grupo de investigación está constituido por dos investigadores formado, tres en formación, y cuatro estudiantes de grado haciendo sus tesis de final de carrera y prácticas profesionales. El proyecto entra en 2023 en el segundo año de su desarrollo.

El estudio de OBC para nano satélites ha sido declarado dominio de relevante interés estratégico comprendido en el Plan Nacional de Ciencia, Tecnología e Innovación Argentina 2030.

El proyecto forma parte de un marco inter institucional consistente en poner en órbita un satélite argentino del tipo CubeSats [6-9], con el objetivo de promover un laboratorio espacial destinado a la provisión de servicios de *Internet de las Cosas (IoT, por Internet of Things en inglés)*. La Universidad Nacional de Avellaneda, a través de la carrera de Ingeniería en Informática será la encargada del estudio y selección de computadoras de a bordo para *Cubesats*. El grupo de investigación está trabajando conjuntamente con el Consejo Profesional de Ingeniería en Telecomunicaciones, Electrónica y Computación.

Finalmente, es importante destacar que existe también la posibilidad de transferencia y servicios al sector industrial y productivo.

1. Introducción

La industria aeroespacial se ha constituido en un polo de atracción, tanto para industria como para la academia. De acuerdo a un reporte de Stanley Morgan las ganancias esperadas en este sector son de 22 mil millones de dólares para 2024 y de 41 mil millones de dólares para 2029 [1,4].

En este contexto el desarrollo de los denominados nanos satélites se ha destacado debido a sus bajos costos de producción y lanzamiento. En los últimos 10 años el segmento de nano satélites creció por un factor de 10x, de apenas 20 lanzamientos en 2011 a casi 200 en 2019, y se estima que entre 1800 y 2400 serán lanzados en los próximos cinco años [2,3].

Uno de los nanos satélites más utilizados son los llamados *CubeSats*, un proyecto iniciado en la Universidad de Stanford en 1999. Más de mil misiones CubeSats han sido realizadas en los últimos 20 años [2,5].

Este tipo de satélites se utiliza principalmente para misiones de órbita baja [1-5,11-12]. Esto representa un gran desafío computacional con respecto al flujo de comunicación entre el satélite y los dispositivos de IoT encargados de recibir y enviar información desde y hacia el satélite.

Es importante en este punto mencionar que es primordial tener cobertura total para la interconexión de todos los actores involucrados, asegurando especialmente la comunicación con dispositivos IoT ubicados en zonas rurales o remotas [8].

Otro aspecto relevante es el diseño y análisis de la Computadora de a Bordo (OBC). Por ejemplo, cómo y cuánto de la batería del satélite consume cada servicio que se ofrece es un requerimiento fundamental a considerar.

Una OBC está a cargo del funcionamiento de todos los servicios del satélite, y de todos los sub sistemas que lo componen. También es la

responsable de manejar y dirigir todo el flujo de comunicación desde y hacia el satélite [8].

Dado este marco resulta indispensable el estudio del diseño de OBC para nano satélites y su posterior verificación formal con herramientas de Ingeniería de Software, el cual constituye el pilar de la presente investigación.

2. Líneas de Investigación y Desarrollo

En el presente trabajo se explorarán las siguientes líneas de investigación:

- Aplicación de métodos formales de Ingeniería de Software al diseño y verificación de OBC.
- Aplicación de métodos formales de Ingeniería de Software en sistemas basados en Internet de las Cosas.
- Garantizar la comunicación, interacción y protocolos desde y hacia el satélite.
- Explorar las demandas de hardware y software para OBC.
- Continuar y profundizar el desarrollo de herramientas de software que den aplicabilidad a los conocimientos adquiridos.
- Posibilidad de expandir, modificar, o adaptar sistemas operativos de OBC como FreeRTOS [13] o KubOS [14].
- Posibilidad de diagramar nuevas funcionalidades requeridas por un sistema operativo para una OBC.
- Estudio de Diseño Arquitectónico Innovadores para cumplir con todos los requerimientos de protocolos y comunicación de nano-satélites.

☒

3. Resultados Obtenidos/Esperados

El objetivo principal del presente proyecto de Investigación es el diseño y desarrollo de OBC para nano satélites.

Como resultados obtenidos se pueden mencionar tres hitos importantes dentro del desarrollo actual del presente proyecto de investigación.

Por un lado, se analizó formalmente el sistema operativo FreeRTOS [15]. A través de la utilización de una herramienta y lenguaje de especificación denominado FVS [18,19], desarrollado por el grupo de investigación, se logró modelar toda la funcionalidad de este popular sistema operativo para OBC. Esta exploración formal del sistema operativo es clave para el desarrollo saludable de OBC.

En segundo término, se desarrolló un análisis inicial de *brokers* o servidores dedicados a OBC para nano satélites [16]. Este estudio resultó clave para tener un entendimiento cabal de las posibilidades y recursos de cada opción disponible.

Finalmente, se realizó un experimento de tipo empírico también sobre performance de *brokers* y servidores para OBC. Los relevamientos de tipo empírico aportan una información sumamente valiosa a la hora de diseñar y desarrollar OBC [17].

Como resultados futuros esperados mencionar:

- a) Relevar y establecer qué tipo de servicios pueden estar disponibles o no en los satélites para que puedan garantizar su funcionamiento
- b) Estudio y exploración de sistemas operativos para OBC.
- c) Estudio y análisis de protocolos de comunicación para dispositivos IoT.
- d) Estudio pormenorizado de requerimientos de performance, consumo de energía, y de procesamiento, a nivel hardware y software.
- e) Aplicar herramientas y técnicas formales de la Ingeniería de Software para el desarrollo OBC.
- f) Explorar y analizar un estudio comparativo y selección de diferentes

computadoras de abordo (OBC) que ofrece el mercado [18,20-21].

- g) Estudio y selección del sistema operativo y firmware a embeber en la computadora de abordo [22-23].
- h) Divulgar los resultados de la investigación en congresos y revistas científicas de interés para los temas de la investigación.
- i) Consolidar los recursos humanos en inicios de las tareas de investigación como estudiantes avanzados de la carrera, mediante la participación activa en el proyecto o mediante la realización de tesis de final de carrera de grado.
- j) Dirección de tesis de final de carrera y supervisión de prácticas profesionales (PPS).
- k) Analizar y explorar posibilidades de transferencia, ya que se abren grandes posibilidades de generación de trabajo pues existe una demanda creciente e insatisfecha en este segmento de mercado. El modelo de negocios se sustenta en el desarrollo de conocimientos y habilidades que aplican a la industria satelital con un bajo costo.

4. Formación de Recursos Humanos

El presente proyecto buscará potenciar a los investigadores en formación y de posgrado.

En cuanto a los estudiantes de grado, se buscará en primer lugar que finalicen su tesis de grado o su práctica profesional.

Se los pretende motivar como para que puedan continuar dentro del proyecto mediante la realización de algún posgrado.

Finalmente, se pretende consolidar las líneas de investigación dentro de la carrera de Ingeniería en Informática de la Universidad Nacional de Avellaneda.

☒

5. Bibliografía

1. Saeed, N., Elzanaty, A., Almorad, H., Dahrouj, H., AlNaffouri, T. Y., & Alouini, M. S. (2020). Cubesat communications: Recent advances and future challenges. *IEEE Communications Surveys & Tutorials*, 22(3), 1839-1862.
2. Hanafi, A., Derouich, A., Karim, M., & Lemmassi, A. (2021, January). Design and Implementation of an Open Source and Low-Cost Nanosatellite Platform. In *International Conference on Digital Technologies and Applications* (pp.421-432). Springer, Cham.
3. Williams, C. and DelPozzo S. Nano Microsatellite Market Forecast - 10th Edition, Space-Works Annual Nano/Microsatellite Market Assessment, 2020.
4. Pressman, A. (2019). Why Facebook, SpaceX and dozens of others are battling over Internet access from space. *Fortune*.
5. Kulu E.. "Nanosatellite and cubesat database," 2019. [Online]. Available: <https://www.nanosats.eu/>
6. Danilo José Franzim Miranda, Maurício Ferreira, Fabricio Kucinskis, David McComas. "A Comparative Survey on Flight Software Frameworks for 'New Space' Nanosatellite Missions". *Journal of Aerospace Technology and Management* vol.11, 2019 Epub Oct 24, 2019 On-line version ISSN 2175-9146.
<https://doi.org/10.5028/jatm.v11.1081>
7. Sneps-Snepe, M., Pauliks, R., & Namiot, D. (2021). On Information Technology Issues in the Nanosatellite Era. In *Conference of Open Innovations Association, FRUCT* (No. 28, pp. 638-644).
8. Fernandez, L., Ruiz-De-Azua, J. A., Calveras, A., & Camps, A. (2020). Assessing LoRa for satellite-to-earth communications considering the impact of ionospheric scintillation. *IEEE access*, 8, 165570-165582.
9. Anantachaisilp, P., Muangkham, M., Punpigul, N., & Thammawichai, M. (2020). Store and Forward CubeSat using LoRa Technology and Private LoRaWAN-Server.
10. Alam, M., Khamees, A., Aboelnaga, T., Amer, A., Harbi, A., Alamir, M., & Elsayed, O. A. (2021, August). Design and Implementation of an Onboard Computer and payload for Nano Satellite (CubeSat). In *The International Undergraduate Research Conference* (Vol. 5, No. 5, pp. 361-364). The Military Technical College.
11. Tipaldi, M., Legendre, C., Koopmann, O., Ferraguto, M., Wenker, R., & D'Angelo, G. (2018). Development strategies for the satellite flight software on-board Meteosat Third Generation. *Acta Astronautica*, 145, 482-491.
12. Wenker, R., Legendre, C., Ferraguto, M., Tipaldi, M., Wortmann, A., Moellmann, C., & Rosskamp, D. (2017, June). On-board software architecture in MTG satellite. In *2017 IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)* (pp. 318-323). IEEE.
13. <https://www.freertos.org/>
14. <https://docs.kubos.com/1.21.0/index.html>
15. Asteasuain, F.: Formalizing operating systems for nano satellites on board computers. In: CONAIIISI (2022)
16. D'Angiolo, F., Collado, F., Mayer, R., Caporaletti, G., Loiseau, M., Aparicio, N., ... & Perez, H. (2022). Estudio y recomendación de Computadoras de Abordo para Cubesat. *Cartografías del Sur Revista de Ciencias Artes y Tecnología*, (16).
17. F. Pazos and F. Collado. "Performance evaluation of MQTT broker servers deployed in the cloud: an application on a Cubesat nanosatellite". Artículo a ser presentado en *IEEE Latin-American Conference on Communications 2022 (LATINCOMM 2022)*. Rio de Janeiro, Brasil, 30 November to 02 December 2022.
18. Asteasuain, F., Braberman, V.: Declaratively building behavior by means of scenario clauses. *Requirements Engineering* 22(2), 239-274 (2017), doi:10.1007/s00766-015-0242-2
19. Asteasuain, F., Calonge, F., Dubinsky, M., Gamboa, P.: Open and branching behavioral synthesis with scenario clauses. *CLEI E-JOURNAL* 24(3) (2021).
20. Zufelt, Brian, "CloudSat: IoT Approach to Small Satellite Ground Infrastructure", Tech.

report, University of NewMexico, 2018.
<https://digitalrepository.unm.edu/>

21. Narayanasamy, A, et al. “Nanosatellites constellation as an IoT communication platform for near equatorial countries”, IOP Conference Series: Materials Science and Engineering, 2017.
<https://iopscience.iop.org/article/10.1088/1757-899X/260/1/012028>

22. Leppinen, Hannu. “Current use of Linux inspacecraft flight software”, IEEE Aerospace andElectronic Systems Magazine, number 32, pp. 4-13,2017.

23. Kelapure, Sarthak, “Design of a Software Architecture for Supervisor System in a Satellite”, Master Thesis, Facultad de Informática, Electrotécnica y Técnicas de la Información, University of Stuttgart, 2020.
DOI: 10.18419/opus-11167

Enfoques y Tendencias en el Desarrollo de Aplicaciones para Dispositivos Móviles.

Pablo Thomas , Leonardo Corbalán , Lisandro Delia , Juan Fernández Sosa ,

Fernando Tesone , Verena Olsowy , Patricia Pesado 

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

526 e/ 10 y 11 La Plata Buenos Aires

{pthomas, corbalan, ldelia, jfernandez, ftesone, volsowy, ppesado}@lidi.info.unlp.edu.ar

RESUMEN

Se presenta una línea de investigación y desarrollo centrada en el estudio de aspectos de Ingeniería de Software aplicados a las diferentes tendencias en el desarrollo de Aplicaciones para Dispositivos Móviles con resiliencia.

Palabras claves: Plataformas para Dispositivos Móviles – Aplicaciones Nativas – Aplicaciones Híbridas – Aplicaciones Interpretadas – Aplicaciones por Compilación Cruzada – Aplicaciones Web Progresivas – Offline First – Instant App – Resiliencia.

CONTEXTO

Esta línea de Investigación forma parte del Proyecto “*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital*” del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa y Latinoamérica.

1. INTRODUCCIÓN

Los avances tecnológicos en comunicación y computación han dado lugar al surgimiento de la “Computación móvil” [1]. El desarrollo de software para dispositivos móviles presenta grandes desafíos. Es necesario tratar con múltiples estándares, protocolos, tecnologías de red y diversas plataformas de hardware y software. Además, se deben considerar las

limitaciones de los dispositivos móviles en cuanto a procesamiento, consumo de energía y espacio de almacenamiento.

La computación móvil ha conducido a nuevas prácticas y metodologías que han promovido el crecimiento de la Ingeniería de Software como disciplina, acompañando este proceso de desarrollo tecnológico.

La competencia en el mercado de aplicaciones móviles es impulsada por la creciente cantidad de dispositivos móviles y empresas que migran sus servicios a la web. Para posicionarse en este sector, las empresas necesitan minimizar el tiempo de desarrollo de las aplicaciones y asegurarse de que funcionen en la mayor cantidad posible de dispositivos. No obstante, la excesiva fragmentación de hardware y software, producto del alto número de dispositivos con sistemas operativos y plataformas de desarrollo propias, dificulta este objetivo.

Esta problemática puede abordarse por medio del desarrollo nativo para cada una de las plataformas destino, utilizando en cada caso los entornos de desarrollo, tecnologías y lenguajes propios de cada plataforma.

El desarrollo nativo de aplicaciones móviles ofrece diversas ventajas, como el acceso sin limitaciones a todas las características del dispositivo, un alto rendimiento y una interfaz gráfica coherente con la plataforma. Además, permite trabajar sin conexión a Internet, ejecutar procesos en segundo plano y distribuir las aplicaciones a través de tiendas en línea. Sin embargo, estas ventajas tienen un alto

costo: el código fuente no se puede reutilizar entre plataformas, lo que aumenta el esfuerzo y los costos de desarrollo, actualización y distribución de nuevas versiones.

Como alternativa al desarrollo nativo, existen los enfoques multiplataforma que permiten compartir el código fuente entre las versiones para las distintas plataformas [2]. La construcción de aplicaciones Web Móviles es uno de estos enfoques. Estas aplicaciones se diseñan para correr dentro de un navegador, se desarrollan con tecnologías web bien conocidas (HTML, CSS y JavaScript), no necesitan adecuarse a ningún entorno operativo móvil específico; su puesta en marcha es rápida y sencilla.

Las aplicaciones web móviles presentan desventajas en su rendimiento debido a los tiempos de respuesta dilatados por la interacción cliente-servidor. Además, las restricciones de seguridad limitan el acceso a todas las capacidades del dispositivo y la experiencia de usuario no es tan atractiva como en las aplicaciones nativas.

Un nuevo concepto ha surgido en los últimos años denominado Aplicaciones Web Progresivas (PWA por sus siglas en inglés). Una PWA es una aplicación web que utiliza las últimas tecnologías de los navegadores para ofrecer una experiencia móvil similar a una aplicación nativa. Buscan alto rendimiento, carga rápida, y resistencia a la pérdida de conexión (*offline first*).

La tecnología TWA de Google permite integrar PWAs en aplicaciones Android distribuidas desde Google Play Store. Las TWA ejecutan una pantalla completa de navegador web sin mostrar la interfaz del navegador.

El enfoque de desarrollo híbrido es otra alternativa multiplataforma basada en tecnologías web estándar. Estas aplicaciones se ejecutan en un contenedor web especial con mayor acceso a las características del dispositivo a través de una API específica. También es posible la distribución de las aplicaciones a través de las *app stores* correspondientes. Sin embargo, el uso de

componentes no nativos perjudica la experiencia de usuario y ralentiza la ejecución debido a la carga asociada al contenedor web.

El enfoque interpretado para el desarrollo de aplicaciones móviles se caracteriza por generar aplicaciones compiladas en su mayor parte a código nativo, mientras que el resto se interpreta en tiempo de ejecución. Estas aplicaciones se implementan de forma independiente de las plataformas mediante tecnologías y lenguajes como Javascript, Typescript y XML. Las interfaces nativas son una de las principales ventajas de este enfoque, aunque definir nuevas componentes puede ser complejo debido a la necesidad de crear abstracciones compatibles con múltiples plataformas.

Finalmente, las aplicaciones generadas por compilación cruzada también constituyen un enfoque de desarrollo multiplataforma. A partir de un único código fuente, estas aplicaciones se compilan de manera nativa creando una versión específica de alto rendimiento para cada plataforma destino.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Enfoques de desarrollo de Aplicaciones para Dispositivos Móviles.
- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones Nativas en Android [3].
- Aplicaciones Nativas en iOS [3].
- Aplicaciones Web Móviles.
- Aplicaciones Móviles Híbridas (PhoneGap [4], Ionic [5]).
- Aplicaciones Móviles Interpretadas (Appcelerator Titanium [6], NativeScript [7]).
- Aplicaciones Móviles generadas por compilación cruzada (Xamarin [8], Corona [9]).
- PWA, Instant App y Offline First.
- TWA (Trusted Web Activities),

- Análisis y estudio comparativo de requerimientos no funcionales, tales como rendimiento, consumo de energía, tamaño de software, entre otros, en los distintos enfoques de Aplicaciones Móviles.
- Experiencia de usuario en Aplicaciones Móviles generadas con distintos enfoques de desarrollo.
- Internet de las Cosas (IoT), Internet de Todo (IoE) y aplicaciones de sensado móvil o sensado urbano.
- Ciudades inteligentes: El rol de las aplicaciones móviles y de la tecnología *blockchain*.
- El rol de los dispositivos móviles en aplicaciones de *crowdsourcing* o colaboración abierta distribuida.
- El rol de las aplicaciones móviles como herramientas para hacer frente a la pandemia de COVID-19.
- Aplicaciones Móviles resilientes. Escenarios adversos: conectividad inestable, pérdida de conexión, ralentización del tráfico de red, ciberataques. Estrategias: *offline first*, almacenamiento local, memoria cache, *retry pattern* (reintento) *throttling pattern* (estrangulamiento), lineamientos para el desarrollo de apps móviles resistentes a ciberataques.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos/esperados se pueden resumir en:

- Se ha estudiado el impacto que tiene el enfoque de desarrollo elegido sobre el rendimiento de las aplicaciones construidas. Se consideraron las plataformas iOS y Android junto a varios *frameworks* de desarrollo multiplataforma. Los resultados y conclusiones fueron publicados en [10] y [11].
- Se desarrolló la aplicación móvil "*Informática UNLP*" [12] para mejorar la comunicación entre alumnos y docentes de

la Facultad de Informática de la Universidad Nacional de La Plata. La aplicación incluye una cartelera virtual personalizable y tecnología de realidad aumentada para informar sobre la ocupación de las aulas en tiempo real (ver fig. 1). Se encuentra disponible para las plataformas iPhone y iPad y Android.

- En [13] se estudiaron las implicaciones de la elección de un enfoque de desarrollo determinado sobre la eficiencia energética de las aplicaciones generadas. Se contemplaron aplicaciones con alta carga de procesamiento, reproducción de video y acceso a imágenes generadas con diversos enfoques de desarrollo.
- En [14] se han publicado los resultados de un estudio sobre el impacto que tienen 9 tecnologías distintas de desarrollo de aplicaciones móviles sobre 23 características de interés para la Ingeniería de Software.
- En [15] se estudió el modo en que distintos *frameworks* de desarrollo multiplataforma afectan el tamaño de la aplicación construida. Los resultados son relevantes dado que el espacio disponible se ha convertido en un recurso crítico para muchos usuarios.
- Se ha cuantificado el impacto que tiene la elección del *framework* de desarrollo sobre tres de los requerimientos no funcionales más demandados por los usuarios de dispositivos móviles: rendimiento, consumo de energía y uso de espacio de almacenamiento. Los resultados y conclusiones se publicaron en [16] y constituyen una guía para los desarrolladores al momento de elegir el *framework* más adecuado a sus expectativas.
- Las Aplicaciones Web Progresivas o PWA, por sus siglas en inglés, fueron analizadas en [17]. Dada su estrecha relación con las aplicaciones web móviles fueron comparadas con estas últimas explicitando diferencias y similitudes. Se concluyeron una serie de ventajas y

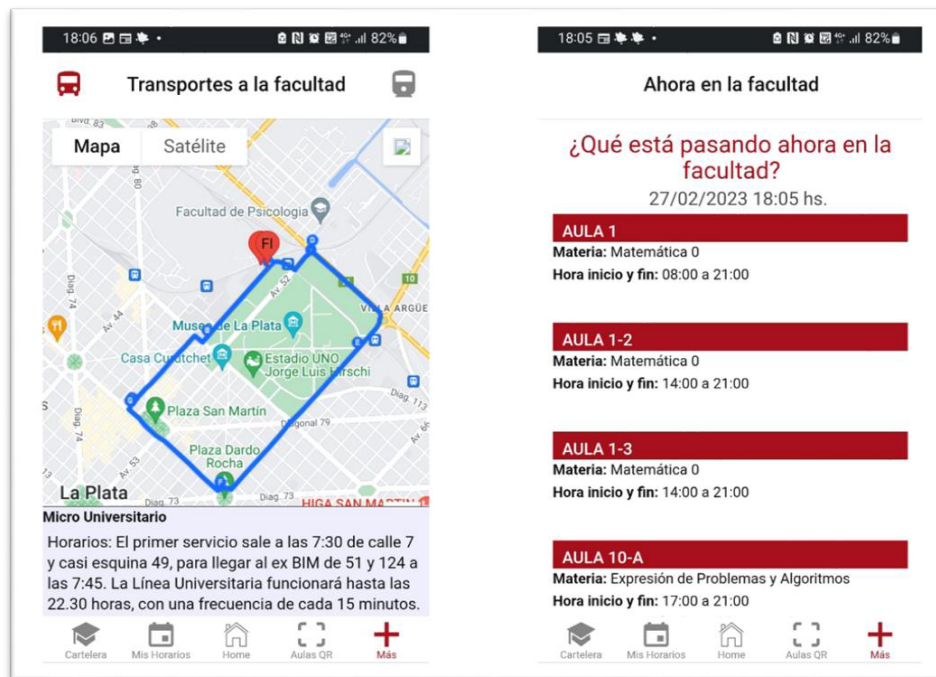


Figura 1. Informática UNLP – pantallas con información de utilidad.

desventajas de un enfoque respecto del otro.

- Se desarrolló “InnovApp”, una PWA que permite realizar una visita virtual guiada a la muestra de Ciencia y Tecnología 2019 llevada a cabo en el Centro de Innovación y Transferencia Tecnológica (CiyTT) de la Facultad de Informática de la UNLP
- Se ha realizado un estudio sobre el uso de la tecnología móvil para enfrentar la pandemia de COVID-19. Para ello se analizaron las aplicaciones promovidas por los gobiernos de 22 países distintos como herramientas tecnológicas para combatir la Pandemia. Los resultados y conclusiones se publicaron en [18].
- Se implementó un prototipo de sistema de sensado móvil participativo que permite al ciudadano monitorear la calidad del agua de red que llega a su domicilio y, en caso de anomalías, alertar a la comunidad y autoridades competentes [19].
- Se han elaborado recomendaciones específicas para aumentar la resiliencia de las aplicaciones móviles ante un aumento

inesperado de la carga computacional del dispositivo y el consumo de energía [20].

- Se continúa con el estudio y análisis de estrategias para aumentar la resiliencia de las aplicaciones móviles considerando distintas circunstancias de adversidad.
- Se continúa con el desarrollo de nuevas características de "*Informática UNLP*", una aplicación multiplataforma en continua evolución y crecimiento.
- Se prosigue con el estudio y el análisis del alcance de la tecnología de sensado móvil en aplicaciones de Internet de las Cosas (IoT), Internet de Todo (IoE) y sensado urbano.
- Se continúa con el análisis sobre el rol de las Aplicaciones Móviles en el contexto de las ciudades inteligentes y se agrega el estudio sobre el potencial de la tecnología *blockchain* para mejorar la calidad de vida de los ciudadanos. En este estudio se evaluará la capacidad de esta tecnología para ofrecer soluciones innovadoras en áreas clave, como la transparencia y la trazabilidad de datos y transacciones. Esto permitirá mejorar la confianza y la

eficiencia de sistemas modulares de una ciudad inteligente, como la gestión de residuos, el seguimiento de la energía y el agua, la gestión de identidad digital y la creación de mercados de energía descentralizados.

- Se examinarán las distintas formas de persistencia de datos en aplicaciones móviles en función del enfoque/*framework* de desarrollo utilizado y como estrategia para resistir ante la inestabilidad de las conexiones.
- Se Promoverá el avance sostenido y continuo de la formación de los miembros involucrados en esta línea de investigación.

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con este proyecto. Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

5. BIBLIOGRAFÍA

1. Hongxing Li, Guochu Shou, Yihong Hu, Zhigang Guo. Mobile Edge Computing: Progress and Challenges. 2016 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud). Oxford UK.
2. Spyros Xanthopoulos, Stelios Xinogalos, *A Comparative Analysis of Cross-platform Development Approaches for Mobile Applications*, BCI' 2013, Greece
3. Tracy, K.W., *Mobile Application Development Experiences on Apple's iOS and Android OS*, Potentials, IEEE, 2012.
4. <http://phonegap.com/>
5. <https://ionicframework.com/>
6. <http://www.appcelerator.com/>
7. <https://www.nativescript.org/>
8. <http://xamarin.com/>

9. <https://coronalabs.com/>

10. Delia L., Galdamez N., Corbalan L., Thomas P, Pesado P., *Un Análisis comparativo de rendimiento en Aplicaciones Móviles Multiplataforma*, CACIC 2015, UNNOBA Junín, Octubre 2015.

11. Delia L., Galdamez N., Corbalan L., Thomas P., Pesado P. *Approaches to Mobile Application Development: Comparative Performance Analysis* SAI Computing Conference (SAI), 2017. Londres, del 18 al 20 de Julio de 2017.

12. Fernández Sosa J., Cuitiño A., Thomas P., Delia L., Caseres G., Corbalán L., Pesado P. *"Informática UNLP" la App de la Facultad de Informática*. CACIC 2017. La Plata, del 9 al 13 de Octubre de 2017.

13. L. Corbalan, J. Fernandez Sosa, A. Cuitiño, L. Delía, G. Cáseres, P. Thomas, P. Pesado. *"Development Frameworks for Mobile Devices: A Comparative Study about Energy Consumption"*. 5th IEEE/ACM International Conference on Mobile Software Engineering and Systems. Gothenburg, Sweden. 27 y 28 de Mayo del 2018.

14. L. Delia, P. Thomas, L. Corbalan, J. Fernandez Sosa, A. Cuitiño, G. Cáseres, P. Pesado. *"Development Approaches for Mobile Applications: Comparative Analysis of Features"* SAI - Computing Conference 2018. London, United Kingdom. 10 al 12 de Julio de 2018.

15. J. Fernández Sosa, P. Thomas, L. Delía, G. Cáseres, L. Corbalán, F. Tesone, A. Cuitiño, P. Pesado. *"Mobile Application Development Approaches: A Comparative Analysis on the Use of Storage Space"*. XV Workshop de Ingeniería de Software - XXIV Congreso Argentino de Ciencias de la Computación CACIC 2018. Tandil, Argentina. 8 al 12 de octubre de 2018.

16. Corbalán, L., Thomas, P., Delía, L., Cáseres, G., Sosa, J. F., Tesone, F., & Pesado, P. (2019, June). A Study of Non-Functional Requirements in Apps for Mobile Devices. In *Conference on Cloud Computing and Big Data* (pp. 125-136). Springer, Cham

17. Aguirre, V., Ortu, A., Delía, L. N., Thomas, P. J., Corbalán, L. C., Cáseres, G., & Pesado, P. M. (2019). PWA para unificar el desarrollo Desktop, Web y Mobile. In *XXV Congreso Argentino de Ciencias de la Computación (CACIC)*(Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019).
18. Fernandez Sosa, J., Aguirre, V., Delía, L., Thomas, P., Corbalán, L., & Pesado, P. M. (2020). COVID-19: un análisis comparativo de Apps. In *XXVI Congreso Argentino de Ciencias de la Computación (CACIC)* (Universidad Nacional de La Matanza, 5 al 9 de octubre de 2020).
19. Fernández Sosa, J. F., Aguirre, V., Corbalán, L. C., Delía, L. N., Thomas, P. J., & Pesado, P. M. (2021). Sensado móvil como estrategia de participación ciudadana en Ciudades Inteligentes. In *XXVII Congreso Argentino de Ciencias de la Computación (CACIC)*(Modalidad virtual, 4 al 8 de octubre de 2021).
20. Corbalán, L. C., Fernández Sosa, J. F., Tesone, F., Dapoto, S. H., Cristina, F., Thomas, P. J., & Pesado, P. M. (2022). Recommendations for Resilient App Development. In *X Jornadas de Cloud Computing, Big Data & Emerging Topics* (La Plata, 2022).

CONSTRUCCIÓN DE COLECCIONES DE PROYECTOS DE SOFTWARE PARA ESTUDIOS EMPÍRICOS

Emanuel Irrazabal; Rubén Alfredo Bernal; Juan Andrés Carruthers; Andrea Lezcano Airaldi; Medina, YaninaM Martín Cardozo, Sebastián Bazterrica, Jeremías González

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
{eirrazabal, jacarruthers}@exa.unne.edu.ar

RESUMEN

Este es el primer año del proyecto F001-2021; una derivación de los proyectos F07-2009, F10-2013 y F018-2017, que desarrollaron modelos, métodos y herramientas para el desarrollo de aplicaciones software teniendo en cuenta aspectos de calidad de software. Este proyecto se enfoca en el diseño y desarrollo de estudios empíricos de calidad de software, como un insumo para la Ingeniería de Software basada en Evidencia.

La línea principal de investigación es la construcción de un catálogo de proyectos software de calidad. Se atiende la necesidad de los grupos de investigación para obtener muestras curadas de proyectos imprescindibles para la generación de resultados confiables y generalizables en la experimentación de estudios empíricos de la calidad de software, proporcionando los insumos y procedimientos necesarios para conseguirlo de manera efectiva.

Otras líneas secundarias de investigación son el desarrollo de estudios empíricos para la toma de decisiones en grupo y el desarrollo de aplicaciones orientadas entornos científicos.

Palabras clave: calidad de software, estudios empíricos, proyecto fuente abierta.

CONTEXTO

Las líneas de Investigación y Desarrollo presentada en este trabajo corresponden al proyecto PI-21F001 “Desarrollo de estudios empíricos en Ingeniería del Software”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2022-2025, y a la beca interna doctoral de CONICET otorgada por RESOL-2021-154-APN-DIR#CONICET para el período 2021-2025.

1. INTRODUCCIÓN

En el desarrollo de software un grupo de trabajo construye aplicaciones software con múltiples versiones [1]. Por lo tanto, muchas de las actividades asociadas con su mantenimiento provocan revisiones para mejorar la funcionalidad o para corregir errores, especialmente en los entornos de programación ágil [2].

En el desarrollo software, la calidad puede estudiarse desde el punto de vista de: i) la calidad del proceso de desarrollo software, y ii) la calidad del código fuente [3]. En este último caso es necesario obtener métodos empíricos para demostrar la calidad del software [4] y utilizar evidencia directamente relacionada con el producto software resultante a partir de métricas e indicadores que se vinculen directamente con la calidad [5]. Sin embargo, esto no siempre es sencillo de realizar, ya que las conclusiones generales de los estudios

empíricos en IS a menudo dependen de una gran cantidad de variables [6].

El uso masivo de repositorios para el código fuente (por ej., SourceForge, GitHub o Maven) le ha otorgado a los investigadores e ingenieros de software el acceso a millones de proyectos y, por lo tanto, datos para el desarrollo de estudios empíricos [7]-[9]. No obstante, la proporción de ruido en una muestra aleatoria tomada de repositorios podría sesgar el estudio, y puede llevar a los investigadores a conclusiones poco realistas, potencialmente inexactas [10].

Esto se contrapone con la condición de reproductibilidad y generalidad de los resultados empíricos, tal y como lo indica el énfasis actual en la IS basada en evidencia. En particular, la reproductibilidad es una condición necesaria, no solo en publicaciones en revistas o conferencias de prestigio de la disciplina, sino también para las empresas de desarrollo de software que a menudo desean analizar la evolución de sus propios proyectos o como patrón comparativo en auditorías de software. En este contexto, una práctica para demostrar la efectividad de las métricas como predictores de las características de calidad del software es la construcción de los denominados corpus o catálogos de proyectos [11].

Los catálogos de proyectos son un insumo para los grupos de trabajo y sirven como mecanismo de comparación para distintos tipos de experimentos. Existen varios ejemplos en la literatura, como los realizados por Barone y R. Sennrich [12], Allamanis y Sutton [13] o Keivanloo [14] y se diferencian por la cantidad, calidad de proyectos que lo componen; como también los criterios y métodos para agruparlos.

Un catálogo popular en IS es el Qualitas corpus [11], cuya última versión es de 2013. En general, se observa que no existe información actualizada en la mayoría de los catálogos, lo que hace necesario revisar los proyectos y sus versiones, y generar nuevas métricas sobre ellos. Una alternativa interesante para ello es la

utilización de herramientas de código abierto, que han demostrado ser de gran utilidad en entornos de trabajo profesionales, tanto tradicionales [15] como aplicadas a nuevas técnicas de desarrollo [16].

Por lo tanto, la línea de trabajo del proyecto tiene que ver con las características de calidad del código fuente del software.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se describe la línea de investigación y desarrollo:

Curaduría de proyectos

Se están estudiando los aspectos metodológicos y criterios considerados por la comunidad científica para conformar los catálogos de proyectos. Para esta etapa de documentación se utilizarán dos métodos de investigación. Por un lado, el método de revisiones sistemáticas [17] para identificar, evaluar, e interpretar toda la información relativa a un tema de investigación en particular, de un modo sistemático y replicable. Como segundo método de documentación se utilizarán las encuestas [18], para recopilar información de los grupos de investigación y los equipos de trabajo de las empresas privadas.

El siguiente paso consiste en la creación de un modelo de procedimientos para la construcción, mantenimiento y curaduría de un cuerpo de proyectos software y sus métricas de calidad de producto. Teniendo como fin proveer una estrategia para la construcción de estudios empíricos en IS que brinde una mejor reproductibilidad, consistencia experimental, y flexibilidad para una evolución gestionada del cuerpo de proyectos en el tiempo. Y, finalmente, se propone implementar el modelo en un ambiente real de trabajo.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el marco de este proyecto y respecto de la línea de curaduría de proyectos y de acuerdo con lo desarrollado en el año 2.022 en el marco del proyecto de investigación anterior se lograron los siguientes objetivos:

1. Mejora y utilización de herramientas para dar soporte en la generación de métricas obtenidas del análisis estático del código fuente.

Sonar Exporting Tool (SET), aplicación web para extraer las métricas de la plataforma Sonar Cloud (SC) en formatos de datos consumibles y difundir el código fuente de los proyectos software analizados. SET permite exportar medidas de métricas de SC en los formatos de datos csv, json y xml; la actualización automática y manual de la base de datos.

Sonar Juploader, aplicación java que permite analizar automáticamente un lote de proyectos con SC por medio del cliente SonarScanner. La aplicación da soporte a la gestión y uso de organizaciones de SC, pre-configuración de proyectos, análisis de proyectos, y visualización de reportes [19].

2. Relevamiento de información a considerar en la selección de proyectos Software y metadatos utilizados en el campo de estudio.

Se realizaron dos mapeos sistemáticos, uno enfocado en reunir evidencias sobre estudios empíricos en IS realizados con colecciones de proyectos [20], y el otro en el uso colecciones existentes en la literatura [21]. El objetivo en ambos casos es identificar las características de los proyectos Software, tipos de metadatos consumidos y las distintas herramientas usadas para recolectar los metadatos de

las colecciones. En total se revisaron 5.000 artículos científicos de las principales revistas y congresos sobre la temática.

En este sentido, se ha finalizado la construcción de un curso que facilite a alumnos de posgrado conceptos esenciales para la experimentación en IS, como también otras áreas del conocimiento. Desde la planificación y diseño de un experimento, hasta la utilización de herramientas estadísticas y visuales para reportarlos de manera efectiva.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo del Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 3 docentes investigadores, dos becarios internos doctorales de CONICET, dos becarios de investigación de pregrado y un estudiante de Licenciatura en Sistemas de Información.

5. REFERENCIAS

- [1] Parnas, D. L. (2001). Some software engineering principles. Hoffman, D. M. y Weiss, D. M. (Ed), Software fundamentals: collected papers by David L. Parnas (pp. 257–266). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
- [2] Irrazabal, E., Vásquez, F., Díaz, R., & Garzás, J. (2011). Applying ISO/IEC 12207: 2008 with SCRUM and Agile methods. In International Conference on Software Process Improvement and Capability Determination (pp. 169-180). Springer, Berlin, Heidelberg.
- [3] Lehman, M. M. (1996). Laws of software evolution revisited. En

- Montangero C. (Ed), *Laws of Software Evolution Revisited* (pp. 108-124). Berlin, Germany: Springer-Verlag.
- [4] Kitchenham, B. y Pfleeger, S. L. (1996). Software quality: the elusive target. *IEEE Software*, 13(1), 12-21.
- [5] Garvin, D. A. (1984). What Does "Product Quality" Really Mean?. *MIT Sloan Management Review*, 25-43.
- [6] Basili, V. R., Shull, F., & Lanubile, F. (1999). Building knowledge through families of experiments. *IEEE Transactions on Software Engineering*, 25(4), 456-473.
- [7] Vidal, S. A., Bergel, A., Marcos, C., & Díaz-Pace, J. A. (2016). Understanding and addressing exhibitionism in Java empirical research about method accessibility. *Empirical Software Engineering*, 21(2), 483-516.
- [8] Vidal, S., Bergel, A., Díaz-Pace, J. A., & Marcos, C. (2016). Over-exposed classes in Java: An empirical study. *Computer Languages, Systems & Structures*, 46, 1-19.
- [9] Vázquez, H. C., Bergel, A., Vidal, S., Pace, J. D., & Marcos, C. (2019). Slimming javascript applications: An approach for removing unused functions from javascript libraries. *Information and software technology*, 107, 18-29.
- [10] Munaiah, N., Kroh, S., Cabrey, C., & Nagappan, M. (2017). Curating github for engineered software projects. *Empirical Software Engineering*, 22(6), 3219-3253.
- [11] Tempero, E., Anslow, C., Dietrich, J., Han, T., Li, J., Lumpe, M., Melton, H y Noble, J. (2010). The Qualitas Corpus: A curated collection of Java code for empirical studies. *IEEE*, 336-345.
- [12] Barone, A. V. M., y Sennrich, R. (2017). A parallel corpus of Python functions and documentation strings for automated code documentation and code generation. *arXiv preprint arXiv:1707.02275*.
- [13] Allamanis, M., y Sutton, C. (2013). Mining source code repositories at massive scale using language modeling. *IEEE Press*, 207-216.
- [14] Keivanloo, I., Rilling, J., y Zou, Y. (2014). Spotting working code examples. *ACM*, 664-675.
- [15] Irrazábal, E., Garzás, J., & Marcos, E. (2011). Alignment of Open Source Tools with the New ISO 25010 Standard-Focus on Maintainability. In *International Conference on Software and Data Technologies* (Vol. 2, pp. 111-116).
- [16] Mascheroni, M. A., & Irrazábal, E. (2018). Continuous testing and solutions for testing problems in continuous delivery: A systematic literature review. *Computación y Sistemas*, 22(3), 1009-1038.
- [17] Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele University*, 33(2004), 1-26.
- [18] Kitchenham, B., & Pfleeger, S. L. (2002). Principles of survey research. *ACM SIGSOFT Software Engineering Notes*.
- [19] Pinto Oppido, J. A., Carruthers, J. A., & Irrazábal, E. A. (2021). Sonar JUploader, aplicación para el análisis, sincronización y actualización automática de proyectos a Sonar Cloud. In *IV Jornadas de Calidad de Software y Agilidad* (pp. 27-36).
- [20] Carruthers, Juan Andrés, Jorge Andrés Díaz-Pace, and Emanuel

Agustín Irrazábal. "How are software datasets constructed in Empirical Software Engineering studies? A systematic mapping study." 2022 48th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE, 2022.

- [21] Carruthers, Juan Andrés, Jorge Andrés Díaz-Pace, and Emanuel Agustín Irrazábal. "Estudios empíricos realizados con colecciones de proyectos software: un mapeo sistemático." *Memorias de las JAHIO 8.3* (2022): 7-20.

BUENAS PRÁCTICAS DE CALIDAD DE SOFTWARE EN EL ÁMBITO PÚBLICO

Emanuel Irrazabal; Gladys Dapozo; Cristina Greiner; Andrea Lezcano Airaldi; Juan Andrés Carruthers; María de los Ángeles Ferraro; Sebastián Bazterrica, Martín Cardozo, Jeremías González

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
{eirrazabal, alezcano}@exa.unne.edu.ar

RESUMEN

Esta línea se enmarca en el Proyecto F005-2021; una continuación de los proyectos F07-2009, F10-2013 y F018-2017 enfocados en el desarrollo de modelos, métodos y herramientas para la calidad del software teniendo en cuenta el entorno regional. Este proyecto aborda los temas emergentes en el área de la calidad de software, en particular, aspectos referidos a la gestión de proyectos en administraciones públicas regionales y el desarrollo de un modelo para la evaluación de visualizaciones con técnicas narrativas o de *storytelling*.

En particular, se está trabajando en el desarrollo de un modelo de Punto Digital Rural, la transformación digital de entidades públicas y la definición de un enfoque para la evaluación de la calidad y mejora asistida de visualizaciones aplicando técnicas narrativas.

Palabras clave: calidad de software, punto digital rural, visualización de la información.

CONTEXTO

Las líneas de Investigación y Desarrollo presentada en este trabajo corresponden al proyecto PI-21F005 “Personalización de buenas prácticas de calidad de software en la región”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2022-2025, la beca interna

doctoral de CONICET otorgada por RESOL-2021-154-APN-DIR#CONICET para el periodo 2021-2025 y el proyecto PFI 2022: “Centro de Integración Tecnológica y Digital en Entornos Rurales”.

1. INTRODUCCIÓN

Una de las líneas de trabajo es la personalización de procedimientos de calidad de software en las administraciones públicas regionales. Se ha venido trabajando con procedimientos de requisitos para empresas yerbateras, [1] el desarrollo de modelos para la entidad ferroviaria local [2] junto con sus procedimientos específicos [3] y en los sistemas de gestión de cuarentena de la provincia de Corrientes [4].

Esto evidencia la alta velocidad de la adopción de tecnologías digitales en entornos donde el “orgullo por lo analógico” era y es lo más común [5]. Esta nueva realidad está construida, en parte, por empresas y pequeños equipos de trabajo ágiles [6]. Sin embargo, no es igualmente fértil para todos; las entidades públicas tienen una tradición analógica y de poco uso de las nuevas tecnologías [7]. En el grupo de investigación se están diseñando un conjunto de procedimientos de soportes y específicos para el desarrollo de un área digital en entornos rurales, tal y como en la actualidad existe la iniciativa de Puntos Digitales [8].

Finalmente, otra línea de investigación es acerca de la narrativa basada en datos,

una estrategia efectiva para transmitir información, siendo su propósito estimular la atención y el compromiso (o *engagement*) de la audiencia y facilitar la toma de decisiones [9]. En el contexto actual, donde las posibilidades de toma de datos de distintos sistemas software ha tenido un crecimiento explosivo (Big Data), la visualización de datos proporciona un medio valioso por permitir un análisis asistido (por ejemplo, mediante controles de sumarización, división de los datos, búsquedas contextuales, o gráficos interactivos) [10]. Más aún, dado que muchas veces se utilizan para la toma de decisiones, es importante considerar las perspectivas de los usuarios finales en la construcción de dichas narrativas. Esta interrelación entre: i) las fuentes de datos que dan soporte a una visualización y ii) los intereses, preferencias y necesidades de los destinatarios de la misma, determina en gran medida la calidad de las narrativas a construir. No considerar estos aspectos suele ocasionar problemas relacionados con: baja calidad del contenido visual, pobre utilización de las fuentes de datos, ineficiencia en el desarrollo del contenido y, en general, poca entrega de valor.

En el contexto actual de desarrollos ágiles, uno de los problemas recurrentes es la poca calidad del contenido visual [11]. Por ello, la implementación de las buenas prácticas de *storytelling* en la construcción de un sistema y su posterior evaluación, pueden tener un impacto significativo en la calidad del producto resultante y en la entrega de valor por parte del sistema a los usuarios finales [12][13].

Se han realizado varios estudios que incluyen las prácticas y pautas generales a seguir para crear visualizaciones efectivas [10], [14][15] o [16]. Si bien se han propuesto algunos criterios para evaluaciones parciales, no se encuentran trabajos exhaustivos, herramientas que asistan a ello o propuestas sistemáticas de modelos para evaluar las narrativas basadas en datos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo principal del proyecto de investigación presentado es la mejora en la calidad del software y su aplicación en la región del nordeste mediante métodos, herramientas y experimentación para el desarrollo y mantenimiento de software; orientados a aumentar la eficiencia, confiabilidad y seguridad en los distintos ámbitos de aplicación de sistemas informáticos. Los objetivos particulares son los siguientes:

1. Elaborar métodos y herramientas la gestión de TIC en administraciones públicas regionales.
2. Desarrollar un modelo para la medición de calidad de visualizaciones de información teniendo en cuenta buenas prácticas de narrativa digital.

En este marco también se propone el desarrollo de transferencias tecnológicas y de formación de recursos humanos en temas relacionados con la mejora de la calidad de software hacia los equipos de trabajo de la región. Esto facilita los estudios de casos y los cuasi experimentos necesarios para la personalización de los desarrollos. Por otro lado, se busca la especialización de recursos humanos en temas relacionados con la calidad del software a través de la elaboración de tesis de carreras de grado, de posgrado y trabajos de investigación realizados en el marco de becas de investigación para alumnos y/o graduados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Transformación digital

La iniciativa de Puntos Digitales logra mejorar la alfabetización digital de los ciudadanos, pero en parte por su éxito, refleja la carencia de iniciativas similares en parajes rurales y localidades pequeñas

que no cuentan con la infraestructura necesaria para mantener un canal de comunicación digital con los ciudadanos o un punto de apoyo físico para el trabajo ciudadano-municipio que mejore el conocimiento y las oportunidades de los vecinos.

El objetivo general de la propuesta es el diseño, desarrollo e implementación de un Punto Digital para Entornos Rurales que cuente con el equipamiento, las condiciones de servicio y un sistema web específico que incluya redes sociales, guía de trámites, transparency, datos abiertos, y participación ciudadana, así como un servicio de acompañamiento para la ayuda a la digitalización y la reducción de la brecha digital.

Los objetivos específicos para el año 2.023 son:

- ☒ Desarrollar un modelo de Punto Digital para Entornos Rurales teniendo en cuenta la situación social y demográfica de las localidades y parajes con menos cantidad de habitantes de la provincia de Corrientes.
- ☒ Diseñar y desarrollar un sistema web personalizable orientado a Municipios con entornos rurales amplios, baja densidad poblacional y poca alfabetización digital, que contemple aspectos de datos abiertos, participación ciudadana y trámites digitales.
- ☒ Realizar la prueba piloto del modelo con la construcción de un Punto Digital en una localidad del interior de Corrientes, la personalización del sitio web y la medición del grado de alfabetización digital que se puede alcanzar.

Narrativa digital

En el marco de esta línea de trabajo y teniendo en cuenta la estrategia de investigación Design Science [17] se llevaron adelante diferentes estudios de

caso que han ido cimentando los resultados posteriores:

Estudio de caso [18] donde se determinaron los beneficios de incluir las buenas prácticas de visualización de datos en el desarrollo de un sistema realizado en el contexto de crisis sanitaria con tiempos límite restringidos. Para ello, se analizaron 16 gráficos de un sistema de información que gestiona el seguimiento del aislamiento y los permisos de circulación en la cuarentena por la pandemia de COVID-19.

Los resultados indican que el uso de técnicas de narrativa basada en datos contribuye a facilitar el proceso de toma de decisiones mediante la implementación de buenas prácticas de visualización que aumentaron la comprensión y la memorabilidad de los gráficos.

Estudio de caso [19] con el objetivo de desarrollar e implementar una aplicación web para un grupo de investigación inserto en una institución académica, utilizando una historia como hilo conductor para capturar y retener la atención de los usuarios.

Los resultados evidenciaron en general, una experiencia positiva al navegar el sitio para realizar las tareas propuestas. Esto podría deberse a que se planteó al usuario como el “protagonista” de la historia, apelando de esta manera a la interacción con el sistema para obtener información.

Se presentó y está pendiente de publicación una Revisión Sistemática de la Literatura con el objetivo de caracterizar las buenas prácticas en la construcción de visualizaciones de datos y las estrategias de evaluación reportadas al momento. En ese contexto, se finalizó la etapa de ejecución, para la cual se analizaron 96 estudios.

Dentro de esta línea de trabajo, se espera además, contribuir a la calidad de software a través del diseño y desarrollo de tableros de control con visualizaciones narrativas, bajo la hipótesis de que “la aplicación de estas técnicas mejorará la experiencia del usuario final y contribuirá

a una mejor comprensión de la información.”

Como trabajo futuro, se prevé desarrollar un modelo de medición de calidad de visualizaciones a partir de la evaluación de las buenas prácticas implementadas.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo del Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 4 docentes investigadores, 2 becarios internos doctorales CONICET y 1 becario de investigación de pregrado.

5. REFERENCIAS

- [1] Coimbra, Rodrigo Pereyra, et al. "Personalización de técnicas ágiles en el desarrollo de software para la obtención de requisitos de calidad en Pymes." *Electronic Journal of SADIO* 19 (2020).
- [2] Acevedo, Joaquín, Andrea Lezcano, and Emanuel Irrazábal. "Revisión sistemática de la literatura sobre implementación de arquitecturas software para sistemas críticos." XXVI Congreso Argentino de Ciencias de la Computación (CACIC)(Modalidad virtual, 5 al 9 de octubre de 2020). 2020.
- [3] Irrazabal, E., Sambrana, I., Pinto Luft, C., Gómez López, M. D. L. Á., Gallina, S. H., Laiuppa, A., ... & Lutenberg, A. (2018). Metodología de desarrollo de aplicaciones ferroviarias según las normas ISO 9001-EN 50126./Lezcano Airaldi, Andrea Fernanda, et al. "Estudio de caso: desarrollos de software en Argentina como soporte para detener la propagación de la pandemia COVID-19 durante el año 2.020." XV Simposio de Informática en el Estado (SIE 2021)-JAIIO 50 (Modalidad virtual). 2021.
- [4] Nadkarni, Swen, and Reinhard Prügl. "Digital transformation: a review, synthesis and opportunities for future research." *Management Review Quarterly* 71.2 (2021): 233-341.
- [5] Thiel, Peter A., and Blake Masters. *Zero to one: Notes on startups, or how to build the future*. Currency, 2014.
- [6] Ismail, Mariam H., Mohamed Khater, and Mohamed Zaki. "Digital business transformation and strategy: What do we know so far." *Cambridge Service Alliance* 10 (2017): 1-35.
- [7] Tekic, Zeljko, and Dmitry Koroteev. "From disruptively digital to proudly analog: A holistic typology of digital transformation strategies." *Business Horizons* 62.6 (2019): 683-693.
- [8] Franz, Fernanda López. "Más allá del acceso: desafíos para la inclusión digital. Un análisis del caso del Programa Punto Digital1." *Diseño y diagramación* (2022): 109.
- [9] Nussbaumer Knaflic, C. *Storytelling with Data: A data visualization guide for business professionals*. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2015.
- [10] Kosara, R., Mackinlay, J. *Storytelling: The Next Step for Visualization*. *Computer*, Vol 46, Issue 5, 44-50 2013
- [11] Nicoletti, M., Schiaffino, S., Díaz Pace, J. A.: An optimization-based tool to support the cost-effective production of software architecture documentation. *J. Softw. Evol. Process.* 27(9): 674-699 (2015).
- [12] Baker, R. *Agile UX Storytelling*. 2017.
- [13] Zimmerman, B. Applying Tufte's principles of information design to creating effective Web sites. *ACM SIGDOC Annu. Int. Conf. Comput. Doc. Proc.*, pp. 309-317, 1997, doi: 10.1145/263367.263406.
- [14] Kosara, R. An Argument Structure for Data Stories. *Proc. Eurographics/IEEE VGTC Symposium on Visualization (EuroVis)*, 2017.

- [15] Tufte, E. R. The Visual Display of Quantitative Information. 1983.
- [16] Ram D., et al. Case Report: Rapid Development of Visualization Dashboards to Enhance Situation Awareness of COVID-19 Telehealth Initiatives at a Multi-Hospital Healthcare System. Journal of the American Medical Informatics Association (2020).
- [17] Johannesson, Paul, and Erik Perjons. An introduction to design science. Vol. 10. Cham: Springer, 2014.
- [18] Lezcano Airaldi A., Diaz-Pace J.A., Irrazábal, E. (2021) Data-driven Storytelling to Support Decision Making in Crisis Settings: A Case Study. JUCS - Journal of Universal Computer Science 27(10): 1046-1068. <https://doi.org/10.3897/jucs.66714>
- [19] Lezcano Airaldi, A., Sandoval, D. & Irrazábal, E. (2021). Storytelling Aplicado al Diseño de Sitios Web: Un Estudio de Caso. En IV Jornadas de Calidad de Software y Agilidad (pp. 37-46).

Evaluación de lenguajes de modelado gráficos de procesos de negocio colaborativos utilizando un método multicriterio

Claudia Verino¹, Marisa Pérez¹, Juan Pablo Ferreyra¹, Diego Cocconi¹, Mario Berón²

¹ Departamento de Ingeniería en Sistemas de Información / Facultad Regional San Francisco / Universidad Tecnológica Nacional (UTN)

Av. de la Universidad 501, San Francisco (C.P. 2400), Córdoba, Argentina, (03564) 431019 / 435402
{cverino, mperez, jpferreyra, dcocconi}@sanfrancisco.utn.edu.ar

² Departamento de Informática / Facultad de Ciencias Físico-Matemáticas y Naturales / Universidad Nacional de San Luis

Ejército de los Andes 950, San Luis (C.P. 5700), Argentina
mberon@unsl.edu.ar

RESUMEN

La adopción actual de tecnologías como Internet y la computación en la nube permite a las organizaciones colaborar entre ellas, resultando en la conformación de redes colaborativas. Dicha colaboración se materializa a través de la definición y ejecución de procesos de negocio colaborativos.

La implementación de procesos de negocio colaborativos implica que las organizaciones puedan llevar a cabo las etapas del ciclo de vida BPM con ellos. Durante las etapas de análisis y diseño, las organizaciones deben definir los procesos colaborativos. En estas etapas se utilizan lenguajes que permiten modelar procesos colaborativos de una manera gráfica y agnóstica, independiente de la tecnología de implementación final.

Sin embargo, al momento de implementar procesos colaborativos en un grupo de organizaciones que conformen una red colaborativa, surge como interrogante si todos los involucrados estarán familiarizados con la notación para representar los modelos. Si se tiene que cambiar la misma en una etapa avanzada es necesario rehacer todos los modelos y la documentación, lo cual implicará costo adicional y/o retrasos.

El presente trabajo describe una línea de investigación para tratar el problema mencionado mediante de la utilización de métodos de toma de decisión multicriterio, definiendo un modelo que permita determinar la mejor opción de lenguaje de modelado gráfico disponible para procesos de negocio colaborativos, considerando todos los involucrados en la red colaborativa.

Palabras clave: gestión de procesos de negocio, redes colaborativas, procesos de negocio colaborativos, lenguajes de modelado, métodos multicriterio.

CONTEXTO

Este trabajo se enmarca en una propuesta de tesis para acceder al título de Magíster en Ingeniería de Software dictado por la Facultad de Ciencias Físico-Matemáticas y Naturales de la Universidad Nacional de San Luis. Dicha propuesta surge como una necesidad producto de las actividades desarrolladas durante la ejecución del proyecto de investigación I+D UTN 7844 “*Optimización organizacional de diferentes unidades de negocio autónomas aplicando modelos de redes colaborativas en PyMEs de la región*” (homologado como

proyecto de investigación y desarrollo por la Secretaría de Ciencia, Tecnología y Posgrado de la UTN). Ese proyecto fue presentado por el grupo de investigación “Grupo Gestión por Procesos” de la UTN Facultad Regional San Francisco (Resolución CD N° 57/2016). También se da continuidad a las actividades de investigación en el contexto del proyecto de investigación consolidado “Ingeniería de Software: Desarrollo de Estrategias de Mantenimiento y Migración de Sistemas de Software de Importante Envergadura” de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

1. INTRODUCCIÓN

Con la amplia adopción actual de tecnologías como Internet y la computación en la nube (*cloud computing*), las organizaciones se pueden ver beneficiadas de colaborar entre ellas, resultando en la conformación de **redes colaborativas** [1] [2].

Una red colaborativa está constituida por organizaciones heterogéneas, autónomas y geográficamente distribuidas que colaboran entre ellas para alcanzar objetivos comunes [3]. Este tipo de redes puede contribuir en forma importante al desempeño de las Pequeñas y Medianas Empresas (PyMEs) [4].

En las redes colaborativas, la colaboración entre organizaciones se materializa a través de la definición y ejecución de **procesos de negocio colaborativos** (*Collaborative Business Processes: CBPs*) [5] [6] [7]. Un proceso de negocio colaborativo, también llamado coreografía de proceso [7] o colaboración [8], especifica las interacciones (intercambio de mensajes) y coordina las acciones entre varias organizaciones participantes en búsqueda de objetivos de negocio comunes.

Los procesos de negocio colaborativos pueden considerarse y definirse según tres perspectivas diferentes [9] [6] [7]: (1) **modelo de coreografía**; (2) **modelos públicos**; y (3) **modelos privados**. El modelo de coreografía proporciona una vista global y pública de las interacciones entre las organizaciones, con sus respectivos roles, en un flujo único que define el orden de dichas interacciones. Los modelos

públicos (o procesos de interfaz) incluyen las actividades públicas y el comportamiento externamente visible de un solo participante (con su propio flujo de control), en términos de dar soporte al intercambio de mensajes con los demás. Desde esta perspectiva, el comportamiento del proceso colaborativo queda especificado por la interacción entre dichos modelos públicos. Finalmente, los modelos privados describen la lógica de negocio interna y privada de cada organización, así como las interacciones con las demás, para poder cumplir el rol que desempeña en el proceso colaborativo. Tales modelos son mantenidos por cada organización, conteniendo información confidencial que debe ser expuesta sólo a los interesados apropiados, y constituyen los verdaderos procesos ejecutables.

En la Figura 1 se muestra un ejemplo de un modelo de coreografía con dos organizaciones participantes, según un diagrama de coreografía de BPMN (*Business Process Model and Notation*), versión 2.0 [10]. Por otro lado, en la parte inferior de la Figura 2 se puede apreciar un modelo público de una de las organizaciones participantes y en la parte superior, un modelo privado de la otra organización, ambos expresados mediante un solo diagrama de colaboración de BPMN 2.0.

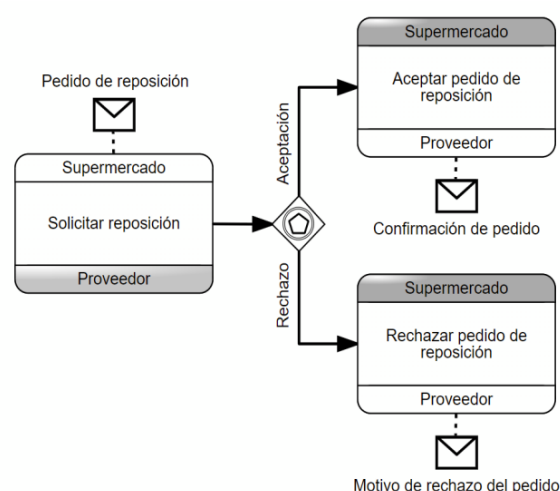


Fig. 1: Modelo de coreografía de un proceso colaborativo de ejemplo con dos organizaciones participantes, modelado en BPMN 2.0 (diagrama de coreografía), extraído de [11].

La implementación de las colaboraciones implica que las organizaciones puedan llevar a

cabo las etapas del *ciclo de vida BPM* (*Business Process Management*) con los procesos colaborativos en los que participen [12] [13]. Durante las etapas de *análisis* y *diseño*, que constituyen el alcance de esta propuesta, las organizaciones deben definir los procesos colaborativos, desde el punto de vista de las tres perspectivas mencionadas anteriormente. En estas etapas se utilizan lenguajes que permiten modelar procesos colaborativos de una manera gráfica y agnóstica, independiente de la tecnología de implementación, como el estándar BPMN 2.0, que proporciona diagramas de coreografía y de colaboración a tal fin (Figuras 1 y 2).

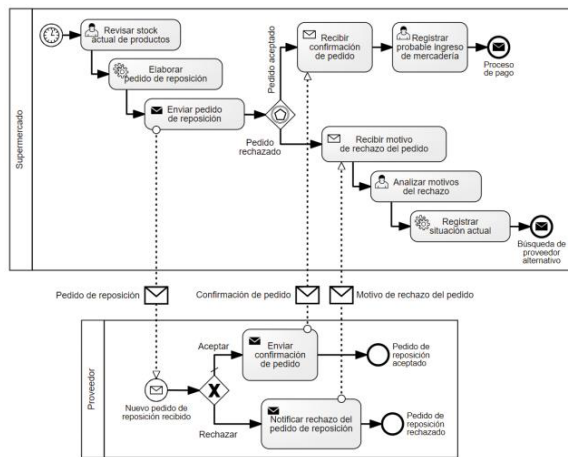


Fig. 2: Modelo público de una de las organizaciones participantes del proceso colaborativo de la Figura 1 (abajo) y modelo privado de la otra (arriba), modelados en BPMN 2.0 (diagrama de colaboración), extraídos de [11].

El estándar BPMN 2.0 es ampliamente aceptado y utilizado hoy en día tanto por la academia como por la industria para modelar gráficamente procesos intra-organizacionales, incluso los modelos públicos y privados de procesos colaborativos, en los que varias organizaciones intercambian mensajes [14] [15] [16]. Sin embargo, a pesar de su amplia difusión, existen desafíos aún. Por ejemplo, pese a que la versión 2.0 de BPMN cuenta con soporte para especificar modelos de coreografías, la gran mayoría de los constructores que más se conocen y utilizan son para especificar procesos intra-organizacionales [15].

En este contexto, al momento de implementar procesos colaborativos en un

grupo de organizaciones que conformen una red colaborativa, surge como interrogante si todos los involucrados estarán familiarizados con la notación para representar modelos de coreografía, comprenderán la semántica de sus constructores y serán capaces de interpretar patrones complejos. Tal como se mencionó anteriormente, los modelos de los procesos colaborativos y su documentación asociada deben ser generados de forma totalmente consensuada por parte de todas las organizaciones participantes [17] en una etapa temprana del ciclo de vida BPM (análisis y diseño). Es ampliamente conocido en Ingeniería de Software que los problemas de requerimientos, errores y cambios que se descubren con antelación son, por lo general, mucho menos costosos de corregir que si se encuentran en etapas posteriores del proceso de desarrollo [18] [19].

¿Cómo tener entonces cierta seguridad que todos los involucrados se comuniquen eficientemente, y se mantengan los modelos y la documentación de los mismos en forma consistente a lo largo de todo el ciclo de vida BPM una vez que se ha elegido una notación específica? Si se tiene que cambiar esta última en una etapa avanzada es necesario rehacer todos los modelos y la documentación, lo cual implicará costo adicional y/o retrasos. De este modo, se propone como objetivo abordar esa problemática mediante un **método de toma de decisión multicriterio** (del inglés *Multi-Criteria Decision Making: MCDM* [20]), definiendo un modelo que permita determinar la mejor opción disponible para modelar gráficamente procesos de negocio colaborativos, considerando todos los posibles involucrados en un grupo de organizaciones que conformen una red colaborativa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación principal abordada por este trabajo se centra en la gestión de procesos de negocio (BPM), la cual se encuentra entre las áreas prioritarias para el desarrollo de las actividades de I+D que se

formalizaron a través de la Resolución CD N° 510/2021 de la UTN Facultad Regional San Francisco. Dentro de esta línea, se hace foco en lenguajes de modelado gráficos para procesos de negocio y modelado de procesos de negocio colaborativos. También se aborda una línea de investigación en el ámbito de la Investigación de Operaciones, específicamente en lo que respecta a métodos de toma de decisión multicriterio, para analizar, evaluar y seleccionar el más apropiado de acuerdo con el objetivo propuesto descrito anteriormente. Existen muchos métodos de este tipo actualmente, y dentro de los más conocidos están bajo consideración: AHP (*Analytic Hierarchy Process*), ELECTRE (*ELimination Et Choix Traduisant la REalité*), PROMETHEE (*Preference Ranking Organization METHod for Enrichment of Evaluations*), LSP (*Logic Scoring of Preference*) [21] [22].

3. RESULTADOS OBTENIDOS/ESPERADOS

Producto de la investigación por concretar se espera definir un modelo mediante el método de toma de decisión multicriterio elegido, para seleccionar el lenguaje de modelado gráfico de procesos de negocio colaborativos adecuado, considerando todos los involucrados en un grupo de organizaciones que conformen una red colaborativa. Actualmente, la investigación se encuentra en una etapa inicial, analizando los métodos de toma de decisión multicriterio para seleccionar el que mejor se adecue a los criterios para poder evaluar los lenguajes de modelado gráficos de procesos de negocio colaborativos, tomando como base un trabajo anterior en el que se realiza un análisis preliminar de tales lenguajes y criterios [11].

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo (Grupo Gestión por Procesos) está conformado por docentes y estudiantes de la carrera de Ingeniería en

Sistemas de Información. Entre los docentes, tres de ellos se encuentran en la etapa de desarrollo de sus tesis de maestría y otro está preparando su tesis de doctorado (mención Ingeniería en Sistemas de Información), todos ellos con temas altamente vinculados a la línea de investigación descripta.

Además, como iniciativa del grupo, se pretende incorporar la experiencia y los conocimientos obtenidos a las cátedras de la carrera de Ingeniería en Sistemas de Información, y también se busca involucrar a estudiantes de la carrera en la realización de actividades del proyecto de investigación, incentivándolos a través de propuestas de becas. De la misma manera, se pretende realizar transferencias de tecnología de procesos a otras áreas de la Facultad, así como dirección y asesoramiento a empresas e industrias locales.

5. BIBLIOGRAFÍA

- [1] C.-M. Chituc, A. Azevedo y C. Toscano, «A Framework Proposal for Seamless Interoperability in a Collaborative,» *Computers in industry*, vol. 60, n° 5, pp. 317-338, 2009.
- [2] C. Durugbo, «Collaborative Networks: A Systematic Review and Multi-level Framework,» *International Journal of Production Research*, vol. 54, n° 12, pp. 3749-3776, 2016.
- [3] L. M. Camarinha-Matos, H. Afsarmanesh, N. Galeano y A. Molina, «Collaborative Networked Organizations – Concepts and Practice,» *Computers & Industrial Engineering*, vol. 57, n° 1, pp. 46-60, 2009.
- [4] B. Andres, P. Macedo, L. M. Camarinha-Matos y R. Poler, «Achieving Coherence between Strategies and Value Systems in Collaborative Networks,» de *Collaborative Systems for Smart Networked Environments: 15th IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2014, Proceedings*, Amsterdam, The Netherlands, 2014.

- [5] I. M. Lazarte, L. H. Thom, C. Iochpe, O. Chiotti y P. D. Villarreal, «A Distributed Repository for Managing Business Process Models in Cross-organizational Collaborations,» *Computers in Industry*, vol. 64, n° 3, pp. 252-267, 2013.
- [6] L. Amdah y A. Anwar, «BPMN Profile for Collaborative Business Process,» de *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, 2018.
- [7] W. Fdhila, D. Knuplesch, S. Rinderle-Ma y M. Reichert, «Verifying Compliance in Process Choreographies: Foundations, Algorithms, and Implementation,» *Information Systems*, vol. 108, p. 101983, 2022.
- [8] W. Fdhila, S. Rinderle-Ma, D. Knuplesch y M. Reichert, «Change and Compliance in Collaborative Processes,» de *2015 IEEE International Conference on Services Computing*, 2015.
- [9] D. Cocconi, J. Roa y P. Villarreal, «A Platform Based on Cloud Computing for Executing Collaborative Business Processes,» *CLEI Electronic Journal*, vol. 21, n° 2, 2018.
- [10] OMG, «Business Process Model and Notation (BPMN) Version 2.0,» Object Management Group (OMG), 2011.
- [11] D. Cocconi, M. Pérez, J. P. Ferreyra, C. Verino, G. Melano, N. Cocconi y A. Biasco, «Propuesta de un framework para la comparación de diferentes lenguajes de modelado gráficos de procesos de negocio en términos de la representación de procesos inter-organizacionales,» de *8vo. Congreso Nacional de Ingeniería Informática – Sistemas de Información (CoNaIISI 2020)*, San Francisco, Córdoba, Argentina, 2020.
- [12] M. Dumas, M. La Rosa, J. Mendling y H. A. Reijers, *Fundamentals of Business Process Management*, Second Edition, Springer-Verlag, 2018.
- [13] M. Weske, *Business Process Management: Concepts, Languages, Architectures*, Third Edition, Springer-Verlag, 2019.
- [14] F. Corradini, A. Morichetta, A. Polini, B. Re y F. Tiezzi, «Collaboration vs Choreography Conformance in BPMN 2.0: from Theory to Practice,» de *2018 IEEE 22nd International Enterprise Distributed Object Computing Conference (EDOC)*, 2018.
- [15] I. Compagnucci, F. Corradini, F. Fornari y B. Re, «Trends on the Usage of BPMN 2.0 from Publicly Available Repositories,» de *Perspectives in Business Informatics Research: 20th International Conference on Business Informatics Research, BIR 2021, Proceedings*, Vienna, Austria, 2021.
- [16] F. Corradini, C. Muzi, B. Re, L. Rossi y F. Tiezzi, «Formalising and Animating Multiple Instances in BPMN Collaborations,» *Information Systems*, vol. 103, p. 101459, 2022.
- [17] A. Saiqa, L.-M. Sanja y M. Nader, «Collaborative Business Process Modeling Approaches: A Review,» de *Proc. of the 2012 IEEE 21st International workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, 2012.
- [18] R. S. Pressman, *Ingeniería del software: Un enfoque práctico*, 7ma. edición, McGraw-Hill, 2010.
- [19] I. Sommerville, *Ingeniería de Software*, 9na edición, Pearson Educación, 2011.
- [20] J. J. Thakkar, *Multi-Criteria Decision Making*, Springer, 2021.
- [21] J. Dujmović y G. De Tré, «Multicriteria Methods and Logic Aggregation in Suitability Maps,» *International journal of intelligent Systems*, vol. 26, n° 10, pp. 971-1001, 2011.
- [22] M. Pereira Basílio, V. Pereira, H. Gomes Costa, M. Santos y A. Ghosh, «A Systematic Review of the Applications of Multi-Criteria Decision Aid Methods (1977–2022),» *Electronics*, vol. 11, n° 11, p. 1720, 2022.

DISEÑO DE UN MARCO DE TRABAJO PARA LA GESTIÓN DEL CONOCIMIENTO DE PRODUCTOS DE SOFTWARE EN EVOLUCIÓN

Laura Zeligueta, Elisa Galdame, Cintia Zacarías, Lucio Martino, Tiziana Angelucci, Joaquín Moyano, Eliana Fontana, Martín Villodas, Poblete Claudia

Instituto de Investigaciones
Facultad de Informática y Diseño
Universidad Champagnat

Mendoza, Argentina, +54 (261) 424-8443

zeligueta.laura@uch.edu.ar, galdame.elisa@uch.edu.ar, zacarias.cintia@uch.edu.ar, mlucio94@gmail.com, tittiangelucci08@gmail.com, joaquinmoyano857@gmail.com, elufontana@gmail.com, martinvillodas@gmail.com, cmapob@gmail.com

Resumen

En este proyecto se está trabajando en el establecimiento de un Marco de trabajo para la generación, adquisición, intercambio, protección, distribución y utilización de conocimientos propios de los productos software que se encuentran en evolución.[1]

Es el primero de esta línea de investigación en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat.

Palabras clave: gestión del conocimiento, software en evolución, marco de trabajo, desarrollo de software, buenas prácticas

Contexto

Dentro de las líneas de investigación del Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat, se encuentra la línea de “Ingeniería de Software”. En ella, se llevaron a cabo diversos proyectos dedicados al desarrollo de software orientado a aspectos.

El proyecto actual, que se denomina “Diseño de un marco de trabajo para la gestión del conocimiento de productos de software en evolución” se inició en setiembre del 2022 cuenta con el financiamiento de la Universidad Champagnat y la participación de la empresa ACP Ingeniería en Sistemas quien posee una

trayectoria de más de 30 años en la industria del software y tiene certificado el proceso de desarrollo de software bajo la norma ISO 9001:2015

Introducción

El proceso de evolución del software es conducido por peticiones de cambios. Estos cambios provienen de requerimientos existentes que no se han implementado en el sistema liberado, de peticiones de nuevos requerimientos, de reportes de bugs de los participantes del sistema, y de nuevas ideas para la mejora del software por parte del equipo de desarrollo del sistema.

Es posible considerar la implementación del cambio como una iteración del proceso de desarrollo, donde las revisiones al sistema se diseñan, se aplican y se ponen a prueba. Sin embargo, una diferencia crítica es que la primera etapa de implementación del cambio puede involucrar la comprensión del programa, tarea que puede presentarse como un reto difícil si *los desarrolladores del sistema original no son los responsables de implementar el cambio*. [2][3]

Es aquí donde se presenta el problema que nos ocupa en el presente proyecto. La explosión de la oferta laboral de alcance mundial en el sector de la industria de las TICs, incrementada durante la pandemia y sostenida en la actualidad, muestra una alta rotación en el

personal de las empresas desarrolladoras de software. Se suma a esto la tendencia de los jóvenes profesionales a convertirse en nómades digitales.

La situación descrita provoca una fuga constante del conocimiento y la consiguiente pérdida del conocimiento como capital de las organizaciones. Esto se resume en grandes dificultades y altos costos para lograr el mantenimiento y evolución de los productos software.

La propuesta del proyecto consiste en crear un Marco de trabajo para institucionalizar el conocimiento que emerge durante la evolución de un producto software partiendo de la base de conocimiento generada durante el desarrollo original.[4][5]

Cabe destacar que "la disciplina de gestión del conocimiento se centra en métodos sistemáticos e innovadores, prácticas y herramientas para tratar la generación, adquisición, intercambio, protección, distribución y utilización de conocimientos, capital intelectual y activos intangibles" (Montana, 2000).

Al respecto la norma ISO N° 30401-2018 – “Sistemas de gestión del conocimiento: Requisitos”, tiene como propósito ayudar a las organizaciones a desarrollar un sistema de gestión que promueva y permita de manera efectiva la creación de valor a través del conocimiento.[6]

Nuestra intención es tomarla como referencia para llegar a diseñar un marco de trabajo para la gestión del conocimiento enfocado a productos de software en evolución evitando la difusión espontánea del conocimiento y permitiendo mantenerse al día con el ritmo del cambio.[7][8]

Líneas de Investigación, Desarrollo e Innovación

Se trata de una investigación aplicada, por tanto, se procurará solucionar problemas reales apoyados en la investigación básica para conseguirlo. En cuanto al ámbito de aplicación se llevará a cabo en una empresa del medio, ACP Ingeniería en Sistemas, en donde se espera incorporar los resultados obtenidos

como una forma de preparación para la certificación en Gestión del Conocimiento.

Resultados y Objetivos

Se considera como Objetivo general la Creación de un marco de trabajo para institucionalizar el conocimiento que emerge durante la evolución de un producto software. En cuanto a Objetivos específicos se pueden enunciar los siguientes: Identificar el conocimiento del producto que debe ser almacenado, Identificar la información de los eventos de cambio o adaptación que deben ser almacenados durante la evolución del producto, Proponer herramientas para su almacenamiento y consolidación, Proponer prácticas que permitan la difusión interna y externa, así como la reutilización del conocimiento.

Así mismo se esperan obtener como resultados: Un conjunto de modelos que muestren las diferentes vistas del producto, Un diseño de registración de los eventos de cambio, Un conjunto de posibles repositorios, Un listado de buenas prácticas para institucionalizar la gestión del conocimiento.

Formación de Recursos Humanos

El equipo se encuentra conformado por investigadores formados, investigadores de apoyo, alumnos de la Lic. En Sistemas de Información que se encuentran cursando 2°, 3°, 4° y 5° año de la carrera. Además, se ha incorporado un docente externo perteneciente a la Ingeniería en Sistemas de Información de la UTN Reg.Mza.

Esta línea de trabajo permitirá generar oportunidades de formación en investigación, aplicada a la industria del software local. Se prevé el inicio de trabajos finales para la Licenciatura en Sistemas de Información.

Bibliografía preliminar

- [1] B. Rubenstein-Montano, J. Liebowitz, J. Buchwalter, D. McCaw, B. Newman, and K. Rebeck, “A systems thinking framework for knowledge

- management,” *Decis. Support Syst.*, vol. 31, no. 1, pp. 5–16, 2001.
- [2] Sommerville Ian, *Ingeniería del Software*, 9 Ed. .
 - [3] SWEBOK, *Guide to the software engineering body of knowledge*. 2004.
 - [4] A. H. Mohamed, “Managing evolution in software-engineering knowledge management systems,” *2006 1st Int. Conf. Digit. Inf. Manag. ICDIM*, no. January 2007, pp. 19–24, 2006.
 - [5] P. M. I. PMI, *PMBOK Guides*, 7 Ed. .
 - [6] International Standard International Standard, “International Standard ISO 30401:2018 Knowledge management systems — Requirements,” *Iso*, vol. 2018, p. 13, 2018.
 - [7] I. Nonaka, *Knowledge management: critical perspectives on business and management*. 2005.
 - [8] H. Nonaka, I. and Takeuchi, *The knowledge-creating company: how Japanese companies create the dynamics of innovation*, Oxford Uni. 1995.

ASPECTOS DE INGENIERÍA DE SOFTWARE, BASES DE DATOS RELACIONALES, Y BASES DE DATOS NOSQL PARA EL DESARROLLO DE SISTEMAS DE SOFTWARE HÍBRIDOS.

Luciano Marrero^{id}, Pablo Thomas^{id}, Ariel Pasini^{id}, Rodolfo Bertone^{id}, Eduardo Ibáñez^{id}, Verónica Aguirre^{id}, Marisa Panizzi¹, Verena Olsowy^{id}, Fernando Tesone^{id}, Patricia Pesado^{id}

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC

¹ **Universidad de Morón**

{lmarrero, pthomas, apasini, pbertone, eibanez, vaguirre, volsowy, ftesone, ppesado}@lidi.info.unlp.edu.ar,
¹marisapanizzi@outlook.com

RESUMEN

Se presenta una línea de investigación que tiene por objeto estudiar las problemáticas actuales que afronta los procesos de diseño de Ingeniería de Software y Bases de Datos (Relacionales y No Relacionales) ante una gran variedad aplicaciones multiplataforma utilizadas por millones de usuario simultáneamente. En la actualidad, la mayoría de las aplicaciones de software son multiplataforma[2, 3, 6]. El avance en la comunicación digital, el acceso constante a la información (internet) y el aumento en la utilización de tecnología móvil, genera que estas aplicaciones sean de uso común por millones de usuarios simultáneamente. Este contexto requiere nuevas formas de pensar los diseños de Bases de Datos en complemento con las tradicionales metodologías ágiles para el desarrollo de Software. Las Bases de Datos No Relacionales surgen como una alternativa de

solución a problemas no resueltos eficientemente por los Sistemas de Gestión de Bases de Datos tradicionales (Relacionales). El término NoSQL no responde a un sólo tipo de Bases de Datos, sino que representa un conjunto de tipos de Bases de Datos, con diferentes formas y características para representar la información [1, 9, 14]. Las metodologías tradicionales de diseño y construcción de Bases de Datos (principalmente las relacionales) han sido ampliamente desarrolladas, estudiadas y aplicadas por décadas. A pesar de los avances en los procesos de la Ingeniería de Software, en donde la documentación ha perdido protagonismo, estas metodologías continúan siendo un eslabón importante en la historia de las Base de Datos, como así también, para recabar requerimientos del usuario y del contexto.

La tecnología móvil impulsó aspectos que hace algunos años no eran considerados en el desarrollo del Software, tales como, movilidad, geolocalización, generación de grandes

volúmenes de información y la diversidad de los dispositivos electrónicos involucrados [1, 2, 5, 8].

Palabras claves: Diseño de Base de Datos, Metodologías ágiles, Bases de Datos Relacionales, Bases de Datos NoSQL, Bases de Datos como servicio en la nube, Aplicaciones Multiplataforma, Aplicaciones Móviles.

CONTEXTO

Esta línea de Investigación forma parte del Proyecto “*Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital*” del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación. Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID. Se utilizan los recursos de Hardware y Software disponibles en el III-LIDI para diseñar, desarrollar y probar diferentes soluciones a problemáticas relacionadas con escenarios a investigar. Como resultado de esto, se espera obtener métricas reales que sirvan como referencia para los investigadores en la comparación de resultados. Las publicaciones científicas generadas y la transferencia continua de resultados concretos, validan esta línea de investigación.

1. INTRODUCCIÓN

“Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la

sociedad actual: educación, ciudades inteligentes y gobernanza digital” del Programa de Incentivos, es un proyecto que propone profundizar las investigaciones que se vienen realizando en el III-LIDI y extender la mirada a nuevos desafíos y cambios que están en gestación. Se organiza en tres subproyectos que permiten atender de manera ordenada el objetivo general propuesto. “*SP1: Ingeniería de Software para escenarios híbridos*”, “*SP2: Ciudades Inteligentes sostenibles (CIS). Gobernanza Digital. Buenas prácticas y calidad*” y “*SP3: Creación de tecnologías digitales para el escenario educativo*”. [22].

Este artículo se centra en el subproyecto “*SP1 - Ingeniería de Software para escenarios híbridos*”. Se orienta a la investigación de metodologías y técnicas de la Ingeniería de Software, Bases de Datos (Relacionales y No Relacionales), Diseño de Base de Datos (Relacionales y No Relacionales), aplicaciones multiplataforma, escenarios híbridos y aplicaciones móviles, entre otras temáticas [22].

En la actualidad existen un conjunto de nuevos desafíos para la Ingeniería de Software y las Bases de Datos. Estos desafíos van desde las etapas de diseño y perduran hasta la puesta en producción de un producto de Software.

Generalmente, un proceso de diseño para una Base de Datos, inicia con el análisis de una especificación de requerimientos en lenguaje natural, la cual deriva de la mirada de expertos sobre una problemática real. La especificación generada, posee la ambigüedad propia del lenguaje, lo que introduce un cierto grado de subjetividad al momento de plantear o pensar en un esquema de datos. Los principios o reglas que se aplican a un esquema de datos Relacional generalmente no resultan apropiados

para una Base de Datos No Relacional (NoSQL). En general, NoSQL se basa en admitir redundancia y desnormalización para obtener eficiencia en las consultas que impactarán sobre el sistema. El modelo relacional y NoSQL proponen visiones diferentes, por lo tanto, su diseño también tendrá que ser diferente [1, 7, 8, 18].

El usuario actual es cada vez más exigente y pretende que las aplicaciones respondan de manera óptima todo el tiempo y desde cualquier punto geográfico en cual posea acceso a Internet. Con el objetivo de satisfacer tales pretensiones, muchas empresas desarrollan aplicaciones que son mundialmente utilizadas poniendo el foco en la escalabilidad de sus recursos y enfrentando nuevos retos para satisfacer la demanda de sus usuarios [2, 14].

Actualmente, existe un amplio abanico de tecnologías, tanto en Bases de Datos (Relacionales y No Relacionales), como así también, en lenguajes para el desarrollo de todo tipo de aplicaciones de Software [10, 11, 12].

En el contexto de Base de Datos, existen distintos motores de Bases de Datos Relacionales y No Relacionales para el desarrollo de aplicaciones móviles, entre ellos, podemos encontrar a:

SQLite, es una librería que implementa un motor de Base de Datos Relacional autocontenido (embebido). Tiene Licencia Public Domain, y puede ser utilizada tanto en el desarrollo nativo de aplicaciones (Android o iOS), como también en aplicaciones multiplataformas [10, 14].

Interbase (Relacional), es embebido, con licencia comercial y conforme al estándar SQL. Se puede utilizar en el desarrollo nativo en Android e iOS [10].

SQLBase (Relacional), tiene licencia de uso comercial. Se encuentra disponible para el desarrollo de aplicaciones móviles nativas, en Android e iOS [10, 14, 15].

Couchbase Lite (No Relacional, Documental), es un DBMS embebido, utiliza JSON como formato de los documentos. Tiene licencia dual, y es posible utilizarlo para el desarrollo de aplicaciones móviles nativas Android e iOS, y multiplataforma [10, 12, 13, 21].

Firebase Realtime Database (No Relacional, Documental), alojado en la nube, almacena los datos en un único JSON, y cuenta con sincronización de datos en tiempo real, manteniéndose disponibles aún sin conexión. Se encuentra disponible para el desarrollo de aplicaciones nativas en Android e iOS, y en el desarrollo de aplicaciones móviles [10, 15].

Google Cloud Firestore (No Relacional, Documental), es un DBMS alojado en la nube, almacena los datos en documentos JSON, cuenta con sincronización de datos en tiempo real y mantiene los datos disponibles aún sin conexión. Es posible utilizarlo en el desarrollo de aplicaciones móviles nativas Android e iOS, y en el desarrollo de aplicaciones móviles.

Oracle Berkeley DB (No Relacional, Clave-Valor), es una familia de Bases de Datos de Clave-Valor embebidas. Tiene licencia open source y se encuentra disponible para el desarrollo de aplicaciones nativas Android e iOS [10, 17, 18].

Como parte de las investigaciones realizadas, se plantea el diseño de Bases de Datos No Relacionales. Generalmente, cuando se trabaja con Bases de Datos que poseen almacenamiento no estructurado de datos, se plantea un diseño directamente de su estructura

a nivel físico teniendo en cuenta el problema a resolver y la flexibilidad de los esquemas. Realizar un esquema a nivel conceptual para una Base de Datos NoSQL es algo que aún es tema de discusión. Otra temática de investigación se obtiene de analizar el proceso de despliegue del Software. Las PyMES en Argentina representan casi el 80% de la industria del software y dada la necesidad de ser competitivas deben mejorar sus métodos y procesos de trabajo. En la mayoría de las empresas el proceso de despliegue no se realiza de manera sistemática y controlada, esto impacta en la finalización del proyecto y la no aceptación del producto final, ocasionando inconvenientes que generan rehacer el trabajo y baja de productividad en su proceso. Ante estas dificultades, se propone realizar un modelo de proceso de despliegue de sistemas de software que permita a las PyMES mejorar la ejecución del proceso de despliegue [4, 17, 18]. En este contexto, la comunicación y la sincronización entre los equipos de trabajo continúan siendo un pilar fundamental para el éxito de un proyecto. La utilización de repositorios de información y sincronización permiten realizar un control de versiones distribuido, trabajando en modo offline o en modo online, con la facilidad de disponer herramientas específicas para la resolución de conflictos entre versiones [3, 4, 6, 16].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Procesos de diseño para Bases de Datos No Relacionales [1, 2].
- Investigar distintos tipos de DBMSs para aplicaciones móviles [8, 10].

- Metodologías y técnicas de la Ingeniería de Software y su aplicación en el desarrollo de Software para escenarios híbridos.
- Desarrollo de casos de estudio y métricas de performance para distintos tipos de motores de Bases de Datos No Relacionales (NoSQL) como por ejemplo MongoDB, Apache Cassandra, Redis, Neo4j, entre otros[14].
- Desarrollo de casos de estudio y métricas de performance para Bases de Datos en la nube. Cloud Firestore (Google), MongoDB Atlas (MongoDB), DataStax Astra (Apache Cassandra), entre otras.
- Investigar, desarrollar casos de estudio e incorporar nuevas pruebas para la escalabilidad horizontal de motores de Bases de Datos No Relacionales.
- Investigar nuevos tipos de motores de Bases de Datos, por ejemplo: NewSQL y Bases de Datos de Serie Temporales.
- Repositorios para la sincronización del trabajo en equipo.
- Modelo de procesos para el despliegue / puesta en producción de sistemas de software

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados esperados/obtenidos se pueden resumir en:

- Capacitación continua de los miembros de las líneas de investigación.
- Estudio y análisis de procesos de diseño para Bases de Datos No Relacionales.

- Estudio y análisis de distintos DBMS Bases de Datos para el desarrollo de aplicaciones móviles [9, 10].
- Estudio y análisis sobre la sincronización de requerimientos no funcionales para una Base de Datos para aplicación móvil con una Base de Datos backend [9, 10, 11].
- Estudio y análisis de Bases de Datos en la nube y otros tipos de Bases de Datos como NewSQL, Series Temporales, entre otras [10, 11, 12].
- Análisis de métricas para diversos casos de estudio realizado entre distintos tipos de motores de Bases de Datos.
- Estudio y análisis sobre escalamiento horizontal para pruebas de performance entre distintos motores de Bases de Datos [13, 15, 20].
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT [11, 12, 20, 19].
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles [11, 12].

Algunas de las transferencias realizadas por el III-LIDI relacionadas con este proyecto, entre otras, son:

- Aplicación móvil para la comunidad de la Facultad de Informática de la UNLP con información sobre horarios de finales, planes de estudio, calendario académico, las clases en tiempo real y las últimas novedades.
- Diseño de Sistemas de Software y Bases de Datos para Instituciones Provinciales.

- Diseño y Gestión de Sistemas de congresos, utilizado por la RedUNCI y por otras entidades.
- Sistemas de Gestión Administrativa de Instituciones Universitarias (Sistema de inscripción y seguimiento de alumnos de la Facultad de Informática, Sistema de gestión administrativa de asignaturas, Sistema de gestión de asistencia de alumnos mediante QR, entre otros).

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además, participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

5. BIBLIOGRAFÍA

1. Un estudio de procesos de diseño de bases de datos NoSQL. Marrero Luciano, Osowy Verena, Tesone Fernando, Thomas Pablo Javier, Corbalán Leonardo César, Fernández Sosa Juan Francisco, Pesado Patricia Mabel. XXVIII Congreso Argentino de Ciencias de la Computación (CACIC) (La Rioja, 3 al 6 de octubre de 2022). ISBN: 978-987-1364-31-2. Páginas: 404-414.
2. Aspectos de ingeniería de software, bases de datos relacionales, no relacionales y como servicios en la nube para el desarrollo de sistemas de software híbridos. Marrero Luciano, Thomas Pablo Javier, Pasini Ariel

- Cristian, Bertone Rodolfo Alfredo, Ibañez Eduardo Javier, Aguirre Verónica, Panizzi Marisa Daniela, Olsowy Verena, Tesone Fernando, Pesado Patricia Mabel. XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza). ISBN: 978-987-48222-3-9. Páginas: 370-376.
3. Ingeniería de Software. Un Enfoque Práctico. Novena Edición. Roger S. Pressman, Bruce R. Maxim. McGraw-Hill Interamericana Editores, S.A. 2021. ISBN: 978-1-4562-8772-6.
 4. Ingeniería de Software Teoría y Práctica. Pfliegger Shari Lawrence. Pearson / Prentice Hall. 2002. ISBN: 9789879460719
 5. Ingeniería de Software. Novena Edición. Ian Sommerville. Addison Wesley / Pearson. 2011. ISBN: 978-607-32-0603-7
 6. Diseño Conceptual de Bases de Datos, un enfoque de entidades-interrelaciones. Carlo Batini, Stefano Ceri, Shamkant B. Navathe.. Addison-Wesley / Díaz de Santos. ISBN 0-201-60120-6 (1994).
 7. Model Driven Extraction of NoSQL Databases Schema: Case of MongoDB. Brahim, A.; Ferhat, R. and Zurfluh, G. (2019). In Proceedings of the 11th International Joint Conference on Knowledge Discovery. V.1: KDIR, ISBN 978-989-758-382-7, pages 145-154.
 8. *NoSQL*: modelos de datos y sistemas de gestión de bases de datos. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018). Universidad Nacional del Nordeste (UNNE). <http://sedici.unlp.edu.ar/handle/10915/67258>
 9. Análisis de performance en Bases de Datos NoSQL y Bases de Datos Relacionales. Luciano Marrero, Verena Olsowy, Fernando Tesone, Pablo Thomas, Lisandro Delia y Patricia Pesado. XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020). Universidad Nacional de La Matanza, del 5 al 9 de octubre del 2020. <https://cacic2020.unlam.edu.ar/es-ar/>
 10. Un Análisis Experimental de Sistemas de Gestión de Bases de Datos para Dispositivos Móviles. 2021. Tesone Fernando, Thomas Pablo, Marrero Luciano, Olsowy Verena, Pesado Patricia. XXVII Congreso Argentino de Ciencias de la Computación (CACIC) (Modalidad virtual, 4 al 8 de octubre de 2021). Universidad Nacional de Salta (UNSA), Argentina. <http://sedici.unlp.edu.ar/handle/10915/130353>
 11. Aspectos de ingeniería de software, bases de datos relacionales y bases de datos no relacionales para el desarrollo de sistemas de software en escenarios híbridos. Luciano Marrero, Pablo Thomas, Ariel Pasini, Rodolfo Bertone, Eduardo Ibañez, Verónica Aguirre, Verena Olsowy, Fernando Tesone, Patricia Pesado. XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020). Universidad Nacional de la Patagonia Austral (UNPA). El Calafate, Santa Cruz (Mayo 2020). <http://sedici.unlp.edu.ar/handle/10915/104026>.
 12. Aspectos de ingeniería de software, bases de datos relacionales, y bases de datos no relacionales y bases de datos como servicios en la nube para el desarrollo de software híbrido. Marrero Luciano, Thomas

- Pablo, Olsow Verena, Pasini Ariel, Bertone Rodolfo, Ibañez Eduardo, Aguirre Verónica, Panizzi Marisa, Pesado Patricia. XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja).
13. Aspectos de ingeniería de software y bases de datos para el desarrollo de sistemas de software en escenarios híbridos. Patricia Pesado, Rodolfo Bertone, Pablo Thomas, Ariel Pasini, Luciano Marrero, Eduardo Ibañez, Alejandra Rípodas, Verónica Aguirre, Verena Olsow, Fernando Tesone. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019). Universidad Nacional de San Juan (UNSJ). (Abril 2019). <http://sedici.unlp.edu.ar/handle/10915/77088>.
 14. Un estudio comparativo de bases de datos relacionales y bases de datos NoSQL. Pesado Patricia, Thomas Pablo, Delia Lisandro, Marrero Luciano, Olsow Verena, Tesone Fernando, Fernández Juan Sosa. XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019). Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019. ISBN 978-987-688-377-1. <http://sedici.unlp.edu.ar/handle/10915/91403>.
 15. A. Nori, "Mobile and embedded databases," in Proceedings of the 2007 ACM SIGMOD International Conference on Management of data, pp. 1175-1177, 2007.
 16. Schema Extraction and Structural Outlier Detection for JSON-based NoSQL Data Stores. M. Klettke, U. Störl, S. Scherzinger, in: BTW conf., 2015, pp. 425–444.
 17. Inferring Versioned Schemas from NoSQL Databases and Its Applications. Diego Sevilla Ruiz, Severino Feliciano Morales, Jesús García Molina. https://link.springer.com/chapter/10.1007/978-3-319-25264-3_35 (Springer).
 18. Aplicaciones para Dispositivos Móviles. Estrategias y enfoques de desarrollo. Thomas Pablo Javier, et. al. XXI Workshop de Investigadores en Ciencias de la Computación (San Juan 2019). ISBN: 978-987-3984-85-3
 19. Which Change Sets in Git Repositories Are Related?. Ramadani, J., Wagner, S. EEE 2016. International Conference on Software Quality, Reliability and Security (Viena, Austria.)
 20. Synchronization and replication in the context of mobile applications. STAGE, A. (2005 Joint Advanced Student School Course 6: Next-Generation User-Centered).
 21. S. Lee, "Creating and using databases for android applications," International Journal of Database Theory and Application, vol. 5, no. 2, 2012..
 22. III-LIDI: <http://weblidi.info.unlp.edu.ar/wp/proyectos/investigacion/>

Desarrollo de Software Público: Un aporte para la mejora de la calidad

Saldarini Javier*, Carrizo Claudio*, Armando Silvana*, Trasmontana Julio*, Mansilla Juan*, Ferreyra Gastón*

Salgado Carlos⁺, Sanchez Alberto⁺, Peralta Mario⁺

Facultad Regional San Francisco Universidad Tecnológica Nacional {saldarinijavier, ejcarrizo77, silvana.armando, julio.trasmontana, juampimansilla17, gastonferreyra49}@gmail.com

⁺Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Luis
e-mail: {csalgado, asanchez, mperalta}@unsl.edu.ar

RESUMEN

Los Organismos que pertenecen al Estado Argentino están convocados a llevar adelante distintos tipos de iniciativas en el marco del Plan de Modernización del Estado; uno de los canales operativos y de soporte para materializar esas iniciativas, lo constituye la Oficina Nacional de Tecnologías de Información (ONTI), específicamente dentro de la estructura de esta oficina existe el Equipo de Software Público, cuya misión es promover y facilitar el desarrollo del Software Público. El término de Software Público hace referencia a aquella solución tecnológica de software desarrollada, usada, modificada y distribuida, que permita facilitar el cumplimiento de objetivos gubernamentales y sociales, por el Estado, y para el Estado.

Dentro de este marco operativo existen distintos instrumentos que asisten a los Organismos para el desarrollo y mantenimiento de software, uno de ellos, es el Código de Buenas Prácticas en el desarrollo de software público, el cual consiste en una guía técnica para el desarrollo sustentable de software en la Administración Pública. Esta guía fue publicada en abril del año 2019.

Visto el marco descripto anteriormente, es que el presente trabajo aborda la necesidad de indagar sobre qué aspectos, y en qué medida, los instrumentos puestos a disposición para el desarrollo del Software Público, contemplan la calidad del software.

Específicamente el desarrollo de este trabajo tiene como principal objetivo desarrollar una propuesta que, desde la perspectiva de la calidad de productos de software, y tomando como marco de referencia lo establecido en la familia de normas ISO/IEC 25000, sirvan como soporte para la identificación y especificación de requisitos de calidad para el desarrollo del Software Público.

Palabras Clave: *Calidad del Software, ISO/IEC 25010, Estado Argentino, Requisitos de Calidad.*

CONTEXTO

La presente línea de I+D se enmarca en el Proyecto de Investigación: “*Calidad de productos de software: Un aporte para el desarrollo de Software en el Estado Argentino.*”

El mencionado Proyecto es una iniciativa conjunta entre la Universidad Nacional de San Luis, a través del equipo que conforma el Laboratorio de Calidad e Ingeniería de Software (LaCIS) – Facultad de Ciencias Físico-Matemáticas y Naturales, y la Universidad Tecnológica Nacional Facultad Regional San Francisco a través del Grupo de I+D “Calidad del Software”.

La evaluación, homologación y financiamiento del proyecto estuvo a cargo de la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, el mencionado

proyecto reconocido bajo el código: SIPPBSF0008185.

1. INTRODUCCIÓN

Es innegable que, en la última década, los avances tecnológicos en el área de las Tecnologías de la Información y las Comunicaciones (TIC's) han cambiado la forma en que las organizaciones llevan a cabo sus actividades para cumplir con sus objetivos. La Administración Pública Nacional (APN) es un tipo de organización que no ha estado ajena a estos cambios, y, según se describe en (GOBIERNO DIGITAL - Hitos significativos y evolución normativa período 1997- 2015) [1], el Estado Nacional Argentino ha venido desarrollando normativas que impulsan la modernización del estado nacional en pos de una gestión más eficaz y eficiente. Al respecto, el decreto presidencial N° 434 del año 2016, impulsa el denominado Plan de Modernización del Estado [2]; también en el III Foro Argentino de Transformación Digital del Estado, organizado por la Cámara de la Industria Argentina del Software (CESSI) [3] se presentaron distintas líneas de acción y casos de éxitos en organismos del Estado en materia de Gobierno Digital.

La modernización del Estado es un proceso continuo en el tiempo que presenta acciones concretas y específicas que buscan mejorar el funcionamiento de las organizaciones públicas.

En tal sentido, resulta necesario aumentar la calidad de los servicios provistos por el Estado incorporando Tecnologías de la Información y de las Comunicaciones, simplificando procedimientos, propiciando reingenierías de procesos y ofreciendo al ciudadano la posibilidad de mejorar el

acceso por medios electrónicos a información personalizada, coherente e integral [1].

Dentro de este esquema de modernización del Estado, los sistemas de información y el software son elementos estratégicos para lograr la transformación digital que se pretende, dado que muchos de los procesos operativos, de comunicación y de toma de decisión están automatizados a través de ellos.

En [4] se expresa que es imposible operar el mundo moderno sin software. Las infraestructuras nacionales y los servicios públicos se controlan mediante sistemas basados en computadoras.

Los Organismos que pertenecen al Estado Argentino están convocados para llevar distintos tipos de iniciativas en el marco del Plan de modernización del estado, uno de los canales operativos y de soporte para materializar esas iniciativas lo constituye la Oficina Nacional de Tecnologías de Información (ONTI) [5], específicamente dentro de la estructura de esta oficina existe el Equipo de Software Público, cuya misión es promover y facilitar el desarrollo del Software Público.

El término de software Público hace referencia a aquella solución tecnológica de software desarrollada, usada, modificada y distribuida para facilitar el cumplimiento de objetivos gubernamentales y sociales por el Estado y para el Estado [6].

Dentro de este marco operativo existen distintos instrumentos que asisten a los Organismos para el desarrollo y mantenimiento de software, uno de ellos, es el Código de Buenas Prácticas en el desarrollo de software público, el cual es una guía técnica para el desarrollo sustentable de software en la Administración Pública, esta guía fue publicada en Abril del año 2019 [6].

Como mencionamos anteriormente los sistemas de información han adquirido gran relevancia para las organizaciones y, asociado a ello, la calidad de estos sistemas informáticos también se ha convertido, hoy en día, en uno de los principales objetivos estratégicos de las organizaciones debido a que, cada vez más, sus procesos más importantes y, por lo tanto, la propia supervivencia de las organizaciones, depende de los sistemas informáticos, según se expresa en [6].

En la literatura se pueden ver diferentes definiciones referidas a calidad del software, entre ellas, se puede citar la propuesta en IEEE [7], en donde se expresa que: “La calidad del software es el grado con el que un sistema, componente o proceso cumple los requerimientos especificados y las necesidades o expectativas del cliente o usuario”.

También se observa que, en ISO/IEC 25010 [8], se la define como el Grado en que el producto software satisface las necesidades expresadas o implícitas, cuando es usado bajo condiciones determinadas.

R. Pressman [9] la define como el cumplimiento de los requisitos de funcionalidad y desempeño explícitamente establecidos, de los estándares de desarrollo explícitamente documentados, y de las características implícitas que se espera de todo software desarrollado profesionalmente.

También Sommerville [4] sostiene que la calidad del software es un concepto complejo que no es directamente comparable con la calidad de la manufactura de productos. En la manufacturación, la noción de calidad viene dada por la similitud entre el producto desarrollado y su especificación. En un mundo ideal, esta definición debería aplicarse a todos sus productos, pero, para sistemas de

software, existen cuestiones específicas que impiden aplicar este mecanismo.

Si bien parece no haber una definición única y aceptada universalmente al respecto, se puede observar que todas ellas mencionan que la calidad del software debería ser el cumplimiento y/o grado de satisfacción respecto de determinadas necesidades y/o requisitos dados, los cuales pueden ser explícitos y/o implícitos, entre otros.

Visto el marco descripto anteriormente, es que surge la necesidad de indagar sobre qué aspectos y en qué medida los instrumentos puestos a disposición por la Administración Pública contemplan la calidad para el desarrollo del Software Público.

Específicamente se pretende desarrollar una propuesta, que, desde la perspectiva de la calidad de productos de software sirvan como soporte para la identificación, evaluación, selección y especificación de requisitos de calidad como soporte al proceso de desarrollo del Software Público.

Si bien la problemática está focalizada en el dominio del desarrollo del Software Público, entendemos que también existen otras organizaciones, como, las empresas desarrolladoras de software o empresas que tienen sus propios equipos de desarrollo, que también podrían ser foco de estudios posteriores con el fin de validar los resultados de este proyecto en otro dominio de aplicación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de esta Línea de I+D están asociados a:

- ☒ Análisis exhaustivo de los documentos que dan soporte al desarrollo del Software Público (Código de buenas prácticas, Decálogo Tecnológico de la ONTI,

- etc.). A través de las metodologías de desarrollo de software propuesto
- ☒ Llevar a cabo un análisis desde la perspectiva de calidad de productos de software, específicamente tomando como referencia lo establecido en la familia de normas ISO/IEC 25000 [10], con el fin de determinar que aportes se pueden realizar desde esa perspectiva a los procesos de desarrollo del software público.
 - ☒ Desarrollo de modelos y/o artefactos que posibiliten arribar a una especificación de requisitos de calidad del software asociados por ejemplo al Código de buenas prácticas para el desarrollo de Software Público.

3. RESULTADOS OBTENIDOS/ESPERADOS

En general se espera que esta línea de I+D colabore con la mejora de la calidad de los productos de software desarrollados en los Organismos pertenecientes a la Administración Pública.

Se pretende indagar sobre los aspectos más determinantes que influyen en la especificación de requisitos de calidad en las distintas metodologías de desarrollo de software y los aportes que desde la perspectiva de la calidad de productos de software puedan hacerse en tal sentido.

De esta manera se espera generar modelos, métodos y/o artefactos que contribuyan para la identificación, evaluación, selección y especificación de requisitos de calidad de software que se desarrolla tanto en el ámbito público.

De manera específica para esta Línea de I+D se obtuvieron resultados parciales los cuales fueron publicados en distintos ámbitos, también cabe destacar que en este marco se finalizó una tesis para la obtención

del título de Magister en calidad del software, la tesis mencionada lleva el título de: *"Estrategia para Especificación de Requisitos de Calidad del Software"*, correspondiente al Ing. Javier D. Saldarini

Como se mencionó anteriormente los resultados alcanzados fueron publicados en los siguientes eventos:

Evento: JAIIO 2022 (Jornadas Argentinas de Informática)

Trabajo evaluado y aceptado: "Software Público Una propuesta para mejorar su calidad a través de ISOIEC 25000"

Evento: WICC 2022 (Workshop de Investigadores en Ciencias de la Computación)

Trabajo evaluado y aceptado: *"Hacia la Adaptación de Scrum para Incorporar Calidad de Datos en el Ámbito del Desarrollo Ágil."*

Evento: Jornada de Ciencia y Tecnología 2022

Trabajo expuesto: *"Modelo Conceptual basado en SQuaRE: Un Aporte para la especificación de Requisitos de Calidad del Software en Entornos de Desarrollos Ágiles"*

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de I+D el equipo de trabajo está conformado por:

2 Docentes Investigadores en formación:

- ☒ Ing. Silvana Armando
- ☒ Ing. Julio Trasmontana

1 Tesista de posgrado

- ☒ Ing. Claudio Carrizo

2 Becarios y tesistas de Grado:

- ☒ Juan Pablo Mansilla
- ☒ Gastón Ferreyra

1 Becario de Grado:

- ☒ *A designar en la convocatoria de becas de UTN para el año 2023*

Tesis de Maestría:

Título de la tesis: "Estrategia para Especificación de Requisitos de Calidad del Software"

Tesista: Ing. Javier Saldarini

Directores: Mg. Ing. Carlos Salgado y Mg. Ing. Alberto Sánchez

Título alcanzado: Magister en calidad del software

Universidad: Universidad Nacional de San Luis

5. BIBLIOGRAFÍA

- [1] Ministerio de Modernización, «GOBIERNO DIGITAL - Hitos significativos y evolución normativa,» 2018 . [En línea].
- [2] Ministerio de Modernización, «Decreto 434/2016 Plan de Modernización del Estado de la República Argentina.,» 2016. [En línea]. Available: https://www.argentina.gob.ar/sites/default/files/cofemod_documentos_plan_de_modernizacion_del_estado_2016.pdf. [Último acceso: 2019].
- [3] CESSI, «III Foro Argentino de Transformación Digital del Estado organizado,» 2018. [En línea]. Available: <http://www.cessi.org.ar/>. [Último acceso: 2019].
- [4] I. Sommerville, Ingeniería de Software, 9° ed., México: PEARSON EDUCACIÓN, 2011.
- [5] OFICINA NACIONAL DE TECNOLOGÍAS DE INFORMACIÓN, «<https://www.boletinoficial.gob.ar>,» 29 08 2019. [En línea]. Available: <https://www.boletinoficial.gob.ar/detalleAviso/primera/215325/20190902?busqueda=2>. [Último acceso: 2019].
- [6] ONTI, «Código de Buenas prácticas en el desarrollo de software público,» ONTI, 2019. [En línea]. Available: <https://www.argentina.gob.ar/onti/codigo-de-buenas-practicas-para-el-desarrollo-de-software-publico>. [Último acceso: 2021].
- [7] INSTITUTE OF ELECTRICAL ELECTRONIC ENGINEERING , *IEEE Std 1061-1998 IEEE Standard for a Software Quality Metrics Methodology –Description*, 1998.
- [8] ISO/IEC 25010:2011, *Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE)-System and software quality models.*, ISO, 2011.
- [9] R. Pressman, Ingeniería de Software., 6 ed., Mcgraw-Hill, 2005.
- [10] J. Calabrese, R. Muñoz, A. Pasini, S. Esponda, M. Boracchia y P. and Pesado, «Asistente para la evaluación de características de calidad de producto de software propuestas por ISO/IEC 25010 basado en métricas definidas usando el enfoque GQM,» de *XXIII Congreso Argentino de Ciencias de la Computación*, La Plata, 2017, pp. 660-671.

CULTURA, TURISMO Y TECNOLOGÍA. EXPERIENCIAS TURÍSTICAS SIGNIFICATIVAS para CIRCUITOS HISTÓRICOS de CALETA OLIVIA

Adriana MARTIN, Viviana SALDAÑO, Gabriela GAETAN, Claudia CARDOZO,
Mariela COLOMBRES, Hernán SOSA

Grupo de Investigación y Formación en Ingeniería de Software (GIFIS)
Instituto de Tecnología Aplicada (ITA)

Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia (UNPA-UACO)
adrianaelba.martin@gmail.com; vsaldanio@gmail.com; gabrielagaetan@yahoo.com.ar;
claudia_yoryi@yahoo.com.ar; mariela_colombres@hotmail.com; hassio_09@hotmail.com

RESUMEN

En este artículo se presentan avances y resultados del proyecto PDTS-UNPA denominado “*Revalorizando el Patrimonio Histórico-Cultural de Caleta Olivia: Una Aplicación Móvil para Circuitos Turísticos y Educativos*”. El proyecto pone foco en diseñar y desarrollar una aplicación móvil de turismo cultural para la ciudad de Caleta Olivia¹, situada en plena Patagonia Argentina², aplicando técnicas de RA y Geolocalización; estas técnicas y herramientas se incorporan con el objetivo de enriquecer el producto móvil, para generar experiencias turísticas significativas.

Palabras clave: *Patrimonio Histórico-Cultural | Aplicación Móvil | Realidad Aumentada | Geolocalización | Circuitos Turísticos y Educativos | Dominio del Turismo | Interacción Humano-Computadora (HCI) | Usabilidad | Accesibilidad | Experiencia de Usuario (UX) | Diseño Centrado en el Usuario (UCD)*

CONTEXTO

El presente proyecto, es desarrollado por integrantes del grupo GIFIS perteneciente al Instituto de Tecnología Aplicada (ITA)³, de la Unidad Académica Caleta Olivia (UACO)⁴, Universidad Nacional de la Patagonia Austral (UNPA)⁵. Este grupo se ha enfocado en el diseño, evaluación y desarrollo de productos Web y móviles, para incluir propiedades de

usabilidad y accesibilidad que mejoren la experiencia de los usuarios. La experiencia recabada en el área de diseño de interfaces y propiedades que propician la satisfacción del usuario les ha permitido llevar adelante la revisión y selección de recomendaciones para el diseño, la evaluación y definición de mejoras, la implementación de rediseño, y recientemente, el asesoramiento y el desarrollo propiamente dicho de productos Web y móviles. Estas actividades se han estado ejecutando en diferentes dominios, tales como: (i) Universidad, (ii) Salud y Discapacidad, (iii) Gobierno y, (iv) Patrimonio Histórico-Cultural en el Turismo, entre otros. Del trabajo realizado por GIFIS, podemos citar las siguientes publicaciones [1][2][4][6][7][8][9]. Referido específicamente al trabajo en el dominio del turismo cultural que requiere el presente proyecto, podemos citar los siguientes avances [3][5][10][16].

La entidad adoptante de este Proyecto es la Municipalidad de la ciudad de Caleta Olivia, a través de la Secretaría de Cultura, Deportes, Turismo y Juventud⁶.

1. INTRODUCCIÓN

La ciudad de Caleta Olivia es un destino emergente con gran potencial turístico. Turistas nacionales e internacionales, pasan por la región durante todo el año para emprender la aventura hacia el sur de la provincia, siendo la temporada de mayor afluencia de visitantes desde Diciembre a Febrero. Es fundamental, aprovechar este caudal de tránsito de viajeros para ofrecerles propuestas interesantes e

¹ Localidad de Caleta Olivia
<<https://www.santacruzpatagonia.gob.ar/centros-de-servicio/caleta-olivia>>

² Patagonia Argentina <<https://www.patagonia-argentina.com/>>

³ ITA-UACO <<https://itauaco.wixsite.com/itauaco>>

⁴ UACO <<http://www.uaco.unpa.edu.ar/uaco/>>

⁵ UNPA <<http://www.unpa.edu.ar/>>

⁶ Secretaría de Cultura, Deportes, Turismo y Juventud
<<https://www.caletaolivia.gov.ar/gobierno/ejecutivo/secretaria-de-deportes-cultura-y-turismo>>

innovadoras, que los atraiga y motive a quedarse y conocer un poco más de la ciudad. La Secretaría de Cultura, Deportes, Turismo y Juventud, ha buscado en los últimos años captar y atender a estos visitantes, definiendo una serie de actividades destinadas a fomentar los conocimientos más profundos de la localidad. Dentro de esas actividades, se cuenta con seis circuitos turísticos, destinados tanto a turistas como a vecinos y estudiantes de la localidad que quieran ver la ciudad con ojos de turistas, y así, conocer más de la ciudad donde viven. Estos circuitos turísticos son implementados, mediante caminatas guiadas por personal de la Dirección de Turismo, en fechas que se publican periódicamente. Además, existen otras propuestas que intentan promover el turismo local, tales como, programas de radio y charlas educativas en las escuelas.

Los objetivos en los que se enmarcan estos circuitos y propuestas son:

- 1º. Fomentar la sensibilización turística ciudadana consolidando el sentido de identidad y pertenencia local a la comunidad de Caleta Olivia.
- 2º. Propiciar actividades de esparcimiento y recreación turística destinadas a la población local, turistas y visitantes temporales.
- 3º. Poner en valor los principales atractivos de la ciudad promocionando y difundiendo su oferta turística.

En este contexto de fortalecimiento y mejora de la calidad de la oferta y servicios turísticos, GIFIS sitúa el desarrollo del presente Proyecto de Desarrollo Tecnológico y Social UNPA, PDTS-UNPA (2022-2024), denominado “Revalorizando el Patrimonio Histórico-Cultural de Caleta Olivia: Una Aplicación Móvil para Circuitos Turísticos y Educativos”.

A la hora de crear productos turísticos para un determinado destino, dicho producto debe [11]:

- ser una manifestación de la identidad cultural, de lo contrario, pierde su principal riqueza;
- ser una carta de presentación al cliente/usuario, siendo capaz de expresar cómo somos y de transmitir lo mejor;
- mantener su fidelidad para con el destino turístico, es decir, para con sus paisajes,

historia y cultura, gastronomía, costumbres, arquitectura, etc.;

- sustentarse en la diversidad, no solo ambiental, como es la moda hoy en día, sino también, social y cultural; y
- destacarse por la singularidad.

Ahora bien, ya mencionamos que un producto turístico se ofrece y/o se comercializa a través de diversos canales de distribución a sus usuarios, pero también por ser un producto, tiene un ciclo de vida.

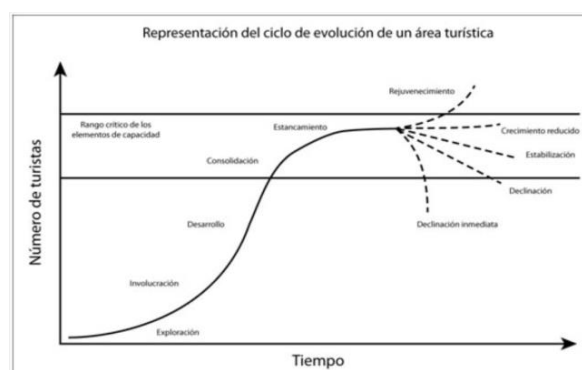


Figura 1: “Modelo Ciclo de Vida de los Destinos Turísticos (MCVDT)” [12]

Richard Butler⁷ propone en 1980 el “*Modelo del Ciclo de Vida de los Destinos Turísticos (MCVDT)* [12]” La Figura 1 ilustra este modelo evolutivo y dinámico debido a los cambios que generan de manera conjunta la oferta y la demanda de un producto turístico. El ciclo propuesto, consta de 6 etapas: exploración, implicación, desarrollo, consolidación, estancamiento y declive o rejuvenecimiento. Desde 1980 hasta la fecha, el MCVDT ha sido utilizado como principal herramienta para el estudio de la evolución de los destinos turísticos, generando desde 1980 hasta la fecha grandes debates; constantes han sido las publicaciones que le hacen referencia y a pesar de las continuas aportaciones de las que es objeto, este modelo mantiene su esencia y sigue siendo aplicado en destinos de todo el mundo [13].

Desde la perspectiva del producto turismo cultural, esta propuesta puede abarcar aspectos e

⁷ Fundador de la “International Academy for the Study of Tourism” <<http://www.tourismscholars.org/directory.php>> // Premiado por la “Organización Mundial del Turismo (OMT)” <https://www.hosteltur.com/lat/117520_richard-butler-distinguido-premio-ulises-omt.html>

implicancias de las áreas de cultura, historia, educación, sociedad y economía. De hecho, esta diversidad de áreas permite clasificar en tipos al producto turístico con enfoque cultural, entre los cuales, podemos mencionar los siguientes [15];**Error! No se encuentra el origen de la referencia.:**

- Comunitario, relacionado con la convivencia en comunidad, disfrutando in situ de su cultura y autenticidad.
- Indígena o Etnoturismo: interesado en destacar atracciones como las artesanías, festividades, gastronomía y folklore.
- Urbano: dirigido a visitar centros urbanos de distintas dimensiones.
- Rural: interesado en desarrollar alojamientos de estilo local en pueblos tradicionales o lugares cercanos donde puedan pernoctar los turistas, disfrutar de la cocina local, observar y compartir actividades populares.
- Agroturismo: derivado del turismo rural, al centrar su oferta también en el paisaje y tradiciones rurales.
- Ecológico o Ecoturismo: inspirado en la conservación de la naturaleza, la educación, la responsabilidad del viajero y la participación activa de la comunidad, generando conciencia ambiental.
- Religioso: relacionado con las personas que se desplazan por motivaciones de carácter religioso, destacando las peregrinaciones que los fieles realizan a lugares considerados santos o con alto valor espiritual.
- Patrimonial: enfocado principalmente en los bienes y expresiones patrimoniales como motivación turística.

Particularmente, el turismo patrimonial, se corresponde con el desplazamiento de personas motivadas por el interés en la historia y patrimonio de un lugar o zona específica. Por lo general, se tiende a confundir el turismo patrimonial y el turismo cultural. Si bien es cierto que la relación entre ambos es muy estrecha, es importante establecer la diferencia, ya que el turismo cultural representa un concepto más amplio que el turismo patrimonial.

Ahora bien, desde la perspectiva tecnológica, los productos turísticos digitalizados han sido ampliamente adoptados y habitualmente son

demandados por los visitantes para dar soporte y enriquecer la experiencia turística y así disfrutar de experiencias significativas. Un producto turístico se puede considerar una experiencia, si es capaz de ser: multisensorial, personal, generar conexión, único y cualitativo emocionalmente. Los diseñadores y clientes deben poner el foco en crear emociones positivas para proporcionar experiencias placenteras. El valor de las nuevas capacidades tecnológicas puede aportar a una integración y cohesión más transparente entre lo real y lo digital, haciendo interacciones más naturales, ágiles y adaptadas a las tecnologías que actualmente las personas dominan, tales como las tecnologías móviles y entornos Web [14].

El presente proyecto PDTs-UNPA, se enmarca en los lineamientos municipales de promoción del turismo local, establecidos por la Secretaría de Cultura, Deportes, Turismo y Juventud a través de la Subsecretaría de Turismo de Caleta Olivia. El producto a desarrollar se trata de una aplicación móvil con funcionalidad similar a la de una Guía Turística, que permita recorrer lugares y circuitos del patrimonio histórico cultural de la ciudad de Caleta Olivia. A los efectos de motivar el recorrido y descubrimiento de los puntos de interés, se incluirán recursos tecnológicos, tales como la Realidad Aumentada y Geolocalización. Así mismo, se evaluarán para el diseño del contenido, técnicas de Storytelling y Gamificación. Enmarcado dentro del paradigma ágil, se está definiendo el proceso de desarrollo a aplicar, tal que: (i) considere las características propias del producto (de turismo cultural) y del dominio, (ii) permita la retroalimentación para atender los requerimientos de las partes involucradas, (iii) entregue valor al cliente (Institución Adoptante) y, (iv) ofrezca una experiencia útil y placentera al usuario (ciudadanos de la región, turistas nacionales y extranjeros, estudiantes).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los integrantes de GIFIS afectados a este proyecto, se encuentran trabajando en los diferentes aspectos requeridos para el desarrollo integral del producto turístico a entregar: una

aplicación móvil para circuitos turísticos y educativos.

Desde el punto de vista tecnológico, el desarrollo de la aplicación móvil se realizará para la plataforma Android, utilizando lenguajes Java y Kotlin.

Desde el punto de vista del proceso de diseño, se llevó a cabo una revisión de los estudios existentes sobre Experiencia de Usuario (UX) en productos de software para Turismo Cultural con RA y de análisis de expertos en UX, para proponer un conjunto preliminar de Guías de Experiencia de Usuario destinadas a Aplicaciones de Turismo Cultural basadas en RA. Además, ya hemos mencionado los antecedentes recopilados por GIFIS tras años de investigación en el campo del diseño de Interfaz de Usuario (UI) con propiedades de calidad, tales como, Usabilidad, Accesibilidad, Interacción Humano-Computadora (HCI) y Diseño de Experiencia de Usuario (UXD).

Finalmente, desde el punto de vista del proceso de desarrollo, se está llevando a cabo un estudio y evaluación de las propuestas existentes para aplicaciones móviles, a los efectos de definir el enfoque dentro del espectro ágil, que responda a las características del dominio y requerimientos de las partes involucradas.

3. RESULTADOS OBTENIDOS/ ESPERADOS

El objetivo general de este proyecto es: *“Revalorizar el Patrimonio Histórico-Cultural de Caleta Olivia por medio de una aplicación móvil que ofrezca experiencias de usuario satisfactorias”*.

Los objetivos específicos (OE) son:

OE1. Categorizar requerimientos de usuarios de guías turísticas de Patrimonio Histórico-Cultural.

OE2. Seleccionar tecnologías de Realidad Aumentada y Geolocalización a incluir en la aplicación móvil.

OE3. Analizar la aplicación de técnicas Storytelling y Gamificación en el diseño de Experiencia de Usuario (UX).

Los resultados parciales obtenidos durante el primer año de ejecución de este proyecto son:

- ☒ En cuanto al objetivo específico OE1, se realizaron entrevistas a usuarios de la

entidad adoptante, y con los requisitos elicitados se elaboró un prototipo de la aplicación, que fue validado por los usuarios.

- ☒ En cuanto al objetivo específico OE3, se estuvo trabajando con la geolocalización. Se realizaron diversas pruebas con las librerías de Google Maps, para representar los puntos de interés y las rutas hacia esos puntos.

4. FORMACIÓN DE RECURSOS HUMANOS

En la actualidad, GIFIS tiene 9 integrantes: 6 docentes-investigadores UNPA, 1 docente-investigador invitado externo UTN-FRTDF y, 2 alumnos de grado UNPA. De estos integrantes, 5 están afectados al presente proyecto PDTs-UNPA.

Como ya se ha mencionado, los integrantes de GIFIS están en permanente capacitación y actualización para dar soporte a la ejecución de los proyectos en curso y ofrecer cursos de formación en las áreas y temáticas que se están trabajando.

5. BIBLIOGRAFÍA

- [1] Martin, A., Gaetán, G., Saldaño, V., Cardozo, C., Villagra, S., Sosa, H., Colombres, M. Experiencia del Ciudadano: Mejorando la Interacción en Contextos Digitales. XXIV Workshop de Investigadores en Cs. de la Computación (WICC); pp. 278-282. ISBN: 978-987-48222-3-9, <<https://wicc2022.uch.edu.ar/descargas/Libro-de-Actas-WICC-2022-1.pdf>>, 2022.
- [2] Gaetán, G., Villagra, S., Martin, A., Saldaño, V. Estrategia de Contenido para Mejorar la Experiencia de Usuarios de un Sitio Web Universitario. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, 7° EIPA, UNPA 2022.
- [3] Martin, A., Saldaño, V., Gaetán, G., Cardozo, C., Colombres, M., Sosa, H. Revalorizando el Patrimonio Histórico-Cultural de Caleta Olivia: Una Aplicación Móvil para Circuitos Turísticos y Educativos. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, 7° EIPA, UNPA 2022.

- [4] Cicchinelli Fernandez, J. A., Gaetán, G., Villagra, S., Martín, A., Saldaño, V. Rediseño Centrado en el Usuario para un Sitio Web de Gobierno: Un Caso de Estudio. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, 7° EIPA, UNPA 2022.
- [5] Cardozo, C., Martín, A., Saldaño, V., Gaetán, G. Aplicaciones de Turismo Cultural: Definiendo un Enfoque para el Desarrollo. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, 7° EIPA, UNPA 2022.
- [6] Sosa, H., Martín, A., Saldaño, V. Gaetán, G. Atendiendo Necesidades de Comunicación Complejas: Una Aplicación Saac Móvil para un Hospital Zonal. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, 7° EIPA, UNPA 2022.
- [7] Martín, A., Gaetán, G., Saldaño, V., Cardozo, C., Villagra, S., Sosa, H. Mejorando la Experiencia de Usuario en la Interacción con Productos Web y Móviles. VII Encuentro de Investigadores, Becarios y Tesistas de la Patagonia Austral, UNPA 2022.
- [8] Sosa, H., Martín, A., Saldaño, V., Gaetán, G. Desarrollando una Aplicación SAAC Móvil para el Hospital Zonal de Caleta Olivia. 1° Ateneo de Estudiantes del ITA, ISBN 978-987-3714-96-2, Editor (edición digital): UNPA-UARG, 2022.
- [9] Sosa, H., Martín, A., Saldaño, V., Necesidades de Comunicación Complejas: Desarrollando una Aplicación SAAC Móvil para el Hospital Zonal de Caleta Olivia. Libro de Actas del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022), pp. 920-924, ISBN: 978-987-1364-31-2, 2022.
- [10] Montaña Torrico, C., Saldaño, V., Martín, A. Integrando Herramientas en una Aplicación Turística para el Aprendizaje de Realidad Aumentada y Geolocalización. ICT-UNPA, Vol. 14(3), pp. 142-169, ISSN: 1852-4516, DOI: <<https://doi.org/10.22305/ict-unpa.v14.n3.900>>, 2023.
- [11] CEUPE. ¿Qué es un Producto Turístico? Centro Europeo de Postgrado <<https://www.ceupe.com/blog/que-es-un-producto-turistico.html>>, Accedido Marzo, 2023.
- [12] Butler, R. W. The Concept of a Tourist Area Cycle of Evolution: Implications for Management of Resources. In *Canadian Geographer / Le Géographe canadien*. Vol 24(1), pp. 5-12, Wiley online Library ISSN: 1541-0064, DOI: <<https://doi.org/10.1111/j.1541-0064.1980.tb00970.x>>, 1980.
- [13] Sánchez Valdés, A., Vargas Martínez, E., Castillo Nechar, M. Origen, Concepción y Tratamiento del Ciclo de Vida de los Destinos Turísticos: Una Reflexión En torno al Modelo de Butler. *Revista Científica Compendium*. Vol. 20(38), ISSN: 1317-6099, 2017. <<https://www.redalyc.org/articulo.oa?id=88051773005>>
- [14] Gomez Oliva, A., Server Gomez, M., Jara, A., Parra-Meroño, M. Turismo Inteligente y Patrimonio Cultural: Un Sector a Explorar en el Desarrollo de las Smart Cities. *International Journal of Scientific Management and Tourism*. Vol. 3(1), pp 389-411, 2017. Institutional Repository UCAM <<https://core.ac.uk/download/pdf/159419531.pdf>>
- [15] Concejo de Cultura (2011-10) Guía metodológica para proyectos y productos de turismo cultural sustentable. [en línea]. Disponible en: <http://repositorio.cultura.gob.cl/handle/123456789/4385>, Accedido Marzo, 2023.
- [16] Vega, M., Gaetán, G., Martín, A., Guías de Experiencia de Usuario para Aplicaciones de Turismo Cultural basadas en Realidad Aumentada. ICT-UNPA. Vol. 13(2), pp. 26-43 ISSN:1852-4516, DOI: <<https://doi.org/10.22305/ict-unpa.v13.n2.812>>, 2021.

DESARROLLO Y EVALUACIÓN DE PRODUCTOS WEB Y MÓVIL CENTRADOS EN LA EXPERIENCIA DE USUARIO

Adriana MARTIN, Gabriela GAETAN, Viviana SALDAÑO, Hernán SOSA, Silvia
VILLAGRA, Claudia CARDOZO

Grupo de Investigación y Formación en Ingeniería de Software (GIFIS)

Instituto de Tecnología Aplicada (ITA)

Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia (UNPA-UACO)

adrianaelba.martin@gmail.com; gabrielagaetan@yahoo.com.ar; vsaldanio@gmail.com;

hassio_09@hotmail.com; svillagra@uaco.unpa.edu.a; claudia_yoryi@yahoo.com.ar.

RESUMEN

En los últimos años, cada vez es más relevante la inversión en productos Web y móvil para que sus consumidores tengan una experiencia positiva. Ya no alcanza con desarrollar productos que sólo se enfoquen en el aspecto visual de la interfaz de usuario, sino que además, deben ofrecer una funcionalidad e interacción superiores, creando experiencias realmente satisfactorias. Mejorar la usabilidad, la accesibilidad y la interacción del usuario es fundamental, pero además, existen otros factores a considerar cuando se trata de diseñar una buena experiencia de usuario. Como si esto fuera poco, tecnologías tales como la geolocalización y la realidad aumentada, han irrumpido ofreciendo experiencias innovadoras al consumidor de productos digitales. Los investigadores se enfrentan al desafío de considerar todos esos factores y ofrecer propuestas para el desarrollo (creación, diseño e implementación), que permitan mejorar la experiencia inmersiva en su totalidad.

En este artículo se presentan avances y resultados del proyecto de investigación PI 29/B285 (2022-2025), denominado “Desarrollo y Evaluación de Proyectos Web y Móviles Centrado en la Experiencia de Usuario”. El proyecto utiliza los antecedentes del grupo de investigación, para proponer y aplicar estrategias de desarrollo con foco en la audiencia y su satisfacción.

Palabras clave: Experiencia del Usuario (UX) | Productos Web | Productos Móviles | Usabilidad | Accesibilidad | Estrategias de Contenido | Aplicación Móvil | Aplicación Web | Diseño Centrado en el Usuario (UCD) |

Realidad Aumentada | Geolocalización | Sistemas Aumentativos y Alternativos de Comunicación (SAAC)

CONTEXTO

A partir de proyectos vinculados y consecutivos, los integrantes de “Grupo de Investigación y Formación en Ingeniería de Software (GIFIS)”, han estado enfocados en mejorar las interfaces y la interacción de productos Web y móvil para contribuir a experiencias de usuario más satisfactorias. Particularmente, desde el PI 29/B256 (2020-2022), denominado “Contextos Digitales para Asistencia de los Ciudadanos: Enfoques de Experiencia de Usuario”, hemos estado enfocados en mejorar la experiencia del usuario ciudadano en diferentes dominios de aplicación. Actualmente, este trabajo continúa en el PI 29/B285 (2022-2025), denominado “Desarrollo y Evaluación de Proyectos Web y Móviles Centrado en la Experiencia de Usuario”, dirigido y codirigido por la Dra. Adriana Martín y la Mg. Gabriela Gaetán, respectivamente; y en el PDTs (2022-2024), denominado “Revalorizando el Patrimonio Histórico-Cultural de Caleta Olivia: Una Aplicación Móvil para Circuitos Turísticos y Educativos”, dirigido y codirigido por la Dra. Adriana Martín y la Mg. Viviana Saldaño, respectivamente. GIFIS, es un grupo de investigación perteneciente al Instituto de Tecnología Aplicada (ITA)¹, de la Unidad Académica Caleta Olivia (UACO)², Universidad Nacional de la Patagonia Austral (UNPA)³. Desde 2012 y a través de varios

¹ ITA-UACO <<https://itauaco.wixsite.com/itauaco>>

² UACO <<http://www.uaco.unpa.edu.ar/uaco/>>

³ UNPA <<http://www.unpa.edu.ar/>>

proyectos relacionados y consecutivos, los integrantes de GIFIS han estado trabajando en el diseño, evaluación y desarrollo de productos Web y móvil, para incluir propiedades de usabilidad y accesibilidad que mejoren la experiencia de los usuarios. La experiencia recabada en el área de diseño de interfaces y propiedades que propician la satisfacción del usuario, les ha permitido llevar adelante la revisión y selección de recomendaciones para el diseño, la evaluación y definición de mejoras, la implementación de rediseño, y recientemente, el asesoramiento y el desarrollo propiamente dicho de productos Web y móvil. Estas actividades se han estado ejecutando en diferentes dominios, tales como: (i) Universidad, (ii) Salud y Discapacidad, (iii) Gobierno y, (iv) Patrimonio Histórico-Cultural en el Turismo. Del trabajo realizado por GIFIS, podemos citar [1][2][3][4][5][6][7][8][9][10][11], entre otras.

1. INTRODUCCIÓN

Si bien no es Don Norman el creador del término eXperiencia de Usuario (UX), sin lugar a dudas es quien lo populariza al establecer, que hablar sólo de interfaz de usuario y de usabilidad, es demasiado limitado para cubrir todos los aspectos de la experiencia de las personas con un sistema. Posteriormente, Don Norman y Jakob Nielsen⁴, dos indiscutidos referentes en esta área del conocimiento, establecen que la UX engloba todos los aspectos de la interacción del usuario final con la entidad/organización/negocio, sus servicios y sus productos [12]. La UX no se trata solamente de diseñar-rediseñar buenas interfaces, sino que además, de aportar valor al usuario a través de esas interfaces. El trabajo en el campo de la UX ha demostrado, que se requiere entender el entorno que rodea al uso del producto y sacar el máximo provecho de las herramientas de las que se dispone, para aportarle valor al mismo. Como desarrolladores de productos Web y móvil, debemos ver el diseño-rediseño de la UX como un relato, que nos permite recorrer todos los caminos y situaciones por los que pasa el usuario. Sobre esta base, se trata de trabajar para

ofrecer interacciones lo más intuitivas posibles, capaces de generar experiencias satisfactorias. Y específicamente, el Diseño de Experiencia de Usuario (UXD⁵) fue creado para satisfacer este mandato, a los efectos de diseñar productos, servicios y soluciones basados en la UX.

Ahora bien, *¿cómo se evoluciona a la demanda UX por parte del consumidor de productos digitales?* En [13], se explica este fenómeno a partir de establecer una analogía con la famosa pirámide de Maslow [14]. La Figura 1, pone en jerarquía de niveles los siguientes aspectos primordiales que definen la UX: (1) lo “funcional”, es decir, que el producto ofrezca una solución de valor al usuario; (2) lo “confiable”, es decir, que se pueda confiar en la solución ofrecida por el producto y; (3) lo “usable”, es decir, que el producto ofrezca una interacción fácil e intuitiva al usuario.

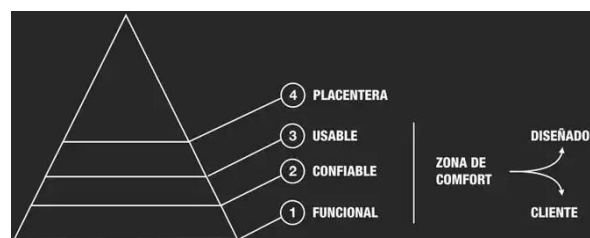


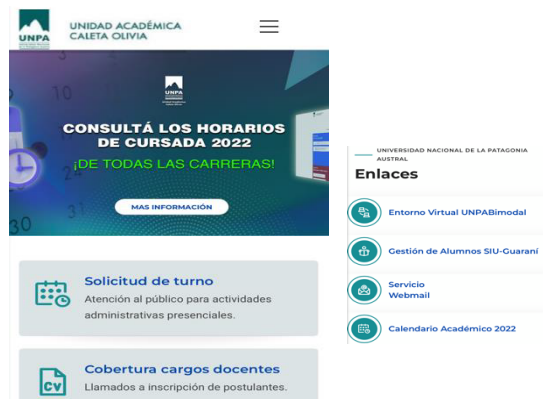
Figura 1: Maslow y los Usuarios [13]

Alcanzados estos tres primeros niveles, el próximo nivel se corresponde con el punto máximo de la satisfacción del usuario: una experiencia “placentera”. De acuerdo con la pirámide, inicialmente se buscaba la ergonomía, la cual surge oficialmente en el año 1949. En los años 70, el usuario se enfrentaba a un conjunto finito de pantallas y recién en los 80, se apoyan proyectos para la Interacción Humano-Computador (HCI), surge la “calidad en uso” con las normas ISO y se acuña el término “ingeniería de usabilidad”; la usabilidad propiamente dicha, florece en la década del 90 con la aparición de los sitios Web. Una vez implementados los paradigmas de ergonomía y usabilidad, en la actualidad, se trata de crear emociones positivas.

Y en este punto es oportuno preguntarse *¿cómo mejorar la UX del consumidor de productos digitales?* Ya mencionamos, que los factores de calidad como la usabilidad, la accesibilidad y la

⁴ Pioneers of User Experience - NN/g
<<https://www.nngroup.com/about/>>

⁵ UXD siglas en inglés de “User eXperience Design”



interacción del usuario son fundamentales, pero existen otros aspectos, y en este proceso se debe tener presente: (i) aportar “valor” al usuario, lo que implica entender el entorno que rodeará (o rodea actualmente) a un producto; (ii) identificar y aplicar las “ventajas de las herramientas” de las que disponemos para desarrollar un producto que brinde la solución requerida; (iii) armar un “relato” que recorra todos los procesos por los que pasa e interactúa el usuario, enumerando y listando todos los casos de esos procesos para convertirlos a lo más intuitivos posible.

A través del presente proyecto PI-UNPA, GIFIS continúa en el desarrollo, diseño-rediseño y evaluación de productos Web y móvil, enfocados en proveer soluciones para diferentes dominios de aplicación y priorizando la entrega de “valor” en una UX satisfactoria. Para ello, GIFIS cuenta con sólidos antecedentes, capitalizados a partir del trabajo realizado durante la ejecución de los PI-UNPA: 29/B167 (2014-2016), 29/B194 (2016-2018), 29/B222 (2018-2020) y 29/B256 (2020-2022).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo general de investigación y desarrollo que establece el presente PI-UNPA es: “Desarrollar Experiencias de Usuario para Productos Web y Móvil que Satisfagan las Expectativas y Necesidades de los Usuarios”. Y para alcanzar el objetivo general, se establecieron las siguientes metas: (i) Diseñar UX considerando la incorporación de tecnologías de geolocalización y realidad aumentada; (ii) Desarrollar aplicación móvil accesible para la comunicación alternativa y aumentativa destinada a personas con necesidades de comunicación complejas en el ámbito de la salud, (iii) Construir y evaluar sitios Web y móvil centrados en la usabilidad del contenido destinados a entidades de educación superior y de gobierno, y (iv) Explorar otras técnicas y herramientas para

mejorar la UX, tales como: storytelling, design thinking y gamificación, entre otras.

3. RESULTADOS OBTENIDOS/ ESPERADOS

A continuación se incluye el trabajo que se ha estado realizando y continúa en el PI actual.



Figura 2: Página de inicio UNPA-UACO Web

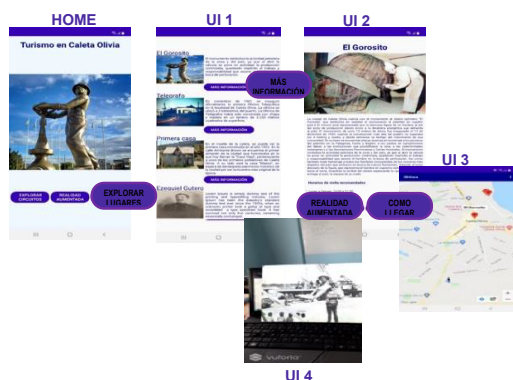
En el ámbito de la educación superior, se terminó el rediseño y se ha implementado el

Figura 3: Página de inicio UNPA-UACO Móvil

sitio Web y móvil de la UNPA-UACO⁶. A tal efecto: (i) se revisaron y reestructuraron todos los contenidos del sitio aplicando estrategias de contenido, y conceptos de ecosistemas de contenido y, (ii) se incluyeron durante el proceso propiedades de usabilidad y accesibilidad Web. Las Figuras 2 y 3, muestran fragmentos de la página de inicio UNPA-UACO Web y móvil respectivamente, para ilustrar como se ha manteniendo un diseño simple, homogéneo y claro para facilitar el acceso e interacción; en [2] se presenta este trabajo. Mientras que en el ámbito de gobierno, se

⁶ UNPA-UACO <<https://www.uaco.unpa.edu.ar/>>

comenzó a trabajar en el desarrollo del sitio Web para la Comisión de Fomento de la ciudad de Cañadón Seco⁷. Este proyecto se está realizando considerando dos aspectos: (i) el metodológico para asegurar que involucre a los ciudadanos, especialmente en la fase de creación-diseño y, (ii) el tecnológico para definir las herramientas a utilizar y asegurar capacitación en el uso óptimo de las mismas; en [3][4] se presenta este trabajo. También, se ha estado trabajando en desarrollar un soporte digital al patrimonio histórico-cultural de la ciudad de Caleta Olivia⁸, Patagonia Argentina. La ejecución de este proyecto involucra: (i) definir el enfoque metodológico para asegurar la participación de los interesados, (ii) analizar y recopilar guías UX específicas para aplicaciones de turismo histórico-cultural basadas en realidad aumentada y, (iii) analizar y capacitar en el aspecto tecnológico para aplicar las herramientas de geolocalización y realidad aumentada más convenientes y que contribuyan a la UX en este dominio; en [5], [10] y [11] se presentan estos trabajos, respectivamente. La Figura 4 ilustra el primer prototipo, que fue utilizado para llevar adelante una primera



interacción con los clientes.

Figura 4: Interfaces de Usuario y Funcionalidades

Finalmente, se ha estado trabajando sobre el concepto de lenguaje aumentativo y alternativo de comunicación, para desarrollar una aplicación móvil destinada a asistir al proceso de comunicación entre los profesionales de la salud y los pacientes con necesidades complejas

de comunicación. Este proyecto utiliza el conjunto de pictogramas de ARASAAC⁹, y su objetivo es proveer soporte al proceso de atención de pacientes con discapacidad comunicativa en las áreas de guardia e internación del Hospital Zonal Caleta Olivia; en [6][8][9] se presenta este trabajo. La Figura 5 ilustra el soporte que brindará la aplicación SAAC resultante.



Figura 5: Uso de la Aplicación - Profesional-Paciente

4. FORMACIÓN DE RECURSOS HUMANOS

En la actualidad, GIFIS tiene 9 integrantes: 6 docentes-investigadores UNPA, 1 docente-investigador invitado externo UTN-FRTDF y, 2 alumnos de grado UNPA. Durante 2022, los integrantes docentes-investigadores de GIFIS han estado presentes en la siguientes actividades: (i) evaluadores de becas de iniciación a la investigación para alumnos de grado y postgrado, (ii) revisores de artículos para revistas y eventos científicos-técnicos y revisores de proyectos y tesis, (iii) directores y codirectores de proyectos y tesis. Asimismo, los integrantes de GIFIS están en permanente actualización, para dar soporte a la ejecución de los proyectos vigentes y a la formación de sus integrantes. A tal efectos, se estuvieron tomando y dictando cursos y talleres de capacitación. La plataforma elegida para los desarrollos móvil, es Android Studio¹⁰, utilizando su lenguaje nativo Kotlin¹¹; lo que está demandando tiempo y esfuerzo de capacitación de recursos humanos.

AGRADECIMIENTOS

⁷ Municipio de Cañadón Seco <<https://www.municipalidad-argentina.com.ar/municipalidad-canadon-seco.html>>

⁸ Localidad de Caleta Olivia

<<https://www.santacruzpatagonia.gob.ar/centros-de-servicio/caleta-olivia>>

⁹ Centro Aragonés para la CAA <<https://arasaac.org/>>

¹⁰ Android developers <<https://developer.android.com/>>

¹¹ Kotlin for Android <<https://kotlinlang.org/docs/android-overview.html>>

A la UNPA por el soporte al PI N°: 29/B285, Período: 2022-2025, Denominado: “*Desarrollo y Evaluación de Proyectos Web y Móviles Centrado en la Experiencia de Usuario*”.

5. BIBLIOGRAFÍA

- [1] Martin, A., Gaetán, G., Saldaño, V., Cardozo, C., Villagra, S., Sosa, H., Colombres, M. Experiencia del Ciudadano: Mejorando la Interacción en Contextos Digitales. XXIV Workshop de Inv. en Cs. de la Computación (WICC); pp. 278-282, ISBN: 978-987-48222-3-9, <<https://wicc2022.uch.edu.ar/descargas/Libro-de-Actas-WICC-2022-1.pdf>>, 2022.
- [2] Gaetán, G., Villagra, S., Martin, A., Saldaño, V. Estrategia de Contenido para Mejorar la Experiencia de Usuarios de un Sitio Web Universitario. VII Encuentro de Inv., Becarios y Tesistas de la Patagonia Austral, Libro de Resúmenes del 7° EIPA, UNPA 2022.
- [3] Martin, A., Saldaño, V., Gaetán, G., Cardozo, C., Colombres, M., Sosa, H. Revalorizando el Patrimonio Histórico-Cultural de Caleta Olivia: Una Aplicación Móvil para Circuitos Turísticos y Educativos. VII Encuentro de Inv., Becarios y Tesistas de la Patagonia Austral, Libro de Resúmenes del 7° EIPA, UNPA 2022.
- [4] Cicchinelli Fernandez, J. A., Gaetán, G., Villagra, S., Martin, A., Saldaño, V. Rediseño Centrado en el Usuario para un Sitio Web de Gobierno: Un Caso de Estudio. VII Encuentro de Inv., Becarios y Tesistas de la Patagonia Austral, Libro de Resúmenes del 7° EIPA, UNPA 2022.
- [5] Cardozo, C., Martin, A., Saldaño, V., Gaetán, G. Aplicaciones de Turismo Cultural: Definiendo un Enfoque para el Desarrollo. VII Encuentro de Inv., Becarios y Tesistas de la Patagonia Austral, Libro de Resúmenes del 7° EIPA, UNPA 2022.
- [6] Sosa, H., Martin, A., Saldaño, V. Gaetán, G. Atendiendo Necesidades de Comunicación Complejas: Una Aplicación Saac Móvil para un Hospital Zonal. VII Encuentro de Inv., Becarios y Tesistas de la Patagonia Austral, Libro de Resúmenes del 7° EIPA, UNPA 2022.
- [7] Martin, A., Gaetán, G., Saldaño, V., Cardozo, C., Villagra, S., Sosa, H. Mejorando la Experiencia de Usuario en la Interacción con Productos Web y Móviles. VII Encuentro de Inv., Becarios y Tesistas de la Patagonia Austral, Libro de Resúmenes del 7° EIPA, UNPA 2022.
- [8] Sosa, H., Martin, A., Saldaño, V., Gaetán, G. Desarrollando una Aplicación SAAC Móvil para el Hospital Zonal de Caleta Olivia. Libro de Resúmenes del 1er. Ateneo de Estudiantes del ITA, ISBN 978-987-3714-96-2, <<https://www.uarg.unpa.edu.ar/ita/attachments/article/282/Libro%20de%20resumenes%20ATENEO%202022.pdf>>, 2022.
- [9] Sosa, H., Martin, A., Saldaño, V., Necesidades de Comunicación Complejas: Desarrollando una Aplicación SAAC Móvil para el Hospital Zonal de Caleta Olivia. Libro de Actas del XXVIII Congreso Argentino de Cs. de la Computación (CACIC 2022), pp. 920-924, ISBN: 978-987-1364-31-2, <<file:///Users/adrianamartin/Downloads/LIBRO%20DE%20ACTAS-CACIC2022Opt2b.pdf>>, 2022.
- [10] Vega, M., Gaetán, G., Martin, A., Guías de Experiencia de Usuario para Aplicaciones de Turismo Cultural basadas en Realidad Aumentada. ICT-UNPA. Vol. 13(2), pp. 26-43 ISSN:1852-4516, DOI: <<https://doi.org/10.22305/ict-unpa.v13.n2.812>>, 2021.
- [11] Montaña Torrico, C., Saldaño, V., Martin, A. Integrando Herramientas en una Aplicación Turística para el Aprendizaje de Realidad Aumentada y Geolocalización. ICT-UNPA, Vol. 14(3), pp. 142-169, ISSN: 1852-4516, DOI: <<https://doi.org/10.22305/ict-unpa.v14.n3.900>>, 2023.
- [12] Norman, D., Nielsen, J. The Definition of User Experience <<https://www.nngroup.com/articles/definition-user-experience/>> Accedido 2023.
- [13] Delgado, Borja. Diseñar en la Cuerda Floja. UXFighters A User Experience and Innovation Event. Madrid, España, 2014 <https://www.interaction-design.org/events/ux-meetups/ux_fighters_a_user_experience_and_innovation_event_sep_13th_2014>, Accedido 2022.

- [14] Maslow, A. A Theory of Human Motivation. Editorial Martino Fine Books. Reprint of 1943 Edition. ISBN-10: 1614274371, 2013.

Evaluación de la Sostenibilidad en un Modelo de Calidad del Software

Leo Rosana*, Salgado Carlos⁺, Sanchez Alberto⁺, Peralta Mario⁺

* Departamento Académico de Ciencias Exactas, Físicas y Naturales. Universidad Nacional de la Rioja. La Rioja, Argentina

leorosana@gmail.com

⁺ Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina

e-mail: {csalgado, mperalta, asanchez}@unsl.edu.ar

RESUMEN

El software, herramienta imprescindible en las más variadas gestiones de la actualidad, también puede ser *sostenible*, si se tiene en cuenta que un producto sostenible es aquel que aporta beneficios ambientales, sociales y económicos, resguardando la salud pública, el bienestar y el medio ambiente en todo su ciclo de vida.

En la línea de investigación se ha definido un modelo de calidad del software, basado en la Norma ISO 25000, que incluye a la sostenibilidad como característica transversal. Además, se ha definido un conjunto de métricas e indicadores asociados al modelo de calidad y en particular a la sostenibilidad. Como continuación de la investigación se está trabajando en la implementación de un algoritmo que permita determinar el grado o nivel de sostenibilidad de un software y su impacto para la sociedad.

Palabras Clave: *Método de evaluación o cálculo, Calidad de Software, Modelos de Calidad, Métricas, Sostenibilidad.*

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-03-2020. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas

universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

La literatura en general coincide que la calidad del software es el cumplimiento y/o grado de satisfacción de requisitos tanto explícitos como implícitos. Pressman la define como: “*Proceso eficaz de software que se aplica de manera que crea un producto útil que proporciona valor medible a quienes lo producen y a quienes lo utilizan*” [1].

La especificación y evaluación de la calidad son importantes para garantizar el valor a todas las partes interesadas del producto software. Para ello se definen características de calidad asociadas con metas y objetivos del sistema, las cuales deben especificarse, medirse y evaluarse mediante métodos de medición validados o ampliamente aceptados [2].

Es importante definir y/o implementar algún modelo, método o herramienta que permita gestionar atributos en el proceso de construcción del software, puesto que la concordancia de los requisitos y su construcción son la base de las medidas de calidad establecidas [3], tal y como lo describe en el modelo de calidad para el producto software la Norma Internacional ISO/IEC 25010:2011 [2], una de las divisiones de la serie SQUARE.

En este sentido la calidad se convierte en algo concreto, que se puede definir, medir y, sobre todo, planificar.

A los fines de este trabajo, se toma como base el modelo definido en [4], que incluye a la sostenibilidad transversalmente en el modelo

que propone SQuaRE [2]. No será una característica aislada, estará inmersa en el modelo, atravesando toda la estructura y complementándose con las restantes características mediante criterios sostenibles en la medición y evaluación de atributos.

El termino sostenibilidad, surge en 1987 en el Informe de Brundtland [5], titulado “Nuestro futuro común” y por la necesidad de estudiar y delimitar el impacto de las actividades humanas sobre el medio ambiente.

En 2015 un manifiesto expone principios y compromisos vinculados con el diseño sostenible (Becker) [6], infiriendo que la sostenibilidad tiene múltiples dimensiones (social, medioambiental, económica, técnica y humana) y que todas deben analizarse.

En 2018 se presenta el “Catálogo de Diseño de Sostenibilidad del Software” [7] como la herramienta que permite integrar la sostenibilidad en el diseño. Es un conjunto de criterios derivados de los nueve principios del manifiesto de Karlskrona, basados en el análisis cruzado de diferentes sistemas. Para cada criterio se derivan indicadores relacionados con las dimensiones de la sostenibilidad y su orden de impactos.

Valorar la sostenibilidad mediante la puntuación o medición en un rango acotado, permitirá tener una idea cuantitativa de la misma. Estableciendo un umbral se podrá determinar en qué grado, el objeto de estudio, es sostenible.

Del trabajo de Condori Fernández y Lago [8], surge un importante aporte al determinar qué atributos de calidad del modelo ISO/IEC 25010 son más relevantes para las distintas dimensiones de la sostenibilidad. Los autores asignan niveles de contribución del atributo a la dimensión, que pueden ser: altamente contributiva, contributiva, ligeramente contributiva o no contributiva. Se comparan los resultados obtenidos en dos instancias diferentes obteniendo tres valores posibles (0, 1 o 2) que indican qué característica es más importante para la dimensión en estudio y cuál es el atributo más relevante.

A los fines de este trabajo, se consideran para inferir la sostenibilidad, los atributos

detallados en Tabla 1, relacionados con las dimensiones ambiental, técnica, económica y social. Para validar la propuesta se toma como caso de estudio al Sistema de Contrataciones de la Provincia de La Rioja [9], particularmente al módulo de Proveedores y al de Publicación de las Contrataciones.

Tabla 1. Características y Atributos de Calidad del Software relacionados con las dimensiones de la sostenibilidad

CARACTERÍSTICA	ATRIBUTO	DIMENSIÓN
Compatibilidad	Interoperabilidad	Técnica
Adecuación Funcional	Corrección funcional	
Mantenibilidad	Capacidad de ser modificado	
Portabilidad	Adaptabilidad	
Eficiencia de Desempeño	Capacidad	Ambiental
	Utilización de recursos	
Seguridad	Confidencialidad	Social
	Integridad	
Usabilidad	Capacidad para ser usado	
	Protección frente a errores de usuarios	
	Accesibilidad	
Fiabilidad	Disponibilidad	Económica

El Sistema Provincial de Contrataciones fue creado por Ley Provincial N° 9341 del 13 de diciembre de 2012 [10]. Tiene por objeto que las compras públicas se realicen con la mejor tecnología, en el momento oportuno y al menor costo posible, coadyuvando al desempeño eficiente de la Administración y al logro de los resultados requeridos por la sociedad, y tiene por finalidad asegurar una eficiente utilización de los Fondos Públicos, mediante la exigencia de la definición previa de las necesidades a satisfacer, la salvaguarda de la libre competencia y la selección de la oferta más ventajosa. Mediante el modelo se aplican métricas en dos momentos, cuando la gestión era manual y en algunos casos semiautomática y actualmente, cuando se digitalizó tanto la gestión de los proveedores del Estado como la difusión de los pliegos a los posibles oferentes conforme al rubro económico en el cual se encuentran registrados. Por ej., en el año 2019

se aprobaron 255 proveedores mediante una gestión combinada, se preinscribían en el sistema [9] y luego debían acercarse a la Dirección General de Sistemas de Contrataciones para terminar el trámite presentando toda la documentación en papel, legajo físico. En 2021, luego de implementar la reingeniería del proceso de registro, se aprobaron 491 proveedores, generándose el legajo electrónico con la documentación digital ingresada en dicho proceso, lográndose digitalizar y despapelizar la gestión.

A continuación, se exponen ejemplos de las métricas mencionadas.

Característica: *Eficiencia de Desempeño*

Subcaracterística: *Utilización de Recursos*

Métrica: *Uso del Papel*

Función de Medición: $X=A/B$

A=Cantidad de informes en papel

B=Cantidad total de informes ($B>0$)

Resultados:

2019: $X=9/10$ $X=0,9$

2021: $X=0/10$ $X=0$

Valor deseado: Lo más cercano a 0 es mejor

Teniendo en cuenta que la dimensión ambiental busca evitar que el desarrollo y uso de sistemas dañe el medio ambiente en el que operan, se define a la Utilización de Recursos como un atributo muy importante. En la valoración se tuvo en cuenta principalmente a la métrica referida al “uso del papel”, debido a que fue un punto muy evidente de mejora y evolución en los dos momentos del sistema que se midieron. No obstante, se busca hacer frente a los retos actuales, reduciendo la demanda energética, el uso de servidores, almacenamiento y transferencia de datos.

Característica: *Usabilidad*

Subcaracterística: *Accesibilidad*

Métrica: *Facilidad de acceso al sistema*

Función de Medición: $X=A/B$

A=Otras capacidades detectadas en los accesos

B=Número de otras capacidades especificadas

Resultados:

2019: $X=1/3$ $X=0,33$

2021: $X=2/3$ $X=0,66$

Valor deseado: Lo más cercano a 1 es mejor

Nos pareció importante también medir la subcaracterística Accesibilidad por lo que en sí misma implica, teniendo en cuenta que la Usabilidad contribuye a la dimensión social. Para la valoración de la métrica se tomó en cuenta la facilidad de acceso de personas con otras capacidades al sistema. Especificamos como otras capacidades la ceguera, la sordera y los problemas motrices, de allí que en el 2019 solo personas sordomudas podían acceder a gestiones que mediana o totalmente requerían tramitación presencial. En cambio, en 2021 al cambiar a una gestión digital también podían acceder personas con problemas motrices que no necesitaban trasladarse; quedando como desafío adaptar el sistema, incorporando funcionalidades o dispositivos necesarios para el acceso de personas con ceguera o problemas visuales.

Luego de obtener los resultados de las métricas, es importante definir los “umbrales” (Tabla 2), puntos de referencia que permitirán determinar el nivel de cumplimiento o las desviaciones, posibles problemas de calidad del producto.

Tabla 2. Valoración de resultados de las métricas

PUNTUACIÓN (valor óptimo: el más cercano a 1)	VALORACIÓN	PUNTUACIÓN (valor óptimo: el más cercano a 0)
1	Muy bueno	0
0,8	Bueno	0,2
0,5	Regular	0,5
0,2	Malo	0,8
0	Muy malo	1

Teniendo en cuenta los ejemplos citados precedentemente, el resultado de la métrica “uso del papel”, para resultar óptimo debe acercarse o ser igual a 0 y se observa en las dos instancias valoradas la evolución favorable de *malo* a *muy bueno*.

En cambio, la “facilidad de acceso al sistema” será mejor cuanto más cerca esté el resultado

de 1. Y si bien evolucionó de *regular* a *bueno*, quedan aspectos importantes por mejorar.

Con los umbrales traducimos los resultados numéricos en una calificación cualitativa que nos permitirá determinar si el producto, objeto de estudio, es o no sostenible. Consideramos que solamente si la valoración es “*muy bueno*”, se puede inferir en que el software es Sostenible.

En base a estas propuestas, y siguiendo esta línea de investigación, surge la necesidad de contar con un método de evaluación que permita la instanciación de un modelo de Calidad con base en la sostenibilidad. Es decir, poder evaluar el grado en que una empresa u organización cumple con un modelo de calidad. El método propuesto tiene sus bases en modelos matemáticos y estadísticos que permiten que los procesos sean sistemáticos y repetibles. Para definirlo, se tuvieron en cuenta diversas herramientas, técnicas, modelos y métodos. Se parte eligiendo un modelo de calidad, en particular se tomó como punto de partida el modelo definido en [4]. Cabe aclarar que dicho modelo, se plantea como inicio para el método, pero no es de ninguna manera estático, es decir, se define de manera que sea aplicable a otros modelos de calidad. Se puede ajustar a las distintas situaciones, tecnologías o reglas de negocio/mercado a medida que vayan surgiendo.

El objetivo del método es poder validar/verificar si el modelo de calidad de la empresa se ajusta, y en qué medida, al modelo de gestión de la calidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de trabajo de esta línea de investigación están asociados a:

- Evaluación de la calidad de productos de software.
- Estudio de modelos conceptuales con base en la sostenibilidad aplicados a la calidad de productos de software.
- Estudio de estándares y metodologías aplicadas a la construcción de MC.

3. RESULTADOS OBTENIDOS/ESPERADOS

La línea de I+D presentada en este trabajo se está desarrollando teniendo en cuenta los objetivos del Proyecto de I+D que la contiene. De manera específica para esta línea investigación se obtuvieron hasta la fecha los siguientes resultados:

- ☒ Estudio de los modelos y normas de calidad aplicados a productos de software.
- ☒ Estudio de metodologías o métodos que guíen la construcción de modelos de calidad.
- ☒ Definición del modelo de calidad con base a la sostenibilidad.
- ☒ Validación del modelo de calidad a través de la evaluación de distintos softwares para saber si cubren las demandas organizacionales y las necesidades de los usuarios, y su impacto en el contexto que está inserto.

Los resultados esperados son:

- ☒ Concluir con la definición del modelo conceptual de calidad aplicado a productos software a través de la metodología y norma de calidad seleccionada, sin perder de vista los criterios de sostenibilidad.
- ☒ Definir e implementar el método o algoritmo para llevar a cabo las distintas evaluaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se trabaja en lo referente a distintos modelos y métodos de evaluación de calidad. Se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) de la Lic. Rosana Leo, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Computación y trabajos finales de carrera

de la Ingeniería en Informática e Ingeniería en Computación de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

5. REFERENCIAS

- [1]. Pressman R.; “Ingeniería de Software. Un enfoque práctico” 9ª Ed. McGraw-Hill Interamericana, 2021
- [2]. ISO/IEC 25010 Systems and software engineering-Systems and Software Quality Requirements and Evaluation (SQUARE) System and software quality models.
- [3]. Callejas-Cuervo, M.; Alarcón-Aldana, A.; Álvarez-Carreño, A. Modelos de calidad del software, un estado del arte. En: Entramado. Enero - junio, 2017. vol. 13, no. 1, p. 236-250
- [4]. Leo R., Salgado C., Peralta M., Sánchez a. “Un Modelo de Calidad de Software con La Sostenibilidad como Característica Transversal” IDETEC 2022.
- [5]. CMMAD. Comisión Mundial de Medio Ambiente y Desarrollo de la Organización de las Naciones Unidas (ONU) (1987). Informe Brundtland: “Nuestro futuro común”
- [6]. Becker C. Manifiesto Karlskrona. Sustainability design and software. 2015
- [7]. Oyedeji, Shola – Seffah Ahmed – Pensestadler Birgit. Catálogo de Diseño de Sostenibilidad del Software (Suiza). 2018.
- [8]. Fernández, N.C. y Lago P. “Characterizing the contribution of quality requirements to software sustainability”. Revista de Sistemas y Software, Vol. 137 pag 289-305. Marzo 2018.
- [9]. Sistema de Contrataciones de la Provincia de La Rioja. 2015.
<https://compras.larioja.gob.ar/index.php>
- [10]. Ley Provincial N° 9341. Régimen General de Contrataciones. 13 de diciembre de 2012 (B.O. 23 de abril de 2013).

CoVaMaT: Modelo conceptual de una herramienta para el soporte a la gestión de variedad en Sistemas Big Data

Agustina Buccella, Alejandra Cechich, LÍam Osycka,
Juan Luzuriaga, Carolina Villegas, Marcos Cruz, Franco Corgatelli,
Rodolfo Martínez, Rafaela Mazalu
GIISCO Research Group
Departamento de Ingeniería de Sistemas
Universidad Nacional del Comahue
Neuquen, Argentina
agustina.buccella@fi.uncoma.edu.ar

1. Resumen

Un cambio importante con respecto a depósitos de datos tradicionales, es que en los Sistemas Big Data (SBDs) la naturaleza no estructurada de algunos datos puede provenir de diversas fuentes, entre ellas sensores, redes sociales, entorno y la misma empresa. La diversidad de esos datos puede analizarse abordando distintas características. Precisamente, la propiedad de los SBDs con respecto a diversidad de los datos se denomina *variedad*. Nuestro proyecto propone modelar variedad mediante casos documentados a través de las variaciones que diferentes variables pueden tomar en un contexto. Sin embargo, para la aplicación de la propuesta, es indispensable una herramienta de soporte que construya incrementalmente repositorios de variaciones a ser reusadas.

Palabras Clave : Reusabilidad - Líneas de Producto de Software - Big Data

2. Contexto

La línea presentada se inserta en el contexto del Proyecto UNComa: Modelado de Variedad en Sistemas Big Data Directora: Dra. Agustina Buccella, Co-directora; Dra. Alejandra Cechich (2022-2025).

3. Introducción

En [3] hemos definido una primera aproximación de los elementos que componen una arquitectura de referencia para SBDs basada en reuso. Uno de los componentes principales de esta arquitectura agrupa los denominados activos de dominio constituidos por artefactos de software que son creados para el dominio en el que se está trabajando. Es importante resaltar que estos activos deben generarse a partir de activos basados en conocimiento. De esta forma, se deben crear artefactos enfocados en que puedan ser reusados en el mismo dominio e incluso en otros dominios relacionados (artefactos para reuso), y/o que puedan desarrollarse en base a otros artefactos ya creados (artefactos con reuso).

Para identificar variedad en activos de dominio, previamente propusimos dos enfoques: (1) identificación desde los datos (bottom-up) y (2) identificación desde los requerimientos (top-down). El enfoque bottom-up de la propuesta parte de la definición de un problema dependiente del dominio e intenta detectar características variantes a partir de los datos existentes en uno o varios repositorios. Por otro lado, en el enfoque top-down se parte del análisis de los requerimientos y del dominio para contrastar suposiciones a través de los datos. Detalles de esta propuesta pueden verse

en [4, 6, 7]. La documentación de activos de dominio se describe mediante un enfoque basado en funcionalidades, donde cada funcionalidad se documenta a través de una hoja de datos funcional (datasheet), representando el conjunto de variaciones que permitan reusar activos de dominio y detectar variedades de contexto.

3.1. CoVaMaT: Context-based Variety Management Tool

La propuesta planteada cobra sentido cuando se tiene una forma de poder almacenar esas variedades formalizadas en forma de datasheets y cuando, además, las variaciones pueden asociarse a casos puntuales. Por ello, se debe plantear una solución que permita almacenar, consultar, modificar y visualizar la variedad de un dominio y todos los casos trabajados, siendo éstos recuperados y reutilizados cuando la variedad de contexto lo permite. La solución se basa entonces en el diseño de una herramienta software, nombrada Context-based Variety Management Tool (CoVaMaT), que tiene la intención de exponer el conocimiento adquirido dentro de un dominio de una forma declarativa y que pueda ser de soporte tanto para el proceso bottom-up como top-down en la identificación de variedad.

Previo a la presentación del modelo conceptual de la herramienta, es necesario introducir las siguientes definiciones:

- **Instanciación de Caso**: consiste en la descripción de un caso de dominio a través de la selección de variaciones que lo describen. Por ejemplo, en el dominio de recursos hídricos, los cuerpos de agua pueden ser ríos, lagos, etc. El caso en cuestión puede corresponder a uno de estos valores (ej. ríos). Además, para ese río en particular, pueden resultar relevantes valores específicos de calidad de agua, como el oxígeno disuelto, pH, etc., y de los cuales se hará también una selección. El conjunto de valores seleccionados para todas las variables que sean relevantes en el dominio, constituirá el caso instanciado.
- **Caso de Reuso**: se utiliza para componer una consulta que recupere aquellos casos

ya instanciados en el repositorio que posean características similares. Esta consulta se compone de un contexto determinado por variaciones seleccionadas (ej. río con variables de calidad de agua con foco en la turbidez). El objetivo es analizar trabajo anterior realizado bajo condiciones especificadas en el caso de reuso.

La arquitectura lógica de CoVaMaT presenta un producto que empaqueta tres servicios, con procesos que los atienden y eventos que los invocan. A continuación se explicarán estos procesos, centrando en el nivel conceptual de negocios de la herramienta, de acuerdo a la notación propuesta por el lenguaje de modelado de procesos ArchiMate.

Documentación de Variedad: S-1, P-1 y E-1

- **S-1 Servicio de documentación de variedad**: le permite al usuario documentar los diferentes tipos de variedad que se encuentran en un dominio cuando se trabaja con el ciclo de actividades de Big Data (adquisición, transformación, análisis y presentación de datos).
- **P-1 Documentar variedad en dominio**: es el proceso que especifica cómo debe trabajarse al momento de identificar un tipo de variedad, y de construir los objetos de negocios que documenten esta variedad.
- **E-1 Comienzo de búsqueda de variedad**: es el evento que se dispara cuando se comienza a una nueva iteración del modelo dentro de un dominio.

Luego de finalizar todas las ejecuciones de (P-1) se tiene un conjunto de datasheets almacenadas representando la variedad en el dominio. Es así, que a partir de estas, puede representarse un caso de dominio seleccionando las variaciones de las diferentes datasheets.

Instanciación de caso: S-2, P-2 y E-2

- **S-2 Servicio de instanciación de caso**: le permite al usuario documentar un caso

¹<https://www.opengroup.org/archimate-forum/archimate-overview>

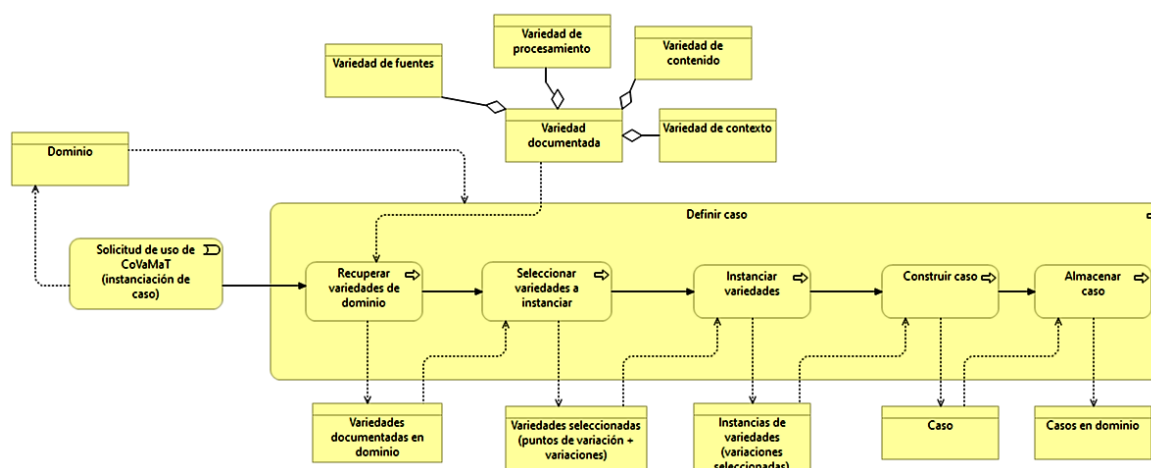


Figura 1. P-2: Definir Caso

de dominio compuesto por instancias de las diferentes variedades encontradas. El caso será identificado particularmente por su variedad de contexto.

- P-2 Definir caso : es el proceso que especifica cómo debe confeccionarse el caso de dominio que será almacenado.
- E-2 Solicitud de uso de CoVaMaT (instanciación de caso): es el evento que se dispara cuando se quiere almacenar un caso en el dominio para su futuro reuso y consulta.

En la Figura 1 puede verse el proceso modelado. Primero entonces, se recuperan todas las variedades almacenadas del dominio, luego se deben seleccionar cuáles se utilizarán para definir el caso y construir al objeto de negocio Variedades seleccionadas (puntos de variación + variaciones). Con este objeto, puede procederse a crear eCaso. Finalmente, el objeto es almacenado dentro de todos los casos del dominio.

Consulta de activos: S-3, P-3, P-4 y E-3

- S-3 Servicio de consulta de activos de dominio: con este servicio se pueden consultar activos de dominio (que serían los casos y las variedades asociadas a ellos) almacenados en la base de conocimiento de CoVaMaT. La consulta será en base a un contexto definido por el usuario que solicita usar el servicio, recuperando entonces los casos que responden al mis-

mo.

- P-3 Definir caso de reuso : en este proceso se construye un caso a partir de la variedad de contexto almacenada en el dominio.
- P-4 Consultar activos de dominio en casos similares : en este proceso se toma el caso de reuso definido en (P-3) para poder recuperar otros con contexto similar y poder observar qué variedades instanciaron.
- E-3 Solicitud de reuso : es el evento que se dispara cuando se quiere recuperar conocimiento del dominio en función de casos trabajados sobre un contexto particular.

En la Figura 2 se observan los pasos a seguir para recuperar el conocimiento de experiencias previas. Primero, se deben recuperar todas las diferentes datasheets que conforman la variedad de contexto del dominio, almacenadas en el objeto Variedad de contexto creado en (P-1). Luego, similar a (P-2), se deben seleccionar qué datasheets se usarán para definir el contexto del caso de reuso, siendo almacenadas en el objeto de negocio Variedades de contexto seleccionadas (punto de variación + variaciones).

Una vez definidas qué datasheets de contexto se usarán, se procede a instanciarlas; y se almacena en el objeto de negocio Instancias de variedades de contexto (variedades seleccionadas). A partir de las datasheets elegidas

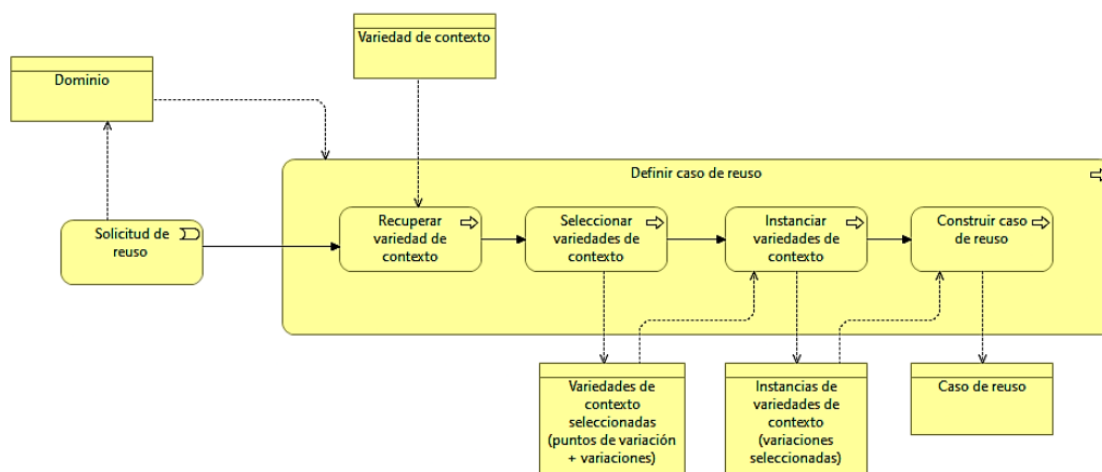


Figura 2. P-3: Definir caso de reuso

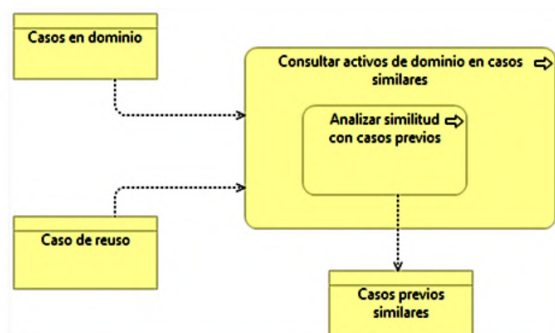


Figura 3. P-4: Consultar activos de dominio en casos similares

y las instanciaciones de cada una, se construye el objeto Caso de reuso. Con esto finaliza la ejecución de (P-3) y comienza la ejecución de (P-4).

En el proceso siguiente, Figura 3, se utilizan los objetos de negocio Casos en dominio (generado en P-2) y Caso de reuso en (P-3) como entradas para realizar una consulta de similitud, con el fin de filtrar dentro de todos los casos de dominio a aquellos que posean la misma variedad de contexto por la que se está consultando. De esta forma puede verse qué se ha hecho previamente en el dominio cuando se presentaron situaciones con un contexto particular.

4. Líneas de Investigación y Desarrollo

En proyectos previos, hemos realizado amplios avances en lo que respecta al área de LPSs definiendo y refinando una metodología de desarrollo a nivel de subdominios [1] que tiene la particularidad de favorecer el reuso basado en una taxonomía de servicios [2, 5].

Actualmente, nuestra línea de investigación migra hacia un cambio con respecto a la construcción de sistemas Big Data con reuso. La variedad en SBDs ha sido relacionada con diversas propiedades como interoperabilidad, seguridad, reusabilidad, etc. En este contexto, y respondiendo a la pregunta de investigación: ¿Cómo puede modelarse la variedad de la información de dominio de manera de incorporar reusabilidad en el desarrollo de SBDs? , nuestra línea de trabajo ha avanzado en los resultados que se listan a continuación.

5. Resultados Obtenidos / Esperados

Al momento, hemos planteado una arquitectura de referencia [3] y un modelo de procesos [6, 7, 4] para la gestión de la variedad en SBDs. Como primeras aproximaciones experimentales, en [8, 9], presentamos caracterizaciones de contexto de variables que influyen la turbidez en los cuerpos de agua y el índice de vegetación normalizado (NDVI). Esto per-

mite combinar datos relevantes del contexto de aplicación, lo que favorecería la identificación de situaciones recurrentes. La propuesta se ejemplifica con un caso de estudio llevado a cabo en el área del Alto Valle del Río Negro.

Continuamos trabajando en la implementación de CoVaMaT y también en colaboración con el Instituto de Tecnología Agropecuaria (INTA)-Alto Valle para la aplicación del proceso de modelado en el análisis de la napa freática, en función de la variedad de fuentes acuíferas de diversas zonas geográficas (variedad contextual)².

6. Formación de Recursos Humanos

El proyecto reúne aproximadamente a 13 investigadores, entre los que se cuentan docentes y alumnos de UNComa, y colaboradoras expertas del dominio de aplicación, específicamente pertenecientes al INTA. A su vez, el proyecto cuenta actualmente con tres docentes y un magister. Varios de los docentes-investigadores de GIISCo-UNComa han terminado o se encuentran próximos a terminar carreras de grado o postgrado.

Referencias

- [1] A. Buccella, A. Cechich, M. Arias, M. Pol'la, S. Doldan, and E. Morsan. Towards systematic software reuse of gis: Insights from a case study. *Computers & Geosciences* 54(0):9 – 20, 2013.
- [2] A. Buccella, A. Cechich, M. Pol'la, M. Arias, S. Doldan, and E. Morsan. Marine ecology service reuse through taxonomy-oriented SPL development. *Computers & Geosciences* 73(0):108 – 121, 2014.
- [3] A. Buccella, J. Luzuriaga, A. Cechich, L. Osycka, F. Paterno, M. Pol'la, M. Cruz, R. Martinez, R. Mazalu, and M. Moyano. Reusabilidad en el contexto de desarrollo de sistemas para big data. In *Actas del XXIII Workshop de Investigadores en Ciencias de la Computación*, Chilecito, La Rioja, pages 525–529, 2021.
- [4] A. Buccella, J. Luzuriaga, A. Cechich, L. Osycka, C. Villegas, M. Cruz, F. Corgatelli, R. Martinez, R. Mazalu, and M. Moyano. Modelado de variedad de activos de dominio en sistemas big data. In *Actas del XXIV Workshop de Investigadores en Ciencias de la Computación*, Mendoza pages 113–117, 2022.
- [5] Agustina Buccella, Alejandra Cechich, Juan Porfiri, and Domenica Diniz Dos Santos. Taxonomy-oriented domain analysis of gis: A case study for paleontological software systems. *ISPRS International Journal of Geo-Information* 8(6), 2019.
- [6] L. Osycka, A. Buccella, and N. A. Cechich. Identificación de variedad contextual en modelado de sistemas big data. In *Memorias del XXVII Congreso Argentino de Ciencias de la Computación (CACIC)*, pages 367–376. Red de Universidades con Carreras en Informática, 2021.
- [7] L. Osycka, A. Buccella, and N. A. Cechich. Data variety modeling: A case of contextual diversity identification from a bottom-up perspective. In *27th Argentine Congress, CACIC 2021*, Salta, Argentina, October 4-8, 2021, Revised Selected Papers. *Communications in Computer and Information Science* 1584, pages 124–138. Springer, 2022.
- [8] Gastón Vidart, Alejandra Cechich, Agustina Buccella, and Ayelén Montenegro. Análisis de turbidez basado en caracterización de contextos. *Memorias de las JAIIO*, 8(4):170–183, dic. 2022.
- [9] C. Villegas, A. Buccella, and A. Cechich. Caracterización de variables para el análisis del Índice de vegetación. In *Memorias del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC) - Workshop Alumnos*, pages 978–982, 2022.

²Convenio de Cooperación Técnica FaIF-INTA Resolución CD FaIF N°109/21

Proceso de validación de requerimientos aplicando técnicas de procesamiento de lenguaje natural en un entorno colaborativo

Sonia Santana , Lucrecia Perero , Alejandro Fernandez , Leandro Antonelli 

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos
sonia.santana, ramona.perero [@uner.edu.ar],
Laboratorio de Investigación y Formación en Informática Avanzada (LIFIA),
Facultad de Informática, Universidad Nacional de La Plata
alejandro.fernandez, leandro.antonelli [@lifia.info.unlp.edu.ar]

Resumen

En el marco de la Ingeniería de Requerimientos (RE por sus siglas en inglés Requirements Engineering) la validación de los requerimientos es una tarea fundamental. Esto es así en cualquier proyecto de Ingeniería de Software y debe ser un proceso continuo en el ciclo de vida del desarrollo del sistema. El principal objetivo de la validación de los requerimientos es confirmar que los mismos sean representaciones de las necesidades y expectativas de los usuarios [1] [2] [3] y que sean completos, correctos y consistentes [3] entre otras características. Trabajar en la validación de requerimientos se está convirtiendo en un desafío para los equipos, clientes y usuarios. Existen diferentes causas que imponen problemas de comunicación, control, intercambio de conocimientos, confianza y retrasos en el desarrollo del software [4].

Hoy en día, los requerimientos se escriben con la participación de un gran número de personas que producen una cantidad variada de artefactos. En este contexto los entornos colaborativos y el procesamiento de lenguaje natural cobran importancia. Este artículo discute los aspectos más importantes de una línea de investigación iniciada en la Universidad Nacional de Entre Ríos (UNER) que tiene como objetivo desarrollar un proceso de validación de requerimientos utilizando, entre otras técnicas,

procesamiento de lenguaje natural y entornos colaborativos.

Palabras clave: Ingeniería de Requerimientos, validación de requerimientos, técnicas de procesamiento de lenguaje natural, entorno colaborativo.

Contexto

El presente PID 7070 se encuadra en la línea de investigación "Ingeniería de Software" y es un proyecto en conjunto entre Facultad de Ciencias de la Administración de la UNER y la Facultad de Informática de la Universidad Nacional de la Plata (UNLP). La línea de investigación es establecida como prioritaria desde la carrera Licenciatura en Sistemas de la Facultad de Ciencias de la Administración de la UNER. Se adecua, además, a una de las prioridades de la UNER considerando que es un proyecto aplicado a la investigación sobre Tecnologías de la Información y la Comunicación.

Introducción

En la industria del software el objetivo fundamental que se debería alcanzar es obtener productos software de calidad aplicando diferentes métodos y modelos bajo procesos estandarizados en el desarrollo de software, combinando técnicas, estrategias y todos aquellos aspectos que contribuyan a la búsqueda

de conseguir calidad en los productos intangibles.

En la actualidad, se observa que los sistemas de software son cada vez más complejos, por lo que las exigencias de calidad en el desarrollo deben ser mayores. Es necesario que el desarrollo de software sea más riguroso para obtener un producto de adecuada calidad. La identificación y tratamiento de errores en etapas tempranas del proyecto de software es crucial para disminuir los costos de operación y evitar errores en etapas posteriores del desarrollo.

Definir y validar los requerimientos del sistema son las actividades más importantes en el desarrollo del software. A menudo requieren la colaboración de múltiples partes interesadas que tienen diferentes necesidades y perspectivas. En muchos casos, es especialmente difícil facilitar la recopilación de requerimientos de manera eficiente y eficaz en un entorno jerárquico y, al mismo tiempo, animar a las partes interesadas del sistema a compartir libremente sus ideas y opiniones [5].

En las últimas dos décadas, las actividades de diseño de productos se han trasladado a un entorno de colaboración entre empresas extendidas, debido a la globalización de los mercados, el avance de las tecnologías, la segregación de las demandas de los clientes y la competencia en el país y en el extranjero [6]. El trabajo colaborativo tiene como resultado una experiencia enriquecedora con requerimientos más completos y consistentes [7]. Además, la interacción de diferentes partes interesadas sobre el mismo problema mejora la calidad de los requerimientos del sistema [8].

Asimismo, es conocido que el 80 % del costo del producto está determinado por las decisiones tomadas en colaboración por las partes interesadas en las primeras etapas del ciclo de vida de los requerimientos. Por lo tanto, la gestión efectiva del diseño del producto es la capacidad de las partes interesadas para comprender lo que se necesita y se espera del producto, ambas juegan un papel importante en todo el ciclo de vida del desarrollo del software [6]. La construcción colaborativa del lenguaje de dominio implica la existencia de descripciones de mejor calidad. Pero, como en cualquier trabajo colaborativo, aparecen conflictos que necesitan ser resueltos para aclarar las descripciones con menos contradicciones.

Goel, afirma, que el lenguaje natural es la única notación que pueden leer y entender las partes interesadas, fomentando así su participación activa, fundamental en los primeros pasos del desarrollo de software y además que el lenguaje natural se usa ampliamente para modelar la información de los requerimientos [9]. Escribir requerimientos es esencialmente un trabajo colaborativo [10] [11], el cliente conoce realmente el problema y el analista lo ayuda a expresarlo en forma correcta y completa. El lenguaje natural es, indudablemente, el lenguaje más natural que el cliente puede usar para expresar lo que espera del sistema, su percepción del problema y un modelo mínimo del entorno en el que funcionará el sistema a desarrollar.

Si bien, existen varios enfoques y criterios para la validación de requerimientos expresado en los más variados formalismos, ninguno de ellos puede garantizar la equivalencia real de lo expresado en los requerimientos con los deseos del cliente [12]. Solo el cliente puede dar la

aprobación final sobre los requerimientos, especialmente en vista de cualquier implicación legal derivada de esta aprobación, por lo que debe estar en condiciones de comprenderlos, evaluarlos y validarlos. Esta necesidad, lleva a considerar los requerimientos del lenguaje natural como la versión de referencia, y elegir el lenguaje natural en sí mismo como la forma preferida de comunicarse con el cliente.

Las técnicas actuales de procesamiento del lenguaje natural desde los primeros intentos exitosos, han evolucionado para conocer conversaciones, puntos de vista, creencias, información contractual y otros puntos sutiles [13]. Los sistemas de comprensión del lenguaje natural están actualmente en funcionamiento, y la traducción de texto casi automática se ha vuelto factible. Sin embargo, esta evolución se ha obtenido a costa de una complejidad considerable. Estos sistemas incluyen, generalmente, componentes de inteligencia artificial, grandes cantidades de información semántica, datos estadísticos, bases de conocimiento, probadores de teoremas, entre otros, con el objetivo de inferir la mayor cantidad de información contextual posible del texto de origen [14].

En esta línea de investigación se analiza el impacto de la utilización de técnicas de procesamiento de lenguaje natural aplicada al proceso de validación de requerimientos en entornos colaborativos.

Líneas de Investigación, Desarrollo e Innovación

- ☒ Enfoques de Validación de Requerimientos
- ☒ Técnicas de Validación de Requerimientos

- ☒ Cualidades de la SRS a evaluar en el proceso de Validación de Requerimientos.
- ☒ Técnicas de procesamiento de lenguaje natural.
- ☒ Procesos Colaborativos.

Resultados obtenidos/esperados

Los resultados obtenidos / esperados se pueden resumir en:

- ☒ Se ha avanzado en una revisión bibliográfica de las principales tendencias de la Validación de Requerimientos del software desde el año 2007 hasta el año 2021. Los artículos analizados fueron obtenidos desde las fuentes IEEE, Elsevier, Springer y ACM Digital Library. Se han preseleccionado 38 trabajos para posteriormente centralizarse en 4 metodologías [15].
- ☒ Se ha avanzado en el análisis comparativo de cuatro metodologías con el fin de identificar las contribuciones al proceso de Validación de Requerimientos [15]:
 - ☒ Funciones, componentes, entornos y características.
 - ☒ Actividades de planificación.
 - ☒ Técnicas de control e indicadores rendimiento.
 - ☒ Definición de estándares.
 - ☒ Aceptación del cliente/usuario.
 - ☒ Dominios de aplicación.
 - ☒ Participación cliente/usuario.
 - ☒ Etapas del ciclo de vida del desarrollo del software donde validan los requerimientos.
- ☒ Se ha avanzado sobre la elaboración de buenas prácticas para el proceso de validación de requerimientos [16]:

- ☒ Determinación de características que se resumen en términos de recomendaciones de buenas prácticas que pueden mejorar el conocimiento en el desarrollo de enfoques para validación de requerimientos.
- ☒ Se ha avanzado el análisis de diversos enfoques de validación de requerimientos sobre las características asociadas al proceso de validación de requerimientos en el ciclo de vida del software: la naturaleza de la información Quién, Qué, Cuando, Por qué y Cómo validar los requerimientos [17]:
 - ☒ Definición de dominios específicos de aplicación
 - ☒ Participación de los usuarios/clientes
 - ☒ Cumplimiento de estándares
 - ☒ Nivel de formalidad, solución y evaluación de la especificación de requerimientos
 - ☒ Utilización de criterios de calidad de los requerimientos
 - ☒ Desempeño de las técnicas de validación de requerimientos
- ☒ Se ha avanzado el análisis del comportamiento de técnicas de validación en el contexto de diferentes dominios de aplicación de software [18]:
 - ☒ Utilización de técnicas de validación
 - ☒ Análisis de las características de los requerimientos 3C (completitud, corrección y consistencia)
 - ☒ Relación sobre el uso de las técnicas de validación y las características de los requerimientos 3C en diferentes dominios de aplicación del software
 - ☒ Uso de herramientas para la validación de requerimientos

- ☒ Avanzar en la capacitación continua de los miembros de la línea de investigación.
- ☒ Avanzar en el aprendizaje de procesos de Ingeniería de Requerimientos.
- ☒ Avanzar en el aprendizaje de técnicas de procesamiento de lenguaje natural teniendo como finalidad aplicarlas al proceso de Validación de Requerimientos.
- ☒ Avanzar en el estudio de las cualidades de la SRS a evaluar en el proceso de Validación de Requerimientos.
- ☒ Avanzar en el estudio de procesos colaborativos teniendo como finalidad aplicarlos al proceso de Validación de Requerimientos.

Formación de Recursos Humanos

Este estudio prevé al menos, dos proyectos de Trabajo Final y el dictado de cursos de créditos de la carrera Licenciatura en Sistemas y la realización de un trabajo de tesis de maestría en la Facultad de Informática de la UNLP.

Referencias

1. P. A. Laplante: Requirements Engineering for Software and Systems, CRC Press (2019).
2. B. H. C. Cheng, J. M. Atlee: Current and Future Research Directions in Requirements Engineering, Design Requirements Engineering A Ten-Year Perspective, Lecture Notes in Business Information Processing, vol. 14, pp. 11–43 (2019).
3. S. L. Pfleeger: Software Engineering – Theory and Practice, Prentice Hall (1998).
4. P. Loucopoulos, V. Karakostas: System Requirements Engineering, McGraw-Hill, London, ISBN 0-07-707843-8 (1995).

5. A. Fruhling, L. Steinhauser, G. Hoff, C. Dunbar: Designing and Evaluating Collaborative Processes for Requirements Elicitation and Validation, 40th Annual Hawaii International Conference on System Sciences (HICSS'07), Waikoloa, HI, USA, pp. 15-15, (2007).
6. X. Sun, Y. Zeng, W. Liu, Formalization of design chain management using environment- based design (EBD) theory, Journal of Intelligent Manufacturing (2011).
7. J. Konate, A. Sahraoui, and G. Kolfshoten: Collaborative requirements elicitation: A processcentred approach. Group Decision and Negotiation, 23(4):847–877, (2014).
8. A. Azadegan, X. Cheng, F. Niederman, and G. Yin: Collaborative requirements elicitation in facilitated collaboration: report from a case study, 46th Hawaii International Conference on System Sciences, pp. 569–578, ISSN 15301605, IEEE, (2013).
9. S. Goel: Transformation from LEL to UML, International Journal of Computer Applications, vol. 48, no. 12, (2012).
10. V. Ambriola, V. Gervasi: An environment for cooperative construction of natural-language requirement bases. In Proceedings of the Eighth Conference on Software Engineering Environments. IEEE Computer Society Press, (1997).
11. K. El Emam, S. Quintin, N. H. Madhavji: User participation in the requirements engineering process: an empirical study, Requirements Engineering Journal, 1 (1), (1996).
12. R. A. DeMillo: Software Testing and Evaluation, The Benjamins Cummings Publishing Company, Inc., (1987).
13. J. Allen: Natural Language Understanding, Addison Wesley, 2da. Edition (1994).
14. V. Ambriola, V. Gervasi: Processing natural language requirements, Proceedings 12th IEEE International Conference Automated Software Engineering, Incline Village, NV, USA, pp. 36-45, (1997).
15. Sonia R. Santana, Leandro Antonelli, Pablo Thomas: Evaluación de metodologías para la validación de requerimientos, XXVII Congreso Argentino de Ciencias de la Computación (CACIC), pp. 419-428, ISBN 978-987-633-574-4, (2021).
16. S.R. Santana, L.R. Antonelli, P.J. Thomas: Best Practices for Requirements Validation Process. In: Pesado, P., Gil, G. (eds) Computer Science – CACIC 2021, CACIC 2021, Communications in Computer and Information Science, vol 1584, Springer, Cham, https://doi.org/10.1007/978-3-031-05903-2_10, (2022).
17. Sonia R. Santana, Leandro Antonelli, Pablo Thomas: Un análisis de enfoques de Validación de Requerimientos, Trabajo de especialización de Ingeniería de Software, Facultad de Informática, Universidad Nacional de la Plata, <http://sedici.unlp.edu.ar/handle/10915/141871>, (2022).
18. Sonia R. Santana, Lucrecia Perero, Alejandro Fernandez, Leandro Antonelli: Un análisis del comportamiento de técnicas de validación en el contexto de diferentes dominios de aplicación de software, XX Jornadas de Administración e Informática, Universidad Nacional de Entre Ríos, Facultad de Ciencias de la Administración, ISBN 978-950-698-538-7, (2022).

Computación afectiva aplicada a la interacción con Interfaces Gráficas de Usuario

M. Claudia Albornoz, Mario Berón, Germán Montejano

Departamento de Informática/Universidad Nacional de San Luis-U.N.S.L./San Luis/Argentina

Ejército de los Andes 950, Tel: +54 (0266) 4520300; int 2102

{albornoz,mberon,gmonte}@unsl.edu.ar

Resumen

La Interfaz Gráfica de Usuario (GUI por su nombre en inglés, Graphical User Interface) es parte fundamental de cualquier aplicación; al comenzar a trabajar con cualquier producto de software, el usuario comienza a interactuar con la interfaz. Es donde comienza la interacción hombre-computadora. El diseño de la interfaz no se lo debe considerar como una tarea secundaria y sin importancia; por el contrario, el equipo de desarrollo debe contar con integrantes especializados en el tema.

En la actualidad, con la virtualidad impuesta a nivel mundial por la pandemia del COVID-19, se debe ofrecer al usuario una interfaz que lo ayude a concretar las tareas de manera rápida, sencilla y satisfactoria. Es la interfaz la responsable de ofrecer una interacción fluida y agradable. Muchas veces, a partir de la interfaz se puede determinar si una aplicación será utilizada o no para resolver los problemas para los cuales fue diseñada.

El objetivo del presente trabajo es mostrar la importancia del diseño de la interfaz desde el punto de vista del usuario; explicar qué es la Interacción Persona Computadora (I.P.C.), fundamentar la importancia del Diseño Centrado en el Usuario y la Ingeniería de la Usabilidad y cómo lograr una interacción satisfactoria entre el usuario y la computadora, según los estados emocionales del usuario.

Palabras Claves:

Interfaz Gráfica de Usuario (GUI), Interacción Persona-Computadora (IPC), Usabilidad, Computación Afectiva

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: “Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”. Dirección: Daniel Riesco código P-032020 de la Facultad de Ciencias Físico, Matemáticas y Naturales, de la Universidad Nacional de San Luis. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se han logrado importantes vínculos con diferentes universidades a nivel internacional y nacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. Introducción

La interfaz es la parte visible de todo producto de software; es la responsable de transmitir o hacerle saber al usuario lo que es capaz de hacer el producto. En la actualidad, la virtualidad se convirtió en protagonista de la vida diaria de las personas. La pandemia del COVID-19 obligó al mundo entero a interactuar con productos de software. Para aprovechar los beneficios de la tecnología, debe existir una buena interacción entre el usuario y la computadora. Pensando en los usuarios se debe diseñar la interfaz [1].

La interfaz se la puede definir como la parte visible de una computadora y su software; que el usuario puede ver, oír, tocar, hablar, o de otra manera entender o dirigir [2].

Cualquier usuario, sin previo conocimiento, puede interactuar con un producto si la interfaz está bien diseñada y

construida. Una interfaz mal diseñada obstaculiza la ‘usabilidad’, es decir aquella característica ‘que hace que la aplicación sea fácil de utilizar y fácil de aprender’ [2].

Al hablar de ‘usabilidad’ se puede considerar que la Interfaz presenta las siguientes características (Fig. 1):

- *Satisfacción*: que el usuario trabaje en un entorno en el que se sienta cómodo.
- *Efectiva*: hace lo que se espera que haga.
- *Eficiente*: realiza las tareas en el tiempo necesario y sin errores.



Fig 1: Usabilidad

El objetivo es lograr una interfaz ‘usable’ para mejorar la Interacción Persona Computadora, concepto que se tratará en la siguiente sección.

2. I.P.C.: Interacción Persona Computadora

Las personas interactúan con las interfaces constantemente: al usar el celular, en el cajero automático, la computadora, el cajero automático, el GPS del auto, etc. Además de la interacción física que realiza el usuario (al teclear, al mover el mouse, al tocar la pantalla digital, etc.) se debe tener en cuenta el nivel cognitivo necesario para que el usuario comprenda el protocolo de interacción que presenta la interfaz [3].

Para el usuario sólo existe la interfaz, no le importa cómo se concretan las tareas, cómo están programadas o implementadas. Los usuarios, se han convertido en clientes exigentes y críticos: esperan un alto grado de elaboración en las Interfaces Gráficas de Usuario (GUI), y luego le dan importancia

al funcionamiento del sistema y al código [9].

Gracias a la invasión tecnológica en la vida de los usuarios, en los últimos tiempos surgió una nueva disciplina que se encarga de estudiar cómo lograr una fluida interacción usuario-máquina. Es lo que se conoce como IPC (Interacción Persona Computadora o IPO por Interacción Persona Ordenador).

Se puede considerar que la IPC es el punto de encuentro de las ciencias humanas (como la psicología, la pedagogía, sociología y otras) con la tecnología o las ciencias exactas como la informática. Se focaliza en estudiar la interacción entre usuarios y sistemas informáticos; su objetivo es proporcionar bases teóricas, metodológicas y prácticas para el diseño y evaluación de aplicaciones interactivas. Esta tarea se realiza para que puedan ser usadas de forma eficaz, segura, eficiente y satisfactoria [4].

Como definición, se puede decir que: la interacción persona computadora es el intercambio observable de información, datos y acciones entre un humano y la computadora, y viceversa [5]. La IPC surge para educar a los primeros trabajadores que se vieron obligados a usar la computadora, ya sea en organismos gubernamentales o grandes empresas. Con el tiempo se fue perfeccionando con el fin de ayudar a todos los usuarios, incluyendo personas con capacidades diferentes [6]. Al referirse a proporcionar bases teóricas, metodológicas y prácticas para el diseño y evaluación de aplicaciones interactivas, se habla de Ingeniería de la Usabilidad y el Diseño Centrado en el Usuario. Ambos abarcan un conjunto de procesos y metodologías que aseguran que se cumplan los niveles de usabilidad requeridos en la aplicación. Conceptos que se tratarán a continuación.

2.1 Ingeniería de la Usabilidad

La Ingeniería de la Usabilidad es multidisciplinar; se nutre de la informática,

de la psicología, de la lingüística, de la sociología, de la antropología y del diseño industrial. Este término se utiliza desde la década del 80 para designar a una nueva disciplina, que se ocupa de proporcionar “métodos sistemáticos y herramientas para la compleja tarea de diseñar interfaces de usuario que sean fáciles de comprender, rápidamente aprendibles y fiablemente operables” [7]. Para el usuario la Interfaz es ‘la aplicación’ o ‘el sistema’ en sí, porque es lo que ve y con lo que interactúa. Si la interacción no es fluida, se la considera inútil, no efectiva, las funcionalidades y la utilidad son limitadas. Cuando esto ocurre el usuario se confunde y se frustra.

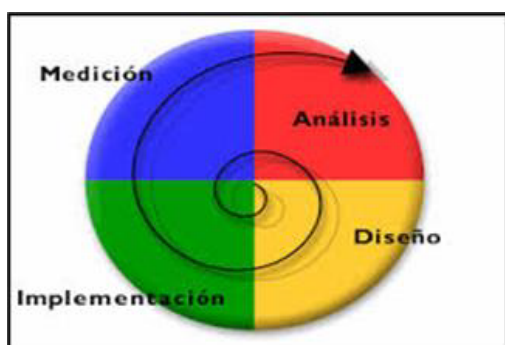


Fig 2: Metodología conceptual y esquemática de la Ingeniería de la Usabilidad.

El objetivo de la Ingeniería de la Usabilidad es minimizar la sobrecarga cognitiva y perceptiva de los usuarios. Utiliza un método de diseño iterativo con prototipado rápido (necesariamente deberá contar con herramientas de ayuda), cuyo ciclo es: "análisis, diseño, implementación, evaluación" (Fig. 2), que se repite varias veces con el fin de ir mejorando progresivamente la aplicación. La etapa de evaluación del prototipo, la cual se realiza con usuarios reales a cada repetición del ciclo, es de suma importancia para obtener resultados dignos de una ingeniería. En esta etapa el usuario es el protagonista.

2.2 Diseño Centrado en el Usuario

El Diseño Centrado en el Usuario (DCU) es el término que se utiliza para

describir el diseño en el que el usuario influye en el resultado final. Se considera que es una filosofía y un proceso. Una filosofía porque sitúa al usuario en el centro con la intención de desarrollar un producto adecuado a sus requerimientos, necesidades y un proceso de diseño porque se centra en los factores cognitivos de las personas y cómo éstos intervienen en sus interacciones con los productos [8].

El usuario final se ve involucrado en cada etapa o fase del proceso de desarrollo, con esto se garantiza que el producto se ajuste a sus necesidades. Ésta forma de diseño le otorga al usuario el rol central del desarrollo. En ocasiones, el usuario es un miembro más del equipo de diseño. El objetivo es obtener productos fáciles de usar, efectivos y eficientes.

Esta metodología consta de diferentes etapas (Fig. 3) [10]. Si bien estas etapas corresponden a todo el proceso de desarrollo, es obvio que el diseño y desarrollo de la interfaz se ven involucrados en ellas. Como es evidente, el usuario se ve involucrado en cada etapa o fase.

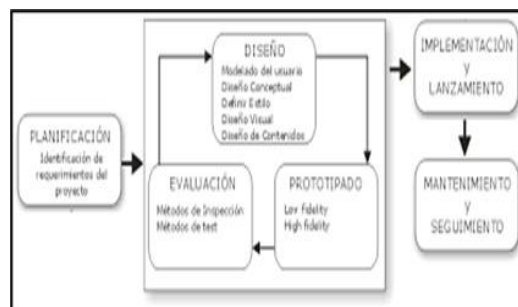


Fig. 3: Etapas del DCU

3. Computación Afectiva: las Emociones

En los últimos tiempos han surgido nuevas investigaciones para lograr no solo una interacción fluida, entre el usuario y la computadora, sino también en que sea satisfactoria y porque no, placentera.

Es por esto que a los usuarios no se los considera sólo como seres con habilidades cognitivas y físicas, también se deben tener en cuenta los estados emocionales (miedos,

esperanzas, valores, etc.). Hoy el objetivo es lograr una excelente interacción entre el usuario y la computadora; así surge una nueva disciplina: la Computación Afectiva.

Esta disciplina es una rama de la Inteligencia Artificial, desarrolla métodos computacionales orientados a reconocer emociones humanas y generar emociones sintéticas. Surge ante la necesidad de optimizar la interacción entre personas y computadoras, pero también se inscribe en la investigación de los procesos inteligentes [11]. La afectividad es fundamental en el comportamiento y la comunicación de las personas. La interacción humana siempre incluye emociones, estados de ánimo, afectos los cuales se transmiten de manera explícita (verbalmente) o implícita (no verbal) a través de gestos, expresiones, actitudes [12]. Esta información que se transmite con actitudes, expresiones y gestos es de gran valor y produce un gran efecto en la comunicación, aún en la comunicación usuario-computadora.

El hecho de que las computadoras puedan comprender nuestras emociones y a su vez que puedan “expresar” (o simular) emociones propias, sería un paso importante para establecer un cambio cualitativo en la interactividad.

Actividades Llevadas a Cabo en la Línea de Investigación

En esta línea de investigación se han llevado a cabo las siguientes tareas:

Revisión Sistemática de la literatura referente al tema de diseño y construcción de Interfaces Gráficas de Usuario; con el fin de evidenciar la importancia de su diseño y cómo lograr una óptima y satisfactoria interacción con los usuarios. Haciendo una mención especial respecto al rol del usuario en el diseño de la interfaz.

Estudio de Métodos de Evaluación Multicriterio. Esta tarea tiene como objetivo seleccionar un método de evaluación para evaluar interfaces la calidad de las interfaces gráficas de usuario.

Hasta el momento se ha encontrado que el método LSP (Logic Scoring of Preference) se adapta muy bien para la evaluación de las GUI.

Elaboración de Criterios de Evaluación. Esta tarea se lleva a cabo con el objetivo de establecer cuáles son las características que las interfaces gráficas de usuario deben tener para que sean consideradas adecuadas para su utilización por parte del usuario. Estos criterios serán la entrada al método de evaluación multicriterio mencionado en el ítem precedente.

Estudio de Reglas, Técnicas y Principios de Diseño de GUI. Esta tarea provee la información necesaria para que los equipos de desarrollo de GUI puedan desarrollar GUIs adecuadas para el usuario.

Análisis de la Influencia Emocional, el objetivo es determinar cómo influye el estado emocional de los usuarios en la interacción con las computadoras y, a su vez, cómo el diseño de la Interfaz influye en el estado emocional del usuario.

Conclusiones y Trabajos Futuros

Luego de la lectura de diferentes trabajos respecto de cómo mejorar la interacción usuarios-computadoras se puede observar que el diseño de una Interfaz Gráfica de Usuario no es una tarea secundaria. Además, no sólo involucra un grupo de expertos en programación; sino que es fundamental que el equipo integre a los usuarios. Diversos estudios demuestran que cada vez son más las investigaciones que se realizan con el objetivo de lograr una óptima interacción. No sólo dejando en claro qué es lo que necesita el usuario (requerimientos), sino también lo que desea y sus estados emocionales.

Como resultado, esta investigación pretende determinar cómo lograr una GUI ‘usable’, comprensible, satisfactoria; con el análisis de los diferentes tópicos que se involucran en la Interacción Persona Computadora.

Futuros trabajos: i) Enriquecer el estudio sistemático de la literatura; ii) Implementar un método de evaluación multicriterio para evaluar GUIs, iii) Definir criterios de evaluación de interfaces gráficas de usuario, iv) Elaborar un manual con las Reglas, Técnicas y Principios de Diseño de GUI y v) Profundizar el estudio sobre la influencia emocional.

Formación de Recursos Humanos

Los progresos obtenidos en esta línea de investigación sirven como base para el desarrollo de tesis de posgrado, ya sea de doctorado, maestrías o especializaciones en Ingeniería de Software; y desarrollo de trabajos finales de las carreras Licenciatura en Ciencias de la Computación, Ingeniería en Informática e Ingeniería en Computación de la Universidad Nacional de San Luis, en el marco de los Proyectos de Investigación.

Referencias

- [1] Chamba, Carrión, and Rober Alonso. *Estándar de usabilidad para la interfaz gráfica de usuario en los proyectos de desarrollo de software*. BS thesis. 2014.
- [2] Galitz, Wilbert O. *The essential guide to user interface design: an introduction to GUI design principles and techniques*. Wiley.com, 2007.
- [3] Toni Granollers i Saltiveri. *MPlu+ a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares*. Universitat de Lleida, 2007. ISBN: 978-84-690-6378-1
- [4] Mascheroni, M. A., Greiner, C. L., Petris, R. H., Dapozo, G. N., and Estayno, M. G. *Calidad de software e ingeniería de usabilidad*. 2012. In XIV Workshop de Investigadores en Ciencias de la Computación.
- [5] Díaz, F. J., Harari, I., and Amadeo, A. P. *Guía de recomendaciones para diseño de software centrado en el usuario*. 2013. Editorial de la Universidad Nacional de La Plata (EDULP).
- [6] Olson, G. M. and Olson, J. S. Human-computer interaction: Psychological aspects of the human use of computing. 2003. Annual review of psychology, 54(1):491-516.
- [7] Lorés, Jesús, and Toni Granollers i Saltiveri. *La Ingeniería de la Usabilidad y de la Accesibilidad aplicada al diseño y desarrollo de sitios web*. Universitat de Lleida, 2004.
- [8] Domingo, Muriel Garreta, and Enric Mor Pera. *Diseño centrado en el usuario*, 2010.
- [9] Marín, José M. Belmonte. *Ingeniería de la Usabilidad. Aplicada al desarrollo de un portal web administrado dinámicamente*. Trabajo Final de Carrera. Universitat de Lleida. 2003.
- [10] Yusef Hassan & Francisco J. Martín Fernández & Ghzala Iazza. *Diseño Web Centrado en el Usuario: Usabilidad y Arquitectura de la Información*. "Hiptertext.net", núm. 2, 2004. <<http://www.hiptertext.net>>
- [11] Causa, Emiliano, and Andrea Sosa. *La computación afectiva y el arte interactivo*. Área Transdepartamental de Artes Multimediales 52 (2007).
- [12] González, J. A., Gardeazabal, L., and Vitoria, N. G. *Mediación emocional aplicada en sistemas de comunicación aumentativa y alternativa*. Inteligencia Artificial: revista iberoamericana de inteligencia artificial, 6(16):65-70. 2002.

Calidad de información en comunidades virtuales

Valeria Zoratto, Gabriela Aranda, Nadina Martinez Carod, Romina Schroeder
Andrés Flores, Natalia Baeza, Lucas Cavaliere, Sandra Lucero

Grupo de Investigación en Ingeniería de Software del Comahue (GIISCO)
<http://giisco.uncoma.edu.ar>

Facultad de Informática. Universidad Nacional del Comahue
Buenos Aires 1400, (8300) Neuquén

Contacto: {vzoratto, gabriela.aranda, nadina.martinez}@fi.uncoma.edu.ar

RESUMEN

La Web 2.0 se caracterizó por no ser simplemente un contenedor de información, sino en un plataforma de trabajo colaborativo donde se desarrollan servicios web, herramientas y aplicaciones que permitieron a los usuarios crear y compartir contenido en la web, como blogs, wikis, redes sociales y plataformas de video y música en línea, sin necesidad de conocimientos técnicos especializados. Al mismo tiempo, emergieron diversas plataformas para la creación de redes sociales con objetivos diversos, como el trabajo en colaboración a distancia, así como de intercambio de conocimiento técnico, académico, científico y social. Estas comunidades virtuales, construidas a través de la participación en línea, fueron reconocidas como un fenómeno sin precedentes en la historia de la comunicación digital.

La Web 3.0 es la siguiente etapa en la evolución de la web y se centra en la creación de una web semántica, en la que la información se organiza y se presenta de una manera más inteligible y útil para los usuarios. La Web 3.0 también se enfoca en la accesibilidad de la información, independientemente del dispositivo utilizado, y en la forma en que las personas interactúan con ella para obtener resultados precisos y relevantes.

Como parte de este tránsito hacia la Web 3.0, nuestro proyecto se centra en proporcionar

modelos de calidad para sistemas de software que permitan la recuperación, análisis, clasificación y reutilización de la información proveniente de comunidades virtuales en la Web.

Palabras clave

Recuperación de información, calidad de datos, comunidades virtuales, participación ciudadana.

CONTEXTO

Esta línea de investigación forma parte del proyecto de investigación “Reuso de información en comunidades virtuales”, de la Universidad Nacional del Comahue, con período de vigencia 2022-2025. Dicha línea extiende y avanza sobre temas desarrollados por el equipo de investigadores principales en proyectos pertenecientes al Programa “Desarrollo de Software basado en Reuso”, Parte I y II, de la Universidad Nacional del Comahue, llevado a cabo entre los años 2013 y 2021. En este nuevo proyecto se continúan las líneas de investigación enfocadas en la recuperación de información disponible en foros de discusión y abarca nuevas tecnologías para soporte a comunidades virtuales con una mirada orientada a la participación ciudadana y la toma de decisiones basada en opinión pública.

1. INTRODUCCIÓN

En 1993, el término "comunidades virtuales" fue acuñado por Howard Rheingold para describir grupos sociales que se forman a través de la interacción en línea en espacios públicos de Internet [1]. Desde entonces una variedad de entornos colaborativos han surgido, incluyendo plataformas de redes sociales como Facebook, Twitter e Instagram, así como las Comunidades de Preguntas y Respuestas (CQA, por sus siglas en inglés).

Las CQA, como los foros de discusión, permiten a los usuarios buscar y compartir conocimientos a través de preguntas y respuestas. Estos sitios han crecido en popularidad en todo el mundo, y son utilizados diariamente por millones de usuarios para encontrar respuestas a preguntas complejas, subjetivas o específicas de un contexto [2, 3]. Sin embargo, con el creciente volumen de información en estos sitios, surge la necesidad de analizar y reutilizar esta información [4, 5]. La integración y reutilización de información (IRI, por sus siglas en inglés) juega un papel clave en la captura, representación, mantenimiento, integración, validación y extrapolación de información que se puede aplicar para mejorar la toma de decisiones en varios dominios de aplicación [6].

En relación a lo anterior, se han llevado a cabo diversos estudios para analizar, evaluar y extraer información de las CQA. Por ejemplo, algunos se centran en analizar los hilos de discusión para determinar los criterios que definen la calidad de las respuestas [7, 8]. Otros investigan técnicas de respuesta a preguntas en CQA y clasifican los sistemas de respuesta en dos categorías: sistemas basados en votación [3] y sistemas basados en modelos [27]. Además, algunos autores sostienen que la calidad de una respuesta se puede evaluar únicamente considerando las características textuales de los hilos [9, 10, 11], mientras que otros combinan dichas características con la red de usuarios [12, 13]. También se ha investigado la experticia de los usuarios que responden, estimando su

experiencia basándose en evidencias como la calidad de sus intervenciones y su interrelación con otros usuarios [14]. En resumen, tanto el análisis del contenido textual de los mensajes como el estudio de las redes sociales subyacentes son temas muy relacionados con este tipo de investigación.

En relación al análisis de redes sociales en CQA, es importante destacar que este tipo de análisis requiere de una combinación de métodos y técnicas que involucran teorías sociológicas y matemáticas. Para analizar las estructuras sociales de estas redes, se utilizan los conceptos, vocabulario y operaciones de la teoría de grafos que permiten probar teoremas sobre los grafos que las modelan, así como deducir y someter determinados enunciados a pruebas [15]. Además, los estudios más recientes en esta área se enfocan en la información proveniente de redes sociales como Twitter [16]. Sin embargo, dada la diversidad de intereses y características de cada red, es importante evaluar el comportamiento de las personas en otros tipos de plataformas colaborativas. Por ejemplo, es relevante identificar a los usuarios que emiten información irrelevante, a los líderes de opinión que son considerados fuentes de conocimiento dentro de su comunidad, o a los boundary spanners, es decir, usuarios que permiten la comunicación entre distintas comunidades [17]. De esta manera, se puede entender mejor la dinámica social y de conocimiento en CQA y cómo estas pueden ser aprovechadas para mejorar la calidad de las respuestas y la experiencia de los usuarios.

Como se ha mencionado, el gran volumen de información generado por las CQA existentes en la Web es propicio para el estudio y definición de técnicas para el reuso e integración de información [6], por lo que nuestro proyecto se enfoca en definir métodos para capturar información, realizar análisis de contenido, así como detectar y clasificar perfiles de usuarios, utilizando para su evaluación corpus de comunidades virtuales

existentes como StackExchange¹, que es una red de webs para CQA que cuenta con más de 173 foros de discusión de diferentes temáticas, y que solo por mes recibe más de 100 millones de usuarios que realizan preguntas o responden a otros. Además, sus bases de datos son accesibles de forma libre para investigación². Luego, el conocimiento adquirido puede ser aplicado en comunidades virtuales más específicas, como por ejemplo, las conformadas a partir de plataformas para la participación ciudadana, que han surgido en los últimos años a partir de los desafíos que enfrentan las ciudades para garantizar la calidad de vida de sus habitantes y mejorar los procesos de toma de decisiones incorporando la opinión pública. Por ello, la participación ciudadana es una herramienta que mejora la gobernanza local y la toma de decisiones, que busca ser una forma directa para conocer las necesidades, demandas e ideas de los individuos que la componen [18]. En muchos casos la toma de decisión se realiza a partir de la opinión de la ciudadanía obtenida mediante herramientas colaborativas, como sitios web y/o aplicaciones móviles, que permiten conocer a corto plazo los efectos que pueden tener los cambios realizados, por ejemplo, en el espacio urbano. Dado que dichas tecnologías se basan en recuperar opiniones de ciudadanos, surgen temáticas de análisis y evaluación que, muchas veces son comunes a otras comunidades virtuales como las que hemos mencionado anteriormente. Con este objetivo, como parte de este proyecto nos proponemos evaluar productos existentes para participación ciudadana (como *decidim*³, *CONSUL*⁴ y *WeLive*⁵ [19, 20]) y trabajar en su adaptación para aplicarlos en el ámbito de barrios de la ciudad de Neuquén. Además, se planea hacer uso de las técnicas y herramientas elaboradas como parte de otras líneas de investigación, para hacer aportes al

reuso de información como soporte para la toma de decisiones basada en la opinión de los ciudadanos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto actual se denomina “Reuso de información en comunidades virtuales” y su objetivo principal es definir técnicas y algoritmos de recomendación para asistencia inteligente a usuarios de comunidades virtuales en la búsqueda de información relevante.

Este proyecto está desarrollado por integrantes del Grupo de Ingeniería de Software de la Universidad Nacional del Comahue, (GIISCo), formado por docentes y estudiantes de la Facultad de Informática de la Universidad Nacional del Comahue, junto con asesoría y colaboración de la Facultad de Ciencias Exactas de la Universidad Nacional del Centro de la Provincia de Buenos Aires, y una colaboradora de la Facultad de Humanidades de la Universidad Nacional del Comahue que se desempeña en proyectos de investigación en participación ciudadana. Además, la participación de docentes pertenecientes a otras áreas de la Facultad (Programación, Ingeniería en Computación, Ingeniería en Sistemas y Teoría de la Computación), permiten abordar la investigación desde ópticas diferentes, enriqueciendo el desarrollo con un trabajo conjunto y colaborativo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Desde 2013, nuestro equipo de investigación ha estado trabajando en temas relacionados con las comunidades virtuales de programadores. En particular, hemos propuesto un modelo de calidad para foros de discusión técnicos y hemos determinado criterios para evaluar la calidad de la información contenida en los hilos de discusión [22]. Estos criterios han sido

¹ <https://stackexchange.com/>

² <https://archive.org/download/stackexchange>

³ <https://decidim.org/es/>

⁴ <https://consulproject.org/en/>

⁵ <https://welive.eu/>

validados mediante encuestas, y hemos formulado variaciones en la parametrización para mejorar los resultados obtenidos [23].

También hemos trabajado en el procesamiento de texto de hilos de discusión en foros técnicos. Para ello, hemos desarrollado una herramienta ad-hoc basada en Lucene, que nos ha permitido recuperar información y analizarla según un conjunto de medidas de calidad que hemos definido, con el fin de proponer un ranking de soluciones posibles para una pregunta [24]. Posteriormente, hemos mejorado esta herramienta incluyendo sinónimos en la base de datos léxica WordNet y el analizador morfológico Stanford POS Tagger para identificar el rol de las palabras en el contexto en el que son utilizadas [25].

Además, hemos avanzado en la clasificación de roles de usuarios activos en un foro, con el objetivo de determinar la jerarquía de roles basados en el nivel de conocimiento de los participantes en los hilos de discusión, según los posts realizados por dichos usuarios [26].

Gracias a los conocimientos adquiridos en nuestros estudios anteriores, y a la incorporación de herramientas de participación ciudadana pretendemos aplicarlos en el análisis de información proveniente de comunidades virtuales creadas específicamente para los barrios de nuestra ciudad. De esta manera, esperamos poder contribuir a la toma de decisiones basadas en la opinión pública.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto se encuentra conformado por docentes de diferentes áreas debido a su naturaleza multidisciplinaria. Las personas que forman parte del proyecto, tanto como colaboradores, asesores o integrantes son:

- Tres docentes investigadores del Departamento de Programación e

Ingeniería de Sistemas, con dedicación exclusiva, con Doctorado en Informática.

- Una docente con dedicación exclusiva del Departamento de Programación, finalizando el Doctorado en Ciencias de la Computación.
- Tres docentes con dedicación simple, pertenecientes a los Departamentos de Programación, Ingeniería de Sistemas e Ingeniería en Computación.
- Una profesora adjunta, asesora local, con dedicación exclusiva, del Departamento de Teoría de la Computación.
- Una docente de la Facultad de Humanidades de la misma universidad, investigadora del Instituto Patagónico de Estudios de Humanidades y Ciencias Sociales - CONICET.
- Una docente investigadora externa, perteneciente al Instituto Superior de Ingeniería del Software (ISISTAN) de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), con experiencia en Sistemas de Recomendación y Recuperación de Información. Doctora en Ciencias de la Computación.
- Tres estudiantes de la carrera de Licenciatura en Ciencias de la Computación que realizan sus tesis dentro del proyecto.

De esta manera, se van incorporando actividades para extender líneas de investigación al proyecto inicial con nuevos enfoques.

5. BIBLIOGRAFÍA

- [1] H. Rheingold. The Virtual Community, revised edition: Homesteading on the Electronic Frontier. MIT press, 2000.
- [2] I. Srba and M. Bielikova. A comprehensive survey and classification of approaches for community question answering. ACM Trans. Web, 10(3), 2016.
- [3] M. Neshati. On early detection of high voted qa on stack overflow. Information Processing Management, 53(4):780–798, 2017.

- [4] G. Cong, L. Wang, C. Lin, Y. Song, and Y. Sun. Finding Question-answer Pairs from Online Forums. In Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '08, pages 467–474, New York, NY, USA, 2008. ACM.
- [5] S. Gottipati, D. Lo, and J. Jiang. Finding relevant answers in software forums. In 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011), Lawrence, KS, USA, November 6–10, 2011, pages 323–332, 2011.
- [6] M. Day, C. Ong, and W. Hsu. An analysis of research on information reuse and integration (2003–2008). International Transactions on Systems Science and Applications, 6(2):146–157, 2010.
- [7] L.T. Le, C. Shah, and E. Choi. Evaluating the Quality of Educational Answers in Community Question-Answering. In Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries, JCDL '16, pages 129–138, New York, NY, USA, 2016. Association for Computing Machinery.
- [8] L. Amancio, C. Dorneles, and D. Dalip. Recency and quality-based ranking question in CQAs: A stack overflow case study. Information Processing Management, 58(4):102552, 2021.
- [9] G. Gkotsis, K. Stepanyan, C. Pedrinaci, J. Domingue, and M. Liakata. It's all in the content: state of the art best answer prediction based on discretisation of shallow linguistic features. In Proceedings of the 2014 ACM conference on Web science, pages 202–210, 2014.
- [10] G. Burel, P. Mulholland, and H. Alani. Structural Normalisation Methods for Improving Best Answer Identification in Question Answering Communities. In Proceedings of the 25th International Conference Companion on World Wide Web, pages 673–678, 2016.
- [11] Y. Pérez-Guadarramas, A. Rodríguez-Blanco, A. S. Cuevas, W. Hojas-Mazo, y J. A. Olivas. Combinando patrones léxico-sintácticos y análisis de tópicos para la extracción automática de frases relevantes en textos. Procesamiento del Lenguaje Natural, (59):39–46, 2017.
- [12] D. Kundu and D. Prasad Mandal. Formulation of a hybrid expertise retrieval system in community question answering services. Applied Intelligence, 49(2):463–477, 2019.
- [13] H. Fu and S. Oh. Quality assessment of answers with user identified criteria and data-driven features in social qa. Information Processing Management, 56(1):14–28, 2019.
- [14] M. Neshati, Z. Fallahnejad, and H. Beigy. On dynamicity of expert finding in community question answering. Information Processing Management, 53(5):1026–1042, 2017.
- [15] L. Sanz-Menéndez. Análisis de redes sociales: O cómo representar las estructuras sociales subyacentes. Apuntes de Ciencia y Tecnología, 7:21–29, 06 2003.
- [16] R. Olivares, F. Muñoz, and F. Riquelme. A multiobjective linear threshold influence spread model solved by swarm intelligence-based methods. Knowledge-Based Systems, 212:106623, 2021.
- [17] P. Matous and P. Wang. External exposure, boundary-spanning, and opinion leadership in remote communities: A network experiment. Soc. Networks, 56:10–22, 2019.
- [18] D. García Castro, V. De Elizagarate Gutierrez, J. Kazak, S. Szewranski, I. Kaczmarek, and T. Wang. Nuevos desafíos para el perfeccionamiento de los procesos de participación ciudadana en la gestión urbana. retos para la innovación social. Management Letters/Cuadernos de Gestión, 20(1):41–64, 2020.
- [19] I. Peña-López. Shifting participation into sovereignty: the case of decidim.barcelona. 03 2019.
- [20] M. X. Rivera Rásury. Desarrollo de una herramienta de soporte metodológico a los procesos de e-participación. Master, Departamento de Sistemas Informáticos y Computación Universitat Politècnica de Valencia, Valencia, España, 2018.
- [21] J. Levy Moreno and H. H. Jennings. “Statistics of Social Configurations.” *Sociometry*, vol. 1, no. 3/4, pp. 342–374. *JSTOR*, 1938.
- [22] G. Aranda, N. Martínez Carod, S. Roger, P. Faraci, A. Cechich, V. Zoratto. Una herramienta para el análisis de hilos de discusión técnicos. CACIC 2014, Buenos Aires, pp.803–812, 2014.
- [23] G. Aranda, V. Zoratto, N. Martínez Carod, S. Roger, F. Otermin, A. Cechich. Clasificación de contenido de hilos de discusión mediante análisis sintáctico y morfológico. CICC SI 2018. ISBN 9789874568366. Mendoza, 2018, pp. 35–44.
- [24] V. Zoratto, G. Aranda, S. Roger, A. Cechich. Analyzing Discussion Forums Threads About Java Programming Language Usage, Electronic Journal of SADIO, ISSN 1514-6774, 2016.
- [25] V. Zoratto, G. Aranda, N. Martínez Carod, F. Otermin. Evaluación de estrategias para clasificar hilos de foros de discusión según su contenido, ASSE-JAIIO 2021, Argentine Symposium on Software Engineering, Argentina, 2021.
- [26] N. Martínez Carod, G. Aranda, V. Zoratto, C. Murray (2019), Una propuesta para clasificación de roles de usuarios en foros de discusión técnicos. CACIC 2019, Argentina, 2019.
- [27] Khot, T., Sabharwal, A., & Clark, P. (2017). Answering complex questions using open information extraction. arXiv preprint arXiv:1704.05572.

Accesibilidad web. Estrategias y aplicaciones.

Sonia I. Mariño, Pedro L. Alfonzo, Verónica K. Pagnoni, Silvana V. Armana, Maria L. Gronda

*Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste,
Corrientes, Argentina*

simarinio@yahoo.com, plalfonzo@hotmail.com, vero_pagnoni@hotmail.com,
sil_armana@yahoo.com.ar, lauragronda@yahoo.com.ar

*Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste,
Corrientes, Argentina*

RESUMEN

En una sociedad en que las TIC irrumpen en todos los dominios es menester asegurar el acceso a los contenidos. Por ello, en un contexto -que promueve la inclusión- las directrices, los estándares, los métodos y las herramientas de accesibilidad deben contemplarse en las distintas fases del ciclo de vida del software. Por ello, desde las normativas, se detecta una tendencia creciente orientada a promover el diseño e implementación de productos y servicios accesibles y así innovar en las interacciones que posibilitan a sus usuarios al asegurar mejores experiencias de uso.

Palabras clave: recursos humanos, AW, métodos, procedimientos, herramientas TIC.

CONTEXTO

Las líneas de I+D+i se enmarcan en el Programa Sistemas de información y TI: modelos, métodos y herramientas y en los Proyectos acreditados por la SGCyT – UNNE: Sistemas informáticos: modelos, métodos y herramientas (PI 19F014, Res. N° 1015/19 CS.) y Sistemas informáticos y gestión del conocimiento. Modelos, métodos y herramientas (PI 22F025, Res. N° 0931/22 CS UNNE).

1. INTRODUCCIÓN

Los cambios tecnológicos pueden facilitar el acceso a la comunicación y a la información para algunos sujetos, pero pueden también generar brechas más profundas y excluyentes como se menciona en [1].

Los sitios web deben diseñarse y desarrollarse considerando estándares de Accesibilidad

Web móvil con la finalidad de evitar dificultades para las personas con algún tipo de discapacidad [2]. La accesibilidad web y móvil, como elementos de calidad, deben contemplarse en las distintas fases del ciclo de vida del software, con la finalidad de innovar en las interacciones de estos artefactos y sus destinatarios al mejorar las experiencias de usuario.

Contemplando los antecedentes del equipo en esta temática de trascendencia, se abordan las estrategias de AW dado que, a través de los estándares, los métodos y las herramientas se contribuye a su internalización y adaptación, lo que se ve reflejado en un sinnúmero de aplicaciones TI que emergen en la economía del conocimiento.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La relevancia de continuar profundizando en temas de Accesibilidad Web (AW) tiene una perspectiva holística, multifacética, tratada como una mirada integral [3]. Por ello, para contribuir a la construcción de espacios más accesibles a los contenidos de la WWW, los principales ejes temáticos en investigación y desarrollo comprenden:

- la revisión sistemática de la literatura, con la finalidad de actualizar el marco teórico de referencia, y que proporciona posibles elementos de interés y necesidad de estudio y validación, y los argumentos que apoyan los hallazgos
- la teoría, considerando los estándares de AW [4-6] propuestos por el W3C [7] y otras organizaciones que abordan la temática, entre los que mencionan [8-14] y las legislaciones [15-19]

- la metodología, caracterizada -por ejemplo- en estudios empíricos, descriptivos y exploratorios, transversales o longitudinales, entre otros tipos [20], y que situados en un contexto socio-histórico-cultural influye en su abordaje.
- la Ingeniería del Software -y según algunas áreas de la guía SWEBOK [21, 22] - como un aspecto de la calidad de software en contextos públicos y privados, regionales, nacionales, en que se desempeñan los participantes; desde el ciclo de vida del software, incorporando a la AW en todas las fases; desde el mantenimiento del software y centrándose en el mantenimiento correctivo; desde la ingeniería de requerimientos como aquellos no funcionales y que afectan a la calidad del software; desde los métodos, técnicas e instrumentos que se aplican como guías de evaluación [23-26]; desde el desarrollo de software web y móvil, e introduciendo buenas prácticas que aseguren su eficiencia [27].
- la informática educativa, análisis de los recursos empleados para transformarlos en inclusivos empleando herramientas específicas de diseño y validación.
- la responsabilidad social [30, 31] y otros abordajes transversales que emergen de estudios relevados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Con la finalidad de relacionar aspectos prácticos de AW sustentados en temas teóricos, y producir conocimientos para apoyar la toma de decisiones, los resultados obtenidos y aquellos esperados son:

- Sistematizar las producciones en I+D de abordajes de AW por los grupos de investigación en Argentina. Estudios preliminares se expusieron en [32].
- Sistematizar las experiencias del equipo en torno a la AW -iniciadas en 2012- con la finalidad de diseñar escenarios de validación, y difundirlas desde un enfoque regional con perspectiva internacional.
- Profundizar el conocimiento en estándares, métodos, herramientas e instrumentos para evaluación de la AW con la finalidad de construir y mantener sitios accesibles, considerado herramientas automáticas y semi-automáticas [33, 34], definición de instrumentos de evaluación manual, determinación de un método para la evaluación de la AW, especificación de métricas, otros.
- Capitalizar las evaluaciones de AW, con la finalidad de establecer buenas prácticas.
- Transformar los recursos producidos por el equipo y utilizados en capacitaciones y formaciones en materiales accesibles, según los estándares estudiados.
- Adaptar un método de la ISBE [35] [36] para la evaluación y medición de AW en sitios web construidos por terceros y por este equipo, consolidar los datos en una base de datos y aplicar algoritmos para producir conocimiento.
- Capacitar y alfabetizar en temas de AW, y socializar las producciones concretadas.
- Fortalecer actividades de transferencia al medio, acuerdos de trabajo y otros documentos que formalicen las actividades del equipo en pro de entornos más accesibles a la información.
- Promover y participar en redes de cooperación que tratan esta temática de connotación social, económica y cultural.

4. FORMACIÓN DE RECURSOS HUMANOS

Desde la FaCENA, y proyectos acreditados por la SGCyT-UNNE, se continuará fortaleciendo la formación de docentes-investigadores, becarios y tesistas de grado y posgrado en temas de AW. Estos actores se insertan en diversos ámbitos científicos-tecnológicos, sociales y profesionales, y se constituyen en los naturales difusores de estos temas de connotación inclusiva, social e internacional. Como se expresó en [36], las estrategias docentes se sustentan en metodologías activas para el logro de aprendizajes significativos y la aplicación de buenas prácticas para la generación o identificación de recursos accesibles.

Los estudiantes de grado o posgrado profundizan aspectos teóricos, metodológicos, técnicos y legales en torno a la AW siendo el contexto de validación los casos abstraídos de la realidad. Se están elaborando dos tesis de Maestría en Ingeniería del Software [38-39] y en becas de grado y trabajos finales de grado en Sistemas -con la dirección de otros integrantes del equipo, se aplican conceptos, métodos y herramientas en torno de AW. Por ello, se considera un ámbito propicio para la iniciación, continuación o finalización de estudios de grado, posgrado e I+D que se retroalimentan en actividades de docencia y transferencia.

REFERENCIAS

- [1] S. I. Mariño, V. R Bercheñi, Identificación de brechas digitales en pandemia: dos experiencias de grados superiores en la disciplina Informática, *Revista Mendeive. Revista de Educación*, 18(4), 910-922, 2020.
- [2] C. M. Carvajal, Web accessibility evaluation of Chilean universities, *Formacion Universitaria*, 13(5), 69 – 76, 2020.
- [3] J. Díaz, I. Harari y P. Amadeo, “Accesibilidad web Una mirada integral”. Facultad de Informática-UNLP, Edulp, 2020.
- [4] Web Content Accessibility Guidelines (WCAG) 2.0. Disponible en: <http://www.w3.org/TR/WCAG20/>
- [5] Web Content Accessibility Guidelines (WCAG) 2.1. Disponible en: <https://www.w3.org/TR/WCAG21/>
- [6] Web Content Accessibility Guidelines (WCAG) 2.2. Disponible en: <https://www.w3.org/TR/WCAG22/>
- [7] W3C. About W3C. Disponible en: <https://www.w3.org/Consortium/>
- [8] AENOR. Asociación Española de Normalización y Certificación. Disponible en: <https://www.aenor.com/>
- [9] ISO/IEC 40500:2012. Information technology - W3C Web Content Accessibility Guidelines (WCAG) 2.0. Disponible en: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=58625.
- [10] ISO/IEC 9241-171:2008. Ergonomics of human-system interaction – Part 171: Guidance on software accessibility.
- [11] ISO/IEC 9241-20:2008. Ergonomics of human-system interaction – Part 20: Accessibility guidelines or information/communication technology (ICT) equipment and services.
- [12] ISO/IEC 9241-151:2008 Ergonomics of human-system interaction – Part 151: Guidance on World Wide Web user interfaces.
- [13] Fundación Sidar, Fundación Sidar - Acceso Universal-. Disponible en: <http://www.sidar.org/>
- [14] CIDAT. Tecno Accesible, Centro de Investigación, Desarrollo y Aplicación Tiflotécnica. Disponible en: <http://cidat.once.es/>
- [15] Universidad de Alicante. Legislación de la Unión Europea. Disponible en: <http://accesibilidadweb.dlsi.ua.es/?menu=union-europea>.
- [16] Universidad de Alicante, Legislación Estados Unidos. Disponible en: <http://accesibilidadweb.dlsi.ua.es/?menu=estados-unidos>.
- [17] Universidad de Alicante, Legislación Española. Disponible en: <http://accesibilidadweb.dlsi.ua.es/?menu=espanola>.
- [18] Congreso de la Nación Argentina, Ley 26.653, Noviembre 2010. Disponible en: <https://www.argentina.gob.ar/jefatura/innovacion-publica/onti/accesibilidad-web>.
- [19] Jefatura de Gabinete de Ministros Oficina Nacional de Tecnologías de Información, Disposición 6/2019, Septiembre 2019. Disponible en: <https://www.boletinoficial.gob.ar/detalleAviso/primera/217660/20190930>.
- [20] V. K. Pagnoni y S. I. Mariño, Accesibilidad Web centrada en discapacidades visuales. Estudio empírico longitudinal de un portal de formación docente, *XXVIII Congreso Argentino de Ciencias de la Computación*, CACIC 2022.

- [21] P. Bourque y R. E. Fairley. SWEBOK Guide to the Software Engineering Body of Knowledge, Version 3.0. IEEE Computer Society. 2014, Disponible en: <https://www.computer.org/education/bodies-of-knowledge/software-engineering>
- [22] S. I. Mariño y P. L. Alfonzo, Automatic Web accessibility evaluation process. Case study, *Espirales Revista Multidisciplinaria de investigación*, 6(42), 2022.
- [23] S. I. Mariño y P. L. Alfonzo, Evidencias de Accesibilidad Web en la generación de sitios. Propuesta de un método, *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, no. 23, 52-60, 2019.
- [24] V. K. Pagnoni y S. I. Mariño, Una guía de Accesibilidad Web para portales educativos. La revisión de usuarios, *XXVII Congreso Argentino de Ciencias de la Computación*, CACIC 2021.
- [25] SEDIC. Accesibilidad Web: Técnicas y herramientas para evaluar la accesibilidad web. Disponible en: <http://www.sedic.es/autoformacion/accesibilidad/9-tecnicas-herramientas.html>.
- [26] S. I. Mariño, M. V. Godoy, P. Alfonzo, M. E. Sánchez, S. N. Ivaniszyn, V. Pagnoni, R. Alderete, M. Fernández, G. de Los Reyes y J. M. Bordón, Diseño de métodos, procedimientos y herramientas. Aportes al desarrollo regional, *Workshop de Investigadores en Ciencias de la Computación*, WICC 2020.
- [27] S. I. Mariño y V. K. Pagnoni, Accesibilidad Web en Dispositivos móviles. Evaluación de un portal educativo de alcance nacional, *International Journal of Educational Research and Innovation (IJERI)*, 13, 208-225, 2020.
- [28] J. J. Mayol, F. Perales, F. Negre-Bennasar y G. F. Nadal, El diseño web y material didáctico accesible en la enseñanza universitaria. *Revista de educación a distancia (RED)*, 19(60), 2019.
- [29] C. E. P. Jiménez, L.R.A. Rodríguez y L. E. G. Peña, Material audiovisual accesible a personas con discapacidad sensorial de Lima Metropolitana, 2021. *Revista Científica de Comunicación Social*, 4, 101-112, 2022.
- [30] S. I. Mariño, P. L. Alfonzo y M. V. Godoy. La Responsabilidad Social centrada en la accesibilidad web. *European Scientific Journal*, 16(30), 35-50, 2020.
- [31] S. I. Mariño, P. L. Alfonzo y M. V. Godoy. Capítulo: Accesibilidad web. Un aporte de responsabilidad social universitaria, Libro: A Ciencia da Computacao e o desenvolvimento de conteúdo tecnológico relevante para a sociedade. Editorial: Atena, 2020. DOI 10.22533/at.ed.683202003
- [32] S. I. Mariño y P. L. Alfonzo, Estudios de Accesibilidad Web en congresos nacionales. Evidencias en tiempos de pandemia Sars-Cov-2, Simposio Argentino de Educación Superior, *51 Jornadas Argentina de Informática e Investigación Operativa*, 2022.
- [33] V. K. Pagnoni y S. I. Mariño, Validación de contenido de un portal educativo centrado en la Accesibilidad Web, *Revista Iberoamericana de Informática Educativa. IE Comunicaciones*, 29, 14-22, 2019.
- [34] V. K. Pagnoni y S. I. Mariño, Calidad de contenidos en dominios de educación. Evaluación de la accesibilidad Web Mediada por validadores automáticos, *EDMETIC, Revista de Educación Mediática y TIC*, 8(1), 107-127, 2019.
- [35] M. Genero, "Ingeniería del software basada en la evidencia," Escuela Superior de Informática, Universidad de Castilla-La Mancha, 2016. Material en diapositiva
- [36] S. I. Mariño, P. L. Alfonzo, V. K. Pagnoni, M. L. Gronda, J. M. Bordón, 2022, Formación en Accesibilidad Web. *Jornadas Argentinas de Didáctica de las Ciencias de la Computación (JADICC2022)*
- [37] V. K. Pagnoni, "Marco de trabajo de evaluación de la accesibilidad web para

entornos virtuales de enseñanza aprendizaje”, Tesis de Ingeniería del Software, Dir. Mariño, S. I., Rossi, G. Trabajo de Maestría en Ingeniería del Software, UNLP, 2018, (en realización).

- [38] M. L. Gronda, Tesis de Ingeniería del Software, Dir. Mariño, S. I., Rossi, G. Trabajo de Maestría en Ingeniería del Software, UNLP, (en realización).

ESTIMACIÓN DEL TAMAÑO DE UN SISTEMA MEDIANTE EL ANÁLISIS DE DIAGRAMAS UML Y LA APLICACIÓN DE MÉTODOS BASADOS EN PUNTOS DE FUNCIÓN

Nicolás Hernández, Julio Castillo, Marina Cardenas, Gonzalo Albarracín, José Victoria
Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/Facultad
Regional Córdoba/Universidad Tecnológica Nacional
{damiannicolas05, dr.jotacastillo, ing.marinacardenas, galbarracin16, josefvictoria52}@gmail.com

RESUMEN

Este proyecto tiene por objetivo lograr estimaciones precisas del tamaño de un sistema mediante el uso de métodos de estimación que utilicen puntos de función, en las etapas tempranas de un proyecto. En la actualidad existen tres métodos que han ganado mucha relevancia entre los métodos de estimación por puntos de función, estos son COSMIC, IFPUG y NESMA. Para poder utilizar estos métodos se requiere información en las etapas iniciales de un proyecto, y por esta razón se analizarán los diagramas UML, y, en particular, los archivos creados por los diferentes programas que soportan estos diagramas y que brindan archivos con diferentes tipos de estándares tales como los archivos XMI, XML, JSON y SVG.

Palabras Clave Automatización, Mediciones, Estimaciones, COSMIC, IFPUG y NESMA, UML, XMI, XML, JSON, SVG, Puntos de función.

CONTEXTO

El objetivo de este trabajo es el de desarrollar una herramienta que permita realizar estimaciones precisas con la información disponible en las etapas iniciales del desarrollo de un proyecto de software. Dicho trabajo, se encuentra enmarcado dentro del proyecto de investigación homologado por la Secretaría de Ciencia y Tecnología (SCyT) de la UTN “Modelo para el procesamiento de textos estructurados Fase 2” (cód. SIECACO0008518), el cual está dentro del Grupo de Aprendizaje Automático, Lenguajes y Autómatas (GA²LA [1]) conformado en la Universidad Tecnológica Nacional - Facultad Regional Córdoba (UTN- FRC), el cual tiene por objetivo investigar los temas relacionados

con las áreas de redes neuronales, gramáticas y sintaxis [2][3][4], y también el desarrollo de sistemas, su calidad y trazabilidad. En etapas tempranas se cuenta con información especificada principalmente en diagramas de UML (Lenguaje de Modelado Unificado), los cuales son generados con las herramientas disponibles en el mercado y que, por lo general, proveen como archivos de salidas en alguno de los formatos estándares mencionados anteriormente, o bien en algún formato propietario. Para poder obtener y analizar esta información se deben realizar análisis léxicos-sintácticos de estos archivos a los efectos de poder realizar el cálculo de los puntos de función.

El proyecto físicamente se desarrolla en el Laboratorio de Investigación de Software LIS del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC), el cual está compuesto por doctores, doctorandos, ingenieros, docentes - investigadores, pasantes, becarios alumnos y tesis de posgrado.

1. INTRODUCCIÓN

La problemática que afecta a los proyectos de sistemas es el no poder contar con información suficiente para saber cuántos recursos asignar para lograr entregar dentro de los plazos acordados, con las funciones solicitadas y dentro del presupuesto establecido. Según la IFPUG (International Function Point User Group) el 45% de los proyectos exceden su presupuesto, 7% entregan fuera de cronograma y 56% entregan menos requerimientos [5]. Es por esto que las empresas deben poder predecir estas tres

variables con la mayor precisión posible con el fin de disminuir los riesgos y las pérdidas, para lograr el objetivo de entregar un sistema con la mayor calidad posible que cumpla con las especificaciones y requerimientos. Para poder realizar esto, es preciso recurrir a una tarea compleja que permite predecir tiempo, costo y esfuerzo: la estimación [6]. Los principales métodos de estimación pueden dividirse en dos grandes grupos, el primero se basa en experiencias adquiridas en proyectos anteriores y en el conocimiento que se tiene del dominio del negocio, y el segundo grupo incluye los métodos basados en algoritmos de costos. Este último grupo, como se mencionó anteriormente, es el que más relevancia ha obtenido con los métodos basados en puntos de función, los cuales tienen la ventaja de ser independiente tanto de los expertos como de los lenguajes de programación. Entre los más populares se pueden nombrar NESMA FPA [7], IFPUG-FPA [5] y COSMIC [8]. Dado que estos métodos requieren de mucho tiempo y esfuerzo para su realización, se ve la necesidad de automatizar el proceso de estimación. Como existe una gran variedad de software que gráfica los diagramas de UML, se planea desarrollar una herramienta que acepte diferentes tipos de archivos y que los procese automáticamente.

Para esto se seleccionaron, según diferentes características, 25 programas. Los cuales fueron estudiados para determinar si cumplían con las siguientes características:

- Soportar UML total o parcialmente.
- Ser gratuitos, de licencia libre (GNU, GPL, MIT, BSD, EPL, shareware), o contar mínimo con una versión de prueba o con suscripción con características reducidas.
- Los tipos de archivos debían ser XML, XML, JSON, SVG y/o extensiones propias de cada software que fueran similares a los antes mencionados.

Los programas seleccionados fueron:

- BOULM.
- DIA.

- Modelio.
- NClass.
- ModelSphere.
- UMPLE.
- visual paradigm.
- WhiteStarUML.
- app.diagrams.net.
- UML Designer.
- ALTOVA UModel.
- Moqup.
- Creately.
- Visio.
- StarUML.
- JetUML.
- IBM Engineering Systems Design Rhapsody.
- Enterprise Architect.
- Diagramo.
- GenMyModel.

En la actualidad se ha conformado un banco de prueba con datos de proyectos, diagramas y duraciones. Cada diagrama ha sido generado con cada programa para estudiar y analizar sus salidas, como la salida XMI que se muestra en la siguiente figura.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <XMI:PROJECT xmlns:XMI="http://www.staruml.com" version="1">
3 <XMI:HEADER>
4 <XMI:SUBUNITS>
5 </XMI:SUBUNITS>
6 <XMI:PROFILES>
7 <XMI:PROFILE>UMLStandard</XMI:PROFILE>
8 </XMI:PROFILES>
9 </XMI:HEADER>
10 <XMI:BODY>
11 <XMI:OBJ name="DocumentElement" type="UMLProject" guid="QvokAFQZUScd3TuVtQ5wAA">
12 <XMI:ATTR name="Title" type="string">ipavc</XMI:ATTR>
13 <XMI:ATTR name="OwnedElements" type="integer">5</XMI:ATTR>
14 <XMI:OBJ name="OwnedElements[0]" type="UMLModel" guid="Wqy0dJtEk2725UCgUthGA">
15 <XMI:ATTR name="Name" type="string">Use Case View</XMI:ATTR>
16 <XMI:REF name="Namespace">QvokAFQZUScd3TuVtQ5wAA</XMI:REF>
17 <XMI:ATTR name="OwnedDiagrams" type="integer">2</XMI:ATTR>
18 <XMI:OBJ name="OwnedDiagrams[0]" type="UMLUseCaseDiagram" guid="2/eFf10fQqiyVY1ocB3mAA">
19 <XMI:ATTR name="Name" type="string">Diagrama de casos de uso esenciales</XMI:ATTR>
20 <XMI:REF name="DiagramOwner">Wqy0dJtEk2725UCgUthGA</XMI:REF>
21 <XMI:OBJ name="DiagramView" type="UMLUseCaseDiagramView" guid="EecF7wRkvkDnxzVf30Xt5AAA">
22 <XMI:REF name="Model">2/eFf10fQqiyVY1ocB3mAA</XMI:REF>
23 <XMI:REF name="Diagram">2/eFf10fQqiyVY1ocB3mAA</XMI:REF>
24 <XMI:ATTR name="OwnedViews" type="integer">32</XMI:ATTR>
25 <XMI:OBJ name="OwnedViews[0]" type="UMLActorView" guid="vz0M9uXj0qPjn2XW5wAAAA">
26 <XMI:ATTR name="LineColor" type="string">clMaroon</XMI:ATTR>
27 <XMI:ATTR name="FillColor" type="string">50009FFF</XMI:ATTR>
28 <XMI:ATTR name="Left" type="integer">58</XMI:ATTR>
29 <XMI:ATTR name="Top" type="integer">124</XMI:ATTR>
30 <XMI:ATTR name="Width" type="integer">152</XMI:ATTR>
31 <XMI:ATTR name="Height" type="integer">78</XMI:ATTR>
32 <XMI:REF name="Model">4Tn0kKDYkeXq1mba98RQAA</XMI:REF>
33 <XMI:OBJ name="NameCompartment" type="UMLNameCompartmentView" guid="KHG32ncdvEqhqt09ypR2gAA">
34 <XMI:OBJ name="NameLabel" type="LabelView" guid="0RagbU1UzE2kaYqP8fIwAA">
35 <XMI:ATTR name="Frontstyle" type="integer">1</XMI:ATTR>
36 <XMI:ATTR name="Text" type="string">Encargado de Liquidaci&243;n</XMI:ATTR>
37 </XMI:OBJ>
38 </XMI:OBJ>
39 </XMI:OBJ>
40 </XMI:OBJ>
41 </XMI:OBJ>
42 </XMI:OBJ>
43 </XMI:OBJ>
44 </XMI:OBJ>
45 </XMI:OBJ>
46 </XMI:OBJ>
47 </XMI:OBJ>
48 </XMI:OBJ>
49 </XMI:OBJ>
50 </XMI:OBJ>
51 </XMI:OBJ>
52 </XMI:OBJ>
53 </XMI:OBJ>
54 </XMI:OBJ>
55 </XMI:OBJ>
56 </XMI:OBJ>
57 </XMI:OBJ>
58 </XMI:OBJ>
59 </XMI:OBJ>
60 </XMI:OBJ>
61 </XMI:OBJ>
62 </XMI:OBJ>
63 </XMI:OBJ>
64 </XMI:OBJ>
65 </XMI:OBJ>
66 </XMI:OBJ>
67 </XMI:OBJ>
68 </XMI:OBJ>
69 </XMI:OBJ>
70 </XMI:OBJ>
71 </XMI:OBJ>
72 </XMI:OBJ>
73 </XMI:OBJ>
74 </XMI:OBJ>
75 </XMI:OBJ>
76 </XMI:OBJ>
77 </XMI:OBJ>
78 </XMI:OBJ>
79 </XMI:OBJ>
80 </XMI:OBJ>
81 </XMI:OBJ>
82 </XMI:OBJ>
83 </XMI:OBJ>
84 </XMI:OBJ>
85 </XMI:OBJ>
86 </XMI:OBJ>
87 </XMI:OBJ>
88 </XMI:OBJ>
89 </XMI:OBJ>
90 </XMI:OBJ>
91 </XMI:OBJ>
92 </XMI:OBJ>
93 </XMI:OBJ>
94 </XMI:OBJ>
95 </XMI:OBJ>
96 </XMI:OBJ>
97 </XMI:OBJ>
98 </XMI:OBJ>
99 </XMI:OBJ>
100 </XMI:OBJ>
101 </XMI:OBJ>
102 </XMI:OBJ>
103 </XMI:OBJ>
104 </XMI:OBJ>
105 </XMI:OBJ>
106 </XMI:OBJ>
107 </XMI:OBJ>
108 </XMI:OBJ>
109 </XMI:OBJ>
110 </XMI:OBJ>
111 </XMI:OBJ>
112 </XMI:OBJ>
113 </XMI:OBJ>
114 </XMI:OBJ>
115 </XMI:OBJ>
116 </XMI:OBJ>
117 </XMI:OBJ>
118 </XMI:OBJ>
119 </XMI:OBJ>
120 </XMI:OBJ>
121 </XMI:OBJ>
122 </XMI:OBJ>
123 </XMI:OBJ>
124 </XMI:OBJ>
125 </XMI:OBJ>
126 </XMI:OBJ>
127 </XMI:OBJ>
128 </XMI:OBJ>
129 </XMI:OBJ>
130 </XMI:OBJ>
131 </XMI:OBJ>
132 </XMI:OBJ>
133 </XMI:OBJ>
134 </XMI:OBJ>
135 </XMI:OBJ>
136 </XMI:OBJ>
137 </XMI:OBJ>
138 </XMI:OBJ>
139 </XMI:OBJ>
140 </XMI:OBJ>
141 </XMI:OBJ>
142 </XMI:OBJ>
143 </XMI:OBJ>
144 </XMI:OBJ>
145 </XMI:OBJ>
146 </XMI:OBJ>
147 </XMI:OBJ>
148 </XMI:OBJ>
149 </XMI:OBJ>
150 </XMI:OBJ>
151 </XMI:OBJ>
152 </XMI:OBJ>
153 </XMI:OBJ>
154 </XMI:OBJ>
155 </XMI:OBJ>
156 </XMI:OBJ>
157 </XMI:OBJ>
158 </XMI:OBJ>
159 </XMI:OBJ>
160 </XMI:OBJ>
161 </XMI:OBJ>
162 </XMI:OBJ>
163 </XMI:OBJ>
164 </XMI:OBJ>
165 </XMI:OBJ>
166 </XMI:OBJ>
167 </XMI:OBJ>
168 </XMI:OBJ>
169 </XMI:OBJ>
170 </XMI:OBJ>
171 </XMI:OBJ>
172 </XMI:OBJ>
173 </XMI:OBJ>
174 </XMI:OBJ>
175 </XMI:OBJ>
176 </XMI:OBJ>
177 </XMI:OBJ>
178 </XMI:OBJ>
179 </XMI:OBJ>
180 </XMI:OBJ>
181 </XMI:OBJ>
182 </XMI:OBJ>
183 </XMI:OBJ>
184 </XMI:OBJ>
185 </XMI:OBJ>
186 </XMI:OBJ>
187 </XMI:OBJ>
188 </XMI:OBJ>
189 </XMI:OBJ>
190 </XMI:OBJ>
191 </XMI:OBJ>
192 </XMI:OBJ>
193 </XMI:OBJ>
194 </XMI:OBJ>
195 </XMI:OBJ>
196 </XMI:OBJ>
197 </XMI:OBJ>
198 </XMI:OBJ>
199 </XMI:OBJ>
200 </XMI:OBJ>
201 </XMI:OBJ>
202 </XMI:OBJ>
203 </XMI:OBJ>
204 </XMI:OBJ>
205 </XMI:OBJ>
206 </XMI:OBJ>
207 </XMI:OBJ>
208 </XMI:OBJ>
209 </XMI:OBJ>
210 </XMI:OBJ>
211 </XMI:OBJ>
212 </XMI:OBJ>
213 </XMI:OBJ>
214 </XMI:OBJ>
215 </XMI:OBJ>
216 </XMI:OBJ>
217 </XMI:OBJ>
218 </XMI:OBJ>
219 </XMI:OBJ>
220 </XMI:OBJ>
221 </XMI:OBJ>
222 </XMI:OBJ>
223 </XMI:OBJ>
224 </XMI:OBJ>
225 </XMI:OBJ>
226 </XMI:OBJ>
227 </XMI:OBJ>
228 </XMI:OBJ>
229 </XMI:OBJ>
230 </XMI:OBJ>
231 </XMI:OBJ>
232 </XMI:OBJ>
233 </XMI:OBJ>
234 </XMI:OBJ>
235 </XMI:OBJ>
236 </XMI:OBJ>
237 </XMI:OBJ>
238 </XMI:OBJ>
239 </XMI:OBJ>
240 </XMI:OBJ>
241 </XMI:OBJ>
242 </XMI:OBJ>
243 </XMI:OBJ>
244 </XMI:OBJ>
245 </XMI:OBJ>
246 </XMI:OBJ>
247 </XMI:OBJ>
248 </XMI:OBJ>
249 </XMI:OBJ>
250 </XMI:OBJ>
251 </XMI:OBJ>
252 </XMI:OBJ>
253 </XMI:OBJ>
254 </XMI:OBJ>
255 </XMI:OBJ>
256 </XMI:OBJ>
257 </XMI:OBJ>
258 </XMI:OBJ>
259 </XMI:OBJ>
260 </XMI:OBJ>
261 </XMI:OBJ>
262 </XMI:OBJ>
263 </XMI:OBJ>
264 </XMI:OBJ>
265 </XMI:OBJ>
266 </XMI:OBJ>
267 </XMI:OBJ>
268 </XMI:OBJ>
269 </XMI:OBJ>
270 </XMI:OBJ>
271 </XMI:OBJ>
272 </XMI:OBJ>
273 </XMI:OBJ>
274 </XMI:OBJ>
275 </XMI:OBJ>
276 </XMI:OBJ>
277 </XMI:OBJ>
278 </XMI:OBJ>
279 </XMI:OBJ>
280 </XMI:OBJ>
281 </XMI:OBJ>
282 </XMI:OBJ>
283 </XMI:OBJ>
284 </XMI:OBJ>
285 </XMI:OBJ>
286 </XMI:OBJ>
287 </XMI:OBJ>
288 </XMI:OBJ>
289 </XMI:OBJ>
290 </XMI:OBJ>
291 </XMI:OBJ>
292 </XMI:OBJ>
293 </XMI:OBJ>
294 </XMI:OBJ>
295 </XMI:OBJ>
296 </XMI:OBJ>
297 </XMI:OBJ>
298 </XMI:OBJ>
299 </XMI:OBJ>
300 </XMI:OBJ>
301 </XMI:OBJ>
302 </XMI:OBJ>
303 </XMI:OBJ>
304 </XMI:OBJ>
305 </XMI:OBJ>
306 </XMI:OBJ>
307 </XMI:OBJ>
308 </XMI:OBJ>
309 </XMI:OBJ>
310 </XMI:OBJ>
311 </XMI:OBJ>
312 </XMI:OBJ>
313 </XMI:OBJ>
314 </XMI:OBJ>
315 </XMI:OBJ>
316 </XMI:OBJ>
317 </XMI:OBJ>
318 </XMI:OBJ>
319 </XMI:OBJ>
320 </XMI:OBJ>
321 </XMI:OBJ>
322 </XMI:OBJ>
323 </XMI:OBJ>
324 </XMI:OBJ>
325 </XMI:OBJ>
326 </XMI:OBJ>
327 </XMI:OBJ>
328 </XMI:OBJ>
329 </XMI:OBJ>
330 </XMI:OBJ>
331 </XMI:OBJ>
332 </XMI:OBJ>
333 </XMI:OBJ>
334 </XMI:OBJ>
335 </XMI:OBJ>
336 </XMI:OBJ>
337 </XMI:OBJ>
338 </XMI:OBJ>
339 </XMI:OBJ>
340 </XMI:OBJ>
341 </XMI:OBJ>
342 </XMI:OBJ>
343 </XMI:OBJ>
344 </XMI:OBJ>
345 </XMI:OBJ>
346 </XMI:OBJ>
347 </XMI:OBJ>
348 </XMI:OBJ>
349 </XMI:OBJ>
350 </XMI:OBJ>
351 </XMI:OBJ>
352 </XMI:OBJ>
353 </XMI:OBJ>
354 </XMI:OBJ>
355 </XMI:OBJ>
356 </XMI:OBJ>
357 </XMI:OBJ>
358 </XMI:OBJ>
359 </XMI:OBJ>
360 </XMI:OBJ>
361 </XMI:OBJ>
362 </XMI:OBJ>
363 </XMI:OBJ>
364 </XMI:OBJ>
365 </XMI:OBJ>
366 </XMI:OBJ>
367 </XMI:OBJ>
368 </XMI:OBJ>
369 </XMI:OBJ>
370 </XMI:OBJ>
371 </XMI:OBJ>
372 </XMI:OBJ>
373 </XMI:OBJ>
374 </XMI:OBJ>
375 </XMI:OBJ>
376 </XMI:OBJ>
377 </XMI:OBJ>
378 </XMI:OBJ>
379 </XMI:OBJ>
380 </XMI:OBJ>
381 </XMI:OBJ>
382 </XMI:OBJ>
383 </XMI:OBJ>
384 </XMI:OBJ>
385 </XMI:OBJ>
386 </XMI:OBJ>
387 </XMI:OBJ>
388 </XMI:OBJ>
389 </XMI:OBJ>
390 </XMI:OBJ>
391 </XMI:OBJ>
392 </XMI:OBJ>
393 </XMI:OBJ>
394 </XMI:OBJ>
395 </XMI:OBJ>
396 </XMI:OBJ>
397 </XMI:OBJ>
398 </XMI:OBJ>
399 </XMI:OBJ>
400 </XMI:OBJ>
401 </XMI:OBJ>
402 </XMI:OBJ>
403 </XMI:OBJ>
404 </XMI:OBJ>
405 </XMI:OBJ>
406 </XMI:OBJ>
407 </XMI:OBJ>
408 </XMI:OBJ>
409 </XMI:OBJ>
410 </XMI:OBJ>
411 </XMI:OBJ>
412 </XMI:OBJ>
413 </XMI:OBJ>
414 </XMI:OBJ>
415 </XMI:OBJ>
416 </XMI:OBJ>
417 </XMI:OBJ>
418 </XMI:OBJ>
419 </XMI:OBJ>
420 </XMI:OBJ>
421 </XMI:OBJ>
422 </XMI:OBJ>
423 </XMI:OBJ>
424 </XMI:OBJ>
425 </XMI:OBJ>
426 </XMI:OBJ>
427 </XMI:OBJ>
428 </XMI:OBJ>
429 </XMI:OBJ>
430 </XMI:OBJ>
431 </XMI:OBJ>
432 </XMI:OBJ>
433 </XMI:OBJ>
434 </XMI:OBJ>
435 </XMI:OBJ>
436 </XMI:OBJ>
437 </XMI:OBJ>
438 </XMI:OBJ>
439 </XMI:OBJ>
440 </XMI:OBJ>
441 </XMI:OBJ>
442 </XMI:OBJ>
443 </XMI:OBJ>
444 </XMI:OBJ>
445 </XMI:OBJ>
446 </XMI:OBJ>
447 </XMI:OBJ>
448 </XMI:OBJ>
449 </XMI:OBJ>
450 </XMI:OBJ>
451 </XMI:OBJ>
452 </XMI:OBJ>
453 </XMI:OBJ>
454 </XMI:OBJ>
455 </XMI:OBJ>
456 </XMI:OBJ>
457 </XMI:OBJ>
458 </XMI:OBJ>
459 </XMI:OBJ>
460 </XMI:OBJ>
461 </XMI:OBJ>
462 </XMI:OBJ>
463 </XMI:OBJ>
464 </XMI:OBJ>
465 </XMI:OBJ>
466 </XMI:OBJ>
467 </XMI:OBJ>
468 </XMI:OBJ>
469 </XMI:OBJ>
470 </XMI:OBJ>
471 </XMI:OBJ>
472 </XMI:OBJ>
473 </XMI:OBJ>
474 </XMI:OBJ>
475 </XMI:OBJ>
476 </XMI:OBJ>
477 </XMI:OBJ>
478 </XMI:OBJ>
479 </XMI:OBJ>
480 </XMI:OBJ>
481 </XMI:OBJ>
482 </XMI:OBJ>
483 </XMI:OBJ>
484 </XMI:OBJ>
485 </XMI:OBJ>
486 </XMI:OBJ>
487 </XMI:OBJ>
488 </XMI:OBJ>
489 </XMI:OBJ>
490 </XMI:OBJ>
491 </XMI:OBJ>
492 </XMI:OBJ>
493 </XMI:OBJ>
494 </XMI:OBJ>
495 </XMI:OBJ>
496 </XMI:OBJ>
497 </XMI:OBJ>
498 </XMI:OBJ>
499 </XMI:OBJ>
500 </XMI:OBJ>
501 </XMI:OBJ>
502 </XMI:OBJ>
503 </XMI:OBJ>
504 </XMI:OBJ>
505 </XMI:OBJ>
506 </XMI:OBJ>
507 </XMI:OBJ>
508 </XMI:OBJ>
509 </XMI:OBJ>
510 </XMI:OBJ>
511 </XMI:OBJ>
512 </XMI:OBJ>
513 </XMI:OBJ>
514 </XMI:OBJ>
515 </XMI:OBJ>
516 </XMI:OBJ>
517 </XMI:OBJ>
518 </XMI:OBJ>
519 </XMI:OBJ>
520 </XMI:OBJ>
521 </XMI:OBJ>
522 </XMI:OBJ>
523 </XMI:OBJ>
524 </XMI:OBJ>
525 </XMI:OBJ>
526 </XMI:OBJ>
527 </XMI:OBJ>
528 </XMI:OBJ>
529 </XMI:OBJ>
530 </XMI:OBJ>
531 </XMI:OBJ>
532 </XMI:OBJ>
533 </XMI:OBJ>
534 </XMI:OBJ>
535 </XMI:OBJ>
536 </XMI:OBJ>
537 </XMI:OBJ>
538 </XMI:OBJ>
539 </XMI:OBJ>
540 </XMI:OBJ>
541 </XMI:OBJ>
542 </XMI:OBJ>
543 </XMI:OBJ>
544 </XMI:OBJ>
545 </XMI:OBJ>
546 </XMI:OBJ>
547 </XMI:OBJ>
548 </XMI:OBJ>
549 </XMI:OBJ>
550 </XMI:OBJ>
551 </XMI:OBJ>
552 </XMI:OBJ>
553 </XMI:OBJ>
554 </XMI:OBJ>
555 </XMI:OBJ>
556 </XMI:OBJ>
557 </XMI:OBJ>
558 </XMI:OBJ>
559 </XMI:OBJ>
560 </XMI:OBJ>
561 </XMI:OBJ>
562 </XMI:OBJ>
563 </XMI:OBJ>
564 </XMI:OBJ>
565 </XMI:OBJ>
566 </XMI:OBJ>
567 </XMI:OBJ>
568 </XMI:OBJ>
569 </XMI:OBJ>
570 </XMI:OBJ>
571 </XMI:OBJ>
572 </XMI:OBJ>
573 </XMI:OBJ>
574 </XMI:OBJ>
575 </XMI:OBJ>
576 </XMI:OBJ>
577 </XMI:OBJ>
578 </XMI:OBJ>
579 </XMI:OBJ>
580 </XMI:OBJ>
581 </XMI:OBJ>
582 </XMI:OBJ>
583 </XMI:OBJ>
584 </XMI:OBJ>
585 </XMI:OBJ>
586 </XMI:OBJ>
587 </XMI:OBJ>
588 </XMI:OBJ>
589 </XMI:OBJ>
590 </XMI:OBJ>
591 </XMI:OBJ>
592 </XMI:OBJ>
593 </XMI:OBJ>
594 </XMI:OBJ>
595 </XMI:OBJ>
596 </XMI:OBJ>
597 </XMI:OBJ>
598 </XMI:OBJ>
599 </XMI:OBJ>
600 </XMI:OBJ>
601 </XMI:OBJ>
602 </XMI:OBJ>
603 </XMI:OBJ>
604 </XMI:OBJ>
605 </XMI:OBJ>
606 </XMI:OBJ>
607 </XMI:OBJ>
608 </XMI:OBJ>
609 </XMI:OBJ>
610 </XMI:OBJ>
611 </XMI:OBJ>
612 </XMI:OBJ>
613 </XMI:OBJ>
614 </XMI:OBJ>
615 </XMI:OBJ>
616 </XMI:OBJ>
617 </XMI:OBJ>
618 </XMI:OBJ>
619 </XMI:OBJ>
620 </XMI:OBJ>
621 </XMI:OBJ>
622 </XMI:OBJ>
623 </XMI:OBJ>
624 </XMI:OBJ>
625 </XMI:OBJ>
626 </XMI:OBJ>
627 </XMI:OBJ>
628 </XMI:OBJ>
629 </XMI:OBJ>
630 </XMI:OBJ>
631 </XMI:OBJ>
632 </XMI:OBJ>
633 </XMI:OBJ>
634 </XMI:OBJ>
635 </XMI:OBJ>
636 </XMI:OBJ>
637 </XMI:OBJ>
638 </XMI:OBJ>
639 </XMI:OBJ>
640 </XMI:OBJ>
641 </XMI:OBJ>
642 </XMI:OBJ>
643 </XMI:OBJ>
644 </XMI:OBJ>
645 </XMI:OBJ>
646 </XMI:OBJ>
647 </XMI:OBJ>
648 </XMI:OBJ>
649 </XMI:OBJ>
650 </XMI:OBJ>
651 </XMI:OBJ>
652 </XMI:OBJ>
653 </XMI:OBJ>
654 </XMI:OBJ>
655 </XMI:OBJ>
656 </XMI:OBJ>
657 </XMI:OBJ>
658 </XMI:OBJ>
659 </XMI:OBJ>
660 </XMI:OBJ>
661 </XMI:OBJ>
662 </XMI:OBJ>
663 </XMI:OBJ>
664 </XMI:OBJ>
665 </XMI:OBJ>
666 </XMI:OBJ>
667 </XMI:OBJ>
668 </XMI:OBJ>
669 </XMI:OBJ>
670 </XMI:OBJ>
671 </XMI:OBJ>
672 </XMI:OBJ>
673 </XMI:OBJ>
674 </XMI:OBJ>
675 </XMI:OBJ>
676 </XMI:OBJ>
677 </XMI:OBJ>
678 </XMI:OBJ>
679 </XMI:OBJ>
680 </XMI:OBJ>
681 </XMI:OBJ>
682 </XMI:OBJ>
683 </XMI:OBJ>
684 </XMI:OBJ>
685 </XMI:OBJ>
686 </XMI:OBJ>
687 </XMI:OBJ>
688 </XMI:OBJ>
689 </XMI:OBJ>
690 </XMI:OBJ>
691 </XMI:OBJ>
692 </XMI:OBJ>
693 </XMI:OBJ>
694 </XMI:OBJ>
695 </XMI:OBJ>
696 </XMI:OBJ>
697 </XMI:OBJ>
698 </XMI:OBJ>
699 </XMI:OBJ>
700 </XMI:OBJ>
701 </XMI:OBJ>
702 </XMI:OBJ>
703 </XMI:OBJ>
704 </XMI:OBJ>
705 </XMI:OBJ>
706 </XMI:OBJ>
707 </XMI:OBJ>
708 </XMI:OBJ>
709 </XMI:OBJ>
710 </XMI:OBJ>
711 </XMI:OBJ>
712 </XMI:OBJ>
713 </XMI:OBJ>
714 </XMI:OBJ>
715 </XMI:OBJ>
716 </XMI:OBJ>
717 </XMI:OBJ>
718 </XMI:OBJ>
719 </XMI:OBJ>
720 </XMI:OBJ>
721 </XMI:OBJ>
722 </XMI:OBJ>
723 </XMI:OBJ>
724 </XMI:OBJ>
725 </XMI:OBJ>
726 </XMI:OBJ>
727 </XMI:OBJ>
728 </XMI:OBJ>
729 </XMI:OBJ>
730 </XMI:OBJ>
731 </XMI:OBJ>
732 </XMI:OBJ>
733 </XMI:OBJ>
734 </XMI:OBJ>
735 </XMI:OBJ>
736 </XMI:OBJ>
737 </XMI:OBJ>
738 </XMI:OBJ>
739 </XMI:OBJ>
740 </XMI:OBJ>
741 </XMI:OBJ>
742 </XMI:OBJ>
743 </XMI:OBJ>
744 </XMI:OBJ>
745 </XMI:OBJ>
746 </XMI:OBJ>
747 </XMI:OBJ>
748 </XMI:OBJ>
749 </XMI:OBJ>
750 </XMI:OBJ>
751 </XMI:OBJ>
752 </XMI:OBJ>
753 </XMI:OBJ>
754 </XMI:OBJ>
755 </XMI:OBJ>
756 </XMI:OBJ>
757 </XMI:OBJ>
758 </XMI:OBJ>
759 </XMI:OBJ>
760 </XMI:OBJ>
761 </XMI:OBJ>
762 </XMI:OBJ>
763 </XMI:OBJ>
764 </XMI:OBJ>
765 </XMI:OBJ>
766 </XMI:OBJ>
767 </XMI:OBJ>
768 </XMI:OBJ>
769 </XMI:OBJ>
770 </XMI:OBJ>
771 </XMI:OBJ>
772 </XMI:OBJ>
773 </XMI:OBJ>
774 </XMI:OBJ>
775 </XMI:OBJ>
776 </XMI:OBJ>
777 </XMI:OBJ>
778 </XMI:OBJ>
779 </XMI:OBJ>
780 </XMI:OBJ>
781 </XMI:OBJ>
782 </XMI:OBJ>
783 </XMI:OBJ>
784 </XMI:OBJ>
785 </XMI:OBJ>
786 </XMI:OBJ>
787 </XMI:OBJ>
788 </XMI:OBJ>
789 </XMI:OBJ>
790 </XMI:OBJ>
791 </XMI:OBJ>
792 </XMI:OBJ>
793 </XMI:OBJ>
794 </XMI:OBJ>
795 </XMI:OBJ>
796 </XMI:OBJ>
797 </XMI:OBJ>
798 </XMI:OBJ>
799 </XMI:OBJ>
800 </XMI:OBJ>
801 </XMI:OBJ>
802 </XMI:OBJ>
803 </XMI:OBJ>
804 </XMI:OBJ>
805 </XMI:OBJ>
806 </XMI:OBJ>
807 </XMI:OBJ>
808 </XMI:OBJ>
809 </XMI:OBJ>
810 </XMI:OBJ>
811 </XMI:OBJ>
812 </XMI:OBJ>
813 </XMI:OBJ>
814 </XMI:OBJ>
815 </XMI:OBJ>
816 </XMI:OBJ>
817 </XMI:OBJ>
818 </XMI:OBJ>
819 </XMI:OBJ>
820 </XMI:OBJ>
821 </XMI:OBJ>
822 </XMI:OBJ>
823 </XMI:OBJ>
824 </XMI:OBJ>
825 </XMI:OBJ>
826 </XMI:OBJ>
827 </XMI:OBJ>
828 </XMI:OBJ>
829 </XMI:OBJ>
830 </XMI:OBJ>
831 </XMI:OBJ>
832 </XMI:OBJ>
833 </XMI:OBJ>
834 </XMI:OBJ>
8
```

datos XMI, XML, JSON, SVG, que luego permitirá que sean utilizados en el cómputo de los puntos de función al momento de realizar las estimaciones de costo y duración de proyectos software.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto tiene por objetivo investigar temas en el área de las estimaciones basadas en puntos de función, y las metodologías utilizadas para contar dichos puntos, además de investigar sobre formatos de transferencia de datos basados en texto estructurado y su análisis sintáctico. En ese sentido, la línea de investigación asociada a este trabajo se relaciona principalmente con el área de Estimaciones de Software en Ingeniería de Software.

A su vez, este trabajo se encuentra dentro del proyecto de “Modelo para el procesamiento de textos estructurados Fase 2” (cód. SIECACO0008518) el cual está homologado por la Secretaría de Ciencia y Tecnología de la UTN y se enmarca principalmente en el área de la Lingüística Computacional e Ingeniería de Software.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como resultados obtenidos, podemos mencionar que se logró confeccionar un banco de prueba con un gran volumen de información y una amplia variedad de gráficos UML de muestra, todos en archivos con los formatos de salida anteriormente mencionados. También se ha avanzado con el desarrollo del módulo de análisis de archivos y la interpretación de los formatos de datos.

Se espera que se logre completar el software con un módulo que, en base a los datos obtenidos del módulo de análisis, permita realizar las estimaciones que se irán mejorando hasta lograr el mayor grado de precisión.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de investigación dentro del cual se encuentra el trabajo descrito en este artículo participan docentes investigadores ingenieros y alumnos del departamento de Ingeniería en Sistemas de Información de la UTN FRC.

Dicho proyecto sirve como marco para la realización de la tesis de Maestría en Ingeniería en Sistemas de Información del Ing. Nicolás Hernández, cuyo director es el Dr. Julio Castillo y su co-directora, la Mg. Marina E. Cardenas. Además ha permitido que estudiantes becarios realicen diferentes tareas permitiendo que incorporen experiencia desde el punto de vista científico.

5. BIBLIOGRAFÍA

- [1] Vázquez J., Castillo J., Constable L. y Cardenas M.. (2018). XX Workshop de Investigadores en Ciencias de la Computación, WICC 2018, Universidad Nacional del Nordeste.
- [2] Castillo J. y Cardenas M. (2010) Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. Iberamia 2010, LNCS, vol. 6433, pp. 366-375.
- [3] Castillo J., Cardenas M., Curti A., Casco O., Navarro M., Hernández N. y Velazco M. (2017) Desarrollo de sistemas de análisis de texto. XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017).
- [4] Cardenas M., Castillo J., Navarro M., Hernández N. y Velazco M. (2019) Herramientas para el desarrollo de sistemas de análisis de textos no estructurados. XXI Edición de Workshop de Investigadores en Ciencias de la Computación. WICC 2019, UN de San Juan, Argentina, 25 y 26 de abril de 2019.

[5] Bundschuh M. y Dekkers C. (2008) The IFPUG Function Point Counting Method. 10.1007/978-3-540-68188-5_12.

[6] PMI (2013) Project Management Institute, PMBOK-Project Management Body Of Knowledge, Quinta ed., Newton Square, Pennsylvania: PMI, 2013.

[7] NESMA. (2019) Function Point Analysis for software enhancement. Professional guide of the Nesma. <https://nesma.org/> Recuperado Febrero 2023, de <https://nesma.org/wp-content/uploads/2020/10/FPA-for-Software-Enhancement-v2.3-EN.pdf>

[8] COSMIC. (2020). Manual de medición de COSMIC. www.cosmic-sizing.org. Recuperado Febrero 2023, de <https://cosmic-sizing.org/wp-content/uploads/2020/05/Part-1-MM-Rules-v5.0-b-EN-ESPA%C3%91OL-1.pdf>.

[9] XML.sax (2022) Soporte para analizadores SAX2 — documentación de Python - 3.10.6. doc.phyton.org. Recuperado 11 de agosto de 2022, de <https://docs.python.org/es/3/library/xml.sax.html>.

Gobernanza de Sistemas y Tecnologías de la Información (SI/TI) en el ámbito Universitario

Hugo Ramón¹, Nicolás Alonso², Sebastián Mir³, María Celeste Nobal⁴, Gustavo Iglesias⁵,
Franco Bernaldo de Quirós⁷, Paola Santinelli⁷

{hramon}@docentes.unsada.edu.ar

{nicolas.alonso, sebastian.mir, celeste.nobal}@unsada.edu.ar

{giglesias, bfranco, psantinelli}@alumnos.unsada.edu.ar

Universidad Nacional de San Antonio de Areco (UNSAdeA), Buenos Aires, Argentina
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA), Instituto de
Investigación y Transferencia (ITT-CIC)

RESUMEN

La evolución de los Sistemas de Información (SI) y Tecnologías de la Información (TI) genera impactos, transformaciones y modificaciones en el comportamiento de los diversos ámbitos de nuestra vida cotidiana. El flujo ininterrumpido de innovaciones que suceden en el campo de los SI/TI nos generan nuevas formas de relacionarnos, estudiar y trabajar.

Partiendo desde una base que las TI/SI son los motores que dan forma a nuestra vida social, económica, política, cultural y ambiental, las actuales exigencias que atraviesan los profesionales del área aumentan progresivamente, los mismos deben de estar preparados para lograr que la adquisición de la tecnología sea una inversión y no un gasto y además esté alineado con los objetivos del negocio de las organizaciones.

Leyes como las de *Gordon Bell* y la de *Downes y Mui* que enuncian, respectivamente que: “*la plataforma dominante en tecnologías de la información cambia alrededor de cada diez años*” y “*los sistemas sociales, políticos y económicos cambian incrementalmente, pero la tecnología cambia exponencialmente*” hace que sea imprescindible indagar en los campos de gobernabilidad en SI/TI de forma que la inminente velocidad de innovación y transformación no traiga aparejado consecuencias no deseadas en las organizaciones, tales como proyectos que no culminan o que se exceden en el tiempo más allá de lo planificado, costos ocultos, alcances funcionales desacompañados entre las etapas de inicio y finalización del proyecto, entre otras. En estos términos, hablar de gobernanza de SI/TI implica hablar de liderazgo, de estructuras, de procesos organizacionales para que, a través de la tecnología, se puedan lograr los objetivos de una organización.

¹ Doctorando, Magister, Lic. en Informática, Investigador ITT-UNNOBA-CIC, Profesor UNNOBA, UNSADA

² Maestrando, Lic. en Informática, Profesor UNLP, UNNOBA, UNSADA

³ Ing. en Informática, Jefe del Departamento IT, Profesor UNSADA

⁴ Lic. en Trabajo Social, Personal Administrativo UNSADA

⁵ Estudiante avanzado/a de la carrera de Lic. en Informática, UNSADA

Este proyecto propone evaluar los esfuerzos y el impacto de las actividades de *e-government* en un ambiente educativo universitario, definiendo las características y factores críticos de su implementación, para luego poder realizar las recomendaciones que surjan de la evaluación de acuerdo a los lineamientos definidos en el Plan Nacional de Gobierno Electrónico (Decreto 378/2005 del PEN). En el caso de la Universidad Nacional de San Antonio de Areco (UNSAaA) se tomarán las normativas existentes y en caso de ser necesario se realizarán las propuestas necesarias que la complementen de manera de definir un proceso que permita la implementación del denominado "Gobierno Electrónico" en el ámbito de la UNSAaA.

Palabras clave: Gobierno Electrónico, Blockchain, Inteligencia Artificial, Ciencias de Datos.

CONTEXTO

Las líneas de investigación a describir se enmarcan en el proyecto de investigación: *Gobierno digital. Evaluación del uso de la tecnología*, con lugar de trabajo en la Universidad Nacional de San Antonio de Areco (UNSAaA) presentado en la convocatoria de Subsidios a la Investigación del año 2022 ante la Secretaría de Investigación de la Universidad y aprobados mediante Resolución (CS) N° 302/2022. Su objetivo es evaluar las actividades de *e-government* planificadas y desarrolladas en el ámbito de la UNSAaA, y el impacto de las mismas en dicha institución.

1. INTRODUCCIÓN

Al *e-government*, se lo define desde dos perspectivas, una lo define como el uso de las TIC por agencias gubernamentales y otra más

amplia lo define como un catalizador para inducir reformas administrativas y de política en el gobierno. El objetivo de muchos países es tener un *e-government* con mayor presencia y cada vez más eficiente. Esto implica no solamente el uso de las TIC, sino una nueva interacción con los actores y el gobierno e involucra aspectos políticos y culturales, cambios tecnológicos y organizacionales, diseñados para direccionar un cambio profundo en las unidades de gobierno. Esta forma de interacción utiliza la tecnología con el fin de aumentar aspectos como la transparencia, eliminar las distancias y divisiones físicas y capacitar a los ciudadanos para participar en los procesos políticos, aplicando en diferentes dominios.

El *e-government* no significa automatizar viejas prácticas. El uso de computadoras en procesos complejos, puede lograr mejorar la calidad de los servicios entregados por el gobierno y/o promover la participación ciudadana, pero centrarse únicamente en las TIC no va a cambiar la mentalidad de los empleados, que es un aspecto central del éxito de la implementación de cualquier proyecto.

Se define una tendencia tecnológica estratégica como aquella con potencial innovador sustancial que empieza a salir de un estado de emergente hacia un impacto de más amplio uso; o aquellas tecnologías que tienen rápido crecimiento, pero alta volatilidad y pueden alcanzar su punto de inflexión en los próximos años. La empresa experta en consultoría e investigación de las tecnologías de la información, *Gartner Inc.*, enunció que las tecnologías estrategias de los últimos años se pueden resumir en:

1. **Cosas Autónomas:** Interacción más natural con el ambiente.

2. **Analítica Aumentada:** Utilización de algoritmos automatizados para explorar nuevas hipótesis desde los datos.
3. **AI Driven Development:** estudio de herramientas, tecnologías y mejores prácticas para integrar la IA en las aplicaciones y en las mismas herramientas para el proceso de desarrollo.
4. **Gemelos Digitales.** Un gemelo digital es un modelo de software dinámico de un objeto físico o un sensor de datos que entiende el estado, responde a los cambios, mejora las operaciones y agrega valor.
5. **Edge Computing.** El procesamiento de la información y la recopilación y entrega de contenidos se sitúan más cerca de las fuentes de la información.
6. **Tecnologías de Inmersión.** Cambiando la forma en que los usuarios interactúan con el mundo, tecnologías como la realidad aumentada (RA), la realidad mixta (RM) y la realidad virtual (RV) aplican a esto.
7. **Blockchain.** Blockchain es un tipo de libro distribuido, una lista ordenada cronológicamente y en expansión de registros de transacciones firmados criptográficamente, irrevocables y compartidos por todos los participantes en una red.
8. **Smart Spaces.** Un espacio inteligente es un entorno físico o digital, en el que los seres humanos y los sistemas tecnológicos interactúan en ecosistemas, cada vez más abiertos, conectados, coordinados e inteligentes.
9. **Ética Digital y Privacidad.** Los consumidores son cada vez más conscientes del valor de su información personal y están cada vez más preocupados por la forma en que está siendo utilizada por entidades públicas y privadas.
10. **Quantum Computing.** La computación cuántica es un tipo de computación no clásica que se basa en el estado cuántico de las partículas subatómicas que representan información como elementos denotados como bits cuánticos.

Estas tecnologías y estratégicas impactan en cómo se desarrollan aplicaciones, buscando lograr la integración que las mismas deben implementar para descubrir las necesidades de los usuarios, con el fin de presentar la información pertinente en el lugar correcto y en el momento adecuado, con las características de calidad requeridas. La oportunidad de *I+D+i* se da en adaptar/diseñar aplicaciones o a la forma de construir las mismas, las cuales gestionan información procedente de redes de sensores compuestas por diversos nodos distribuidos y localizadas en espacios cerrados (hospitales, fábricas, oficinas, etc.) o abiertos (campos, bosque, etc.). Ambos espacios se encuentran en las diferentes actividades económicas de la región.

Además, la tecnología está incorporada en casi cada aspecto de nuestra vida, aunque en general pueda pasar desapercibida, está en nuestros teléfonos, en nuestros autos, en nuestros dispositivos de entretenimiento, en nuestros electrodomésticos y en casi cualquier objeto de uso cotidiano, pequeñas computadoras capaces de procesar y almacenar información, las cuales no se perciben como objetos diferenciados, son ubicuos.

Dado estos nuevos sistemas de *software*, y teniendo en cuenta el contexto anteriormente definido, la calidad y la innovación juegan un papel muy importante en el desarrollo de productos y servicios. Esto implica que hoy en día incluir al usuario en los procesos de innovación, y ubicarlo en una posición centralizada, está dando lugar a la aparición de nuevos espacios de interacción y comunicación, en los que el actor que consume productos y servicios, deja de considerarse en forma pasiva para convertirse en un actor activo central en los procesos de *I+D+i*. En Argentina, el Decreto 378/2005 del Poder Ejecutivo Nacional (PEN), indica que el Estado Argentino es el mayor productor/consumidor de información del país y establece el *Plan Nacional de Gobierno Electrónico*, el cual define los lineamientos estratégicos para *e-government*.

Es importante mencionar que el *e-government* no significa automatizar viejas prácticas. El uso de computadoras en procesos complejos, puede lograr mejorar la calidad de los servicios entregados por el gobierno y/o promover la participación ciudadana, pero centrarse únicamente en las TIC, no va a cambiar la mentalidad de los empleados, que es un aspecto central del éxito de la implementación de cualquier proyecto.

Los autores Al-Hashmi, A., & Darem, A. B. (2008), identifican diferentes estrategias para la implementar *e-government*:

- Trazar un plan integral a largo plazo.
- Identificar sólo unas pocas áreas clave como el foco de los primeros proyectos.
- Comenzar con proyectos pequeños en fases.

Para la elaboración de planes e iniciativas, algunos autores dividen el proceso de aplicación en diversas fases. Estas fases no son

dependientes entre sí, ni consecutivas, por lo cual no necesitan completar una fase antes de iniciar otra. Ofrecen maneras de pensar acerca de los objetivos.

En Baum, C., & Di Maio, A. (2000), se describe un proceso de cuatro fases, que poseen diferentes objetivos y requerimientos relacionados con la capacitación en el uso de las TIC y costos asociados.

Es un objetivo del presente proyecto evaluar los esfuerzos y el impacto de las actividades de *e-government* en un ambiente educativo universitario. A los fines indicados, se aplicarán las denominadas Tecnologías de la Información y de la Comunicación (TIC) a la organización y a los procedimientos internos de la administración en red. En este contexto, toma relevancia el concepto de *Gobierno Inteligente*, el cual refiere a la incorporación masiva de las TIC para mejorar los procesos internos gubernamentales que impactan en la interacción con las personas, organizaciones y la entrega de servicios. Para ello se debe de desarrollar, mantener y promover los sistemas integrados basados en Internet para la prestación de servicios y la provisión de información; y adoptar las medidas necesarias para que las comunicaciones se efectúen preferentemente mediante tecnologías informáticas, optimizando para ello la utilización de los recursos electrónicos disponibles en las distintas dependencias de la Universidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación de este trabajo consisten en el estudio de casos, de tipo descriptivo-correlacional, que mezcla técnicas cualitativas que van desde entrevistas, observación, revisión de documentos, observación directa de las instalaciones, de los

procesos y de los productos que aplican al objeto de estudio. Se aclara que no hay una intención de realizar una generalización estadística.

El estudio de casos es una de las diversas maneras de investigar que dependen de tres condiciones:

- El tipo de pregunta de investigación.
- El control que un investigador tiene sobre la conducta actual de los eventos.
- El foco sobre lo contemporáneo como opuesto al fenómeno histórico.

Como se indica en Yin, R. (1994), en general los estudios de casos son las estrategias preferidas cuando se realizan las preguntas “cómo” y “porqué” y cuando el foco está en un fenómeno contemporáneo dentro de un contexto de la vida real. En Muñiz M. (2010), se indica que los estudios de caso son importantes en la investigación cualitativa, pero al mismo tiempo se usan cada vez más en una serie de aplicaciones en el campo profesional por ejemplo en la evaluación de un sistema educativo, la organización de una empresa, la participación social en un grupo, etc. Todo esto desde la perspectiva de las personas específicas que forman parte de esos grupos, empresas o instituciones.

La metodología de investigación aplicada que se utilizará en este proyecto consta de las siguientes etapas:

- Analizar el estado actual y los fundamentos teóricos de los temas propuestos.
- Realizar seminarios de discusión internos y externos para identificar problemas tecnológicos. Es importante el intercambio de información con los grupos del país y del exterior que trabajan en proyectos similares.

- Recopilar y analizar la bibliografía existente, así como también, publicaciones realizadas por otros equipos de investigación, revistas y demás producciones digitales disponibles en la Web sobre la temática de estudio.
- Focalizar el objetivo de investigación descriptiva para conocer situaciones, costumbres y actitudes predominantes a través de la descripción exacta de las actividades, objetos, procesos y personas.
- Investigar las alternativas de solución buscando optimizar el empleo de recursos humanos y tecnológicos interactuando con el mundo real obteniendo una solución específica para el problema concreto. Esta etapa resulta particularmente importante en la formación de recursos humanos que se inician en *I+D+i* y es beneficioso el intercambio con otros grupos con proyectos de investigación similares.
- Integrar el trabajo al equipo de investigación en el marco del proyecto al que pertenece.
- Analizar hipótesis, experimentar y obtener resultados e indicadores.
- Evaluar los resultados, eventualmente publicarlos y realizar transferencia al medio.
- Publicar resultados obtenidos.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera que los aspectos investigados contribuyan a propiciar el fortalecimiento en la formación de recursos humanos, en su

rol de investigadores o partícipes activos en equipos de investigación; fomentando la culminación de sus estudios superiores, promoviendo la redacción, exposición y defensa de trabajos finales de grado y postgrado.

A su vez, se pretende difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes congresos, jornadas y workshops de carácter nacional e internacional relacionados con el uso de las tecnologías emergentes.

En lo que refiere a la evaluación y caracterización de las actividades de *e-government*, se espera que los aspectos investigados contribuyan a poder optar por un enfoque cuanti-cualitativo que identifique y describa un estado de situación acerca de las actividades *e-government* en el caso de UNSAdA, que permite luego hacer un mapeo de las decisiones de política institucional que inciden en el *e-government* o el valor público de este. Desde la perspectiva cuantitativa, se espera que, por medio del diseño de un instrumento de medición, cuyos objetivos alcancen cuestiones de infraestructura, prácticas e iniciativas, políticas de decisión institucional respecto al *e-government*, permita la identificación de tendencias acerca de las políticas y de los usos de los servicios entregados por el *e-government* en el contexto de un relevamiento no exhaustivo que contribuya un muestreo valioso para una primera sistematización.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por docentes e investigadores formados y en formación pertenecientes a la Universidad

Nacional de San Antonio de Areco (UNSAAdA), algunos de los cuales dirigieron becas del Consejo Interuniversitario Nacional: (CIN) y de la Comisión de Investigaciones Científicas (CIC), y diversos trabajos finales de la carrera de Analista en Informática. Entre los integrantes del equipo hay Doctores y Magísteres, y otros docentes investigadores se encuentran cursando Maestrías y Especializaciones. Se espera en el 2022 tener dos presentaciones de Becas CIN de estudiantes avanzados y una beca CIC.

5. BIBLIOGRAFÍA

Al-Hashmi, A., & Darem, A. B. (2008). Understanding phases of E-government project. *New Delhi. Disponible en:* http://www.csi-sigegov.org/emerging_pdf/17_152-157.pdf

Baum, C., & Di Maio, A. (2000). Gartner's four phases of e-government model. *Gartner Group, 12.*

Decreto 378/2005 (2005). Plan Nacional de Gobierno Electrónico y Planes Sectoriales de Gobierno Electrónico. Disponible en: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/105000-109999/105829/norma.htm#:~:text=El%20Pl an%20Nacional%20de%20Gobierno%20Elec tr%20C3%B3nico%20impulsar%20C3%A1%20el %20uso%20intensivo,los%20servicios%20p %20C3%BAblicos%20e%20incrementar>

Decreto 996/2018 (2018). Agenda Digital Argentina. Disponible en: <http://servicios.infoleg.gob.ar/infolegInternet/verNorma.do?id=316036>

European Commission (2016). EU eGovernment Action Plan 2016-2020. Accelerating the digital transformation of

government. Disponible en: <https://eur-lex.europa.eu/legal-content/en/all/?uri=celex:52016dc0179>

Gobierno Digital del Departamento de Estado de los EEUU (2012). Digital Government Strategy - United States Department of State. Disponible en: <https://www.state.gov/digital-government-strategy/>

Kraemer, K. L., & King, J. L. (1984). National Policies for Local Government Computing: an Assessment of Experience in Ten OECD Countries. *International Review of Administrative Sciences*, 50(2), 133–147. DOI: <https://doi.org/10.1177/002085238405000204>

Krieger, M. (2009). Los desafíos de cambio en la Administración Pública Argentina. *Schweinheim, Guillermo (compilador), Estado y administración pública: críticas, enfoques y prácticas en la Argentina actual, Buenos Aires, Asociación Argentina de Estudios en Administración Pública*, pág, 71-93.

Muñiz, M. (2010). Estudios de caso en la investigación cualitativa. *División de estudios de posgrado en la Universidad Autónoma de Nuevo León. Facultad de psicología. México*, 1-8.

Valdés, G., Solar, M., Astudillo, H., Iribarren, M., Concha, G., & Visconti, M. (2011). Conception, development and implementation of an e-government maturity model in public agencies. *Government Information Quarterly*, 28(2), 176-187.

Transformación Digital – Propuestas de solución para las Ciudades Inteligentes Sostenibles, la Gobernanza Digital y las Buenas prácticas y calidad.

Silvia Esponda , Ariel Pasini , Marcos Boracchia , Rocío Muñoz ,
Juan Ignacio Torres , Kristian Petkoff Bankoff , Julieta Calabrese , Cesar Estrebow ,
Patricia Pesado 

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 La Plata Buenos Aires

(sesponda, apasini, marcosb, rmunoz, jitorres, kpb, cesarest, ppesado) @lidi.info.unlp.edu.ar

RESUMEN

El III-LIDI (Instituto de Investigación en Informática LIDI) posee una línea de investigación orientada a la transformación digital y mejora de procesos.

La línea se divide en tres ejes principales

1. ciudades inteligentes sostenibles y resiliencia; 2. Gobernanza digital y gobierno abierto; y 3. Buenas prácticas y calidad. Para cada eje se proponen un conjunto de ideas para analizar durante este año, y además, se describen los resultados obtenidos hasta el momento.

Palabras Claves

Ingeniería de Software – Calidad de software - Gobernanza Digital – Ciudades inteligentes – Transformación digital - Resiliencia

CONTEXTO

Esta línea de investigación se enmarca en el proyecto “ Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” (11/F023) (2018-2022) y continua en el nuevo proyecto “Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital” (2023-2026), UNLP, Ministerio de Educación de la Nación.

Asimismo, la temática se encuentra abordada en proyectos aprobados por la Facultad de Informática UNLP (2022-2023).

El Instituto posee diversos acuerdos de cooperación con varias Universidades de Argentina y del exterior y con empresas privadas del sector, interesadas en mejorar sus procesos de desarrollo aplicando mejoras y participando en proyectos internacionales.

Desde esta línea el III-LIDI participa en proyectos internacionales que abordan los temas de Transformación Digital y Ciudades Digitales, en particular en los últimos años se ha trabajado en el proyecto CAP4CITY “Fortalecimiento de la capacidad de gobernanza de ciudades sostenibles inteligentes” (2018-2022), financiado por la Unión Europea dentro del programa ERASMUS+. Que ha sido el punta pie inicial para la creación de un postgrado interuniversitario entre la UNLP y la UNS: “Maestría en Gestión y Tecnología de Ciudades Inteligentes”. La carrera se encuentra en proceso de acreditación de CONEAU y se espera dar inicio al primer ciclo durante el 2023. Además, el Instituto posee una importante presencia en el “Consorcio de I+D+I en Cloud Computing-Big Data & Emergent Topics” (CC-BD&ET) con la participación

de grupos de investigación de Argentina, España, Chile, entre otros.

1. INTRODUCCION

La línea de investigación se focaliza en 3 ejes. El primero basado en el concepto de transformación digital, donde se avanza en la creación y desarrollo de soluciones relacionadas con las necesidades de las ciudades inteligentes sostenibles, la generación de herramientas para capacitar a la comunidad en el desarrollo de CIS y el efecto de la resiliencia en la sociedad. El segundo eje está directamente relacionado con el concepto de CIS. Se profundiza en el uso de la tecnología en la gobernanza digital, gobierno y datos abiertos. Por último, el tercer eje se enfoca en distintos lineamientos referidos a la generación de buenas prácticas, la mejora de procesos relacionados a los organismos públicos, y la calidad de sistemas de información.

Dichos ejes denominados: 1. ciudades inteligentes sostenibles y efectos resiliente; 2. Gobernanza digital y gobierno abierto; y 3. Buenas prácticas y calidad.

Eje 1 - Ciudades inteligentes sostenibles y efecto resiliente

Una CIS se define como: "...una ciudad innovadora que utiliza Tecnologías de Información y Comunicación (TIC) y otros medios para mejorar la calidad de vida, la eficiencia de las operaciones y los servicios urbanos; y la competitividad, al tiempo que se garantiza que satisfaga las necesidades de las generaciones presentes y futuras con respecto a los aspectos económicos, sociales, ambientales y culturales." [1][2][3][4], [5].

Las CIS se componen de cinco pilares: "Social", "Económico", "Ambiental", "Gobernanza" e "Infraestructura urbana".

Las ciudades proveen servicios públicos mediante el uso TICs a su comunidad, si estos servicios son orientados a la innovación y la sostenibilidad, se lo

considera un servicio CIS. Generalmente estos servicios se agrupan en los cinco pilares facilitando la identificación de los receptores que se verán beneficiados.

CIS y Resiliencia

La pandemia hizo visible el concepto ciudad digital resiliente. La transformación digital ha tomado fuerza como detonante de la resiliencia, en un proceso acelerado por la pandemia que apresuró cambios (que ya se evidenciaban) en las interacciones, la navegación de espacios y el modo de acceder a los servicios. Estos cambios se incrementarán, además, pues se espera que, en las décadas venideras, el proceso de urbanización sea prácticamente total. La población se desplaza hacia las ciudades en busca de oportunidades comerciales y económicas.

Frente a estos problemas, las ciudades se enfrentan al riesgo de padecer impactos y tensiones. A fin de lograr el menor trastorno posible y prevenir, minimizar y gestionar estos escenarios de forma efectiva, las ciudades deben hacerse más resilientes. La transformación digital es una herramienta clave de la resiliencia para prepararse para los desastres y las catástrofes y para mejorar el involucramiento ciudadano [6].

Un sistema resiliente debe tener cualidades tales como, ser Flexible, Redundante, Integrado, Inclusivo, Ingenioso, Robusto y Reflexivo.

Tener presentes estas cualidades permite observar el funcionamiento de las ciudades desde una perspectiva de resiliencia y ofrece la capacidad de diseñar estrategias más adecuadas para su crecimiento urbano

Eje 2- Gobernanza Digital y Gobierno Abierto

El concepto de gobernanza digital se basa en el uso de las TICs para mejorar la forma de gobierno, la entrega de servicios públicos y facilitar las interacciones entre el gobierno y el ciudadano [3], [5], [7], [8].

Los portales de gobierno, recientemente, se han utilizado como una herramienta para la difusión de información financiera con el fin de generar una imagen de transparencia gubernamental. Esta práctica usualmente se la confunde con el concepto de “gobierno abierto”. Pero la transparencia no es solo una característica de los gobiernos abiertos, sino que también es la consecuencia de políticas de gobierno abiertas que generan información a disposición de la ciudadanía. Existen diferentes definiciones de gobierno abierto. Sandoval lo define como: “El gobierno abierto debe entenderse como una plataforma tecnológica institucional que convierta los datos gubernamentales en datos abiertos para permitir su uso, protección y colaboración por parte de los ciudadanos en los procesos de decisión pública, rendición de cuentas y mejoramiento de los servicios públicos.” [8]–[10].

De la definición se desprende que el gobierno abierto no es solo “datos” en un sitio web para promocionar la transparencia del gobierno, sino que debe proveer tecnologías de información y comunicación (incluidas las redes sociales) para garantizar la protección de los datos y promover la participación por parte de la comunidad [9], [11].

El gobierno abierto se apoya en tres principios: Transparencia, Participación y Colaboración

Transparencia: Un gobierno transparente proporciona información sobre lo que está haciendo, sobre sus planes de actuación, sus fuentes de datos y sobre aquello de lo que puede responsabilizarse frente a la sociedad.

Participación: Un gobierno participativo promueve el derecho de la ciudadanía a colaborar activamente en la formulación de políticas públicas y facilita el camino para que las administraciones públicas se beneficien del conocimiento, de las ideas y de la experiencia de los ciudadanos.

Colaboración: Un gobierno colaborativo compromete e implica a los ciudadanos y

demás agentes sociales en el esfuerzo por trabajar conjuntamente para resolver los problemas nacionales.

La aplicación de estos conceptos se basa en dos ejes fundamentales: 1) Apertura de los datos: utilización de datos públicos (transparencia) para la innovación, generación de nuevas aplicaciones transformando a los gobiernos en gestores de servicios (permitiendo a la comunidad generar nuevos servicios que agregan valor). 2) Apertura de los procesos: facilitando la comunicación para la toma de decisiones participativas, incorporando el conocimiento y la experiencia de los ciudadanos en el diseño de políticas públicas (colaboración).

Siguiendo estos lineamientos, la sociedad posee exigencias cada vez más elevadas hacia sus gobernantes y sus entes de gobierno, orientando las políticas de gobierno hacia una gobernanza digital, con transparencia y manejo eficiente de los bienes públicos, entre otros. En estos puntos el uso de TICs es un requisito indispensable para incluir al ciudadano. En el mismo sentido se encuentra el concepto de datos abierto [8], [12], donde diferentes organismos coordinan acciones para mejorar la calidad de vida de la sociedad mediante la apertura de datos. En esta línea se generaron diversas aplicaciones y herramientas desde múltiples sectores para brindarles soporte y automatizar o mejorar el proceso de publicación, búsqueda y, en ocasiones, procesamiento de la información a las distintas partes de la sociedad. Entre estas herramientas se destacan los catálogos y portales de datos abiertos, mediante los cuales los organismos publican los datos sobre diferentes aspectos de su ejercicio y del medio ambiente en el cual operan.

Como parte de las áreas de aplicación de este proyecto se trabajará entonces en el diseño y desarrollo de metodologías y herramientas para la creación de sistemas de software aplicadas a estas temáticas.

Eje 3- Buenas Prácticas y Calidad

Una buena práctica es un conjunto de acciones que han sido efectivas en un determinado contexto, y se espera obtener el mismo resultado en contextos similares. Estas buenas prácticas se pueden encontrar tanto en tareas que se realizan de manera cotidiana como en la prestación de servicios. Tomando como premisa estas buenas prácticas, se puede mejorar la prestación de servicios de una organización. Es fundamental hacer un registro exhaustivo de dichas buenas prácticas para que puedan ser transmitidas entre las partes interesadas. Un proceso es un conjunto de actividades que dan un resultado esperado. La mejora de los procesos en la prestación de servicios aplicando las buenas prácticas influye de forma directa en su calidad.

Las líneas relacionadas con CIS y gobernanza digital, se basan en una buena gestión de los servicios de TICs y requieren de un análisis que permita evolucionar de acuerdo a las necesidades de la comunidad y mantenerse en el proceso de mejora continua. Existen distintas normas que coadyuvan al proceso de mejora continua abordando el tema de la calidad en el uso de TICs desde diferentes puntos de vista. En este contexto se cuenta con el aval de la familia de las normas ISO/IEC 25000:2005 - SQuaRE (Requisitos y evaluación de la calidad del producto de software); dentro de la familia, se destacan ISO/IEC 25010 - System and software quality models e ISO/IEC 25040 - Evaluation process. ISO/IEC 25023 - Measurement of system and software product quality, que define específicamente las métricas para realizar la medición de la calidad de productos y sistemas software.

Está demostrado que realizar un proceso de software bajo estándares de calidad desemboca en un producto de calidad. Para ello la familia de las normas ISO/IEC 33000 proporciona un marco de trabajo coherente para la evaluación de procesos software que sustituye las diferentes partes de la norma ISO/IEC 15504. Además, en el caso de las PyMEs se pueden aplicar los

lineamientos de la ISO/IEC 29110 (Systems and software engineering — Lifecycle profiles for Very Small Entities (VSEs)). En ambos casos podría aplicarse la ISO 9001 (Quality management systems — Requirements) bajo las directrices de la ISO/IEC 90003 (Software engineering -- Guidelines for the application of ISO 9001:2015).

Calidad de Datos y Calidad de Información: apunta a la gestión de la carga y almacenamiento de los datos que permite construir información confiable. En este sentido, la norma ISO/IEC 25012, Software engineering-Software product quality requirements and evaluation (SQuaRE) y la ISO/IEC 25024 - Measurement of data quality asisten para mantener los datos con alto nivel de calidad. SQuaRE agrupa las características de los datos bajo tres puntos de vista: 1) inherentes: se refiere al grado en que las características tienen el potencial de satisfacer las necesidades establecidas e implícitas (Exactitud, Completitud, Consistencia Credibilidad Actualidad); 2) Independientes del sistema: hace referencia al grado en que se alcanza y preserva la calidad de los datos en un sistema informático (Disponibilidad, Portabilidad Recuperabilidad) y 3) datos inherente y dependientes del sistema: las características que son pertinentes desde ambos puntos de vista: Accesibilidad, Cumplimiento, Confidencialidad, Eficiencia, Precisión, Trazabilidad y Comprensibilidad [13]–[15].

Calidad de servicio: analiza la calidad sobre la prestación de los servicios, en particular los relacionados con la atención al cliente y en este sentido se pueden aplicar las guías propuestas por la IS/IEC 20000 (Service management).

Calidad de gestión: este componente incluye a todas las partes que involucran la gestión de los servicios de TI. La prestación de los servicios generalmente está diseñada en un modelo basado en proceso, por lo que aplica de forma directa las mejoras que se obtienen al aplicar un

sistema de gestión de la calidad basado en la norma ISO 9001 Quality Management [15]–[16]

LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Análisis y estudio de normas y modelos orientados a la Calidad de Datos.
- Asistencia en el proceso de certificación en organismos públicos y privados según los requisitos de IRAM - ISO 9001.
- Desarrollo de Voto electrónico presencial y Voto por Internet. Arquitecturas adaptadas a la legislación vigente en distintos organismos.
- Análisis, discusión y estudio de herramientas para la aplicación de normas de calidad relacionadas con certificación de servicios gubernamentales.
- Análisis, discusión y estudio de técnicas y herramientas de mejoras de proceso en el desarrollo de software.
- Dentro del concepto de Transformación Digital se desarrollaron actividades relacionadas con el análisis de:
 - innovación y co-creación de servicios públicos
 - aplicaciones móviles para ciudades digitales
 - resiliencia en sistemas para la gobernanza digital

2. RESULTADOS OBTENIDOS/ ESPERADOS

- Se implementó el Sistema de Gestión de la Calidad de la Facultad (SIGCFI), con la aprobación del primer informe por el Consejo Directivo de la Facultad.
- Se mantiene el Portal de Gestión Administrativa para la Facultad de Informática UNLP, donde se reúne información y enlaces sobre los diferentes servicios que brinda la facultad.
- Se realizaron acciones de consultoría y asesoramiento en organismos públicos y privados, en particular se aplicó el Voto por Internet en distintos organismos.

- Se obtuvo la re-certificación bajo la norma IRAM-ISO 9001:2015 del proceso: “Concursos Docentes de la Facultad de Informática”.

- Se lanzó en la plataforma Edex el curso "Smart and Sustainable Cities: New Ways of Digitalization & Governance" que ha sido uno de los resultados del proyecto CAP4CITY

- Se realizaron eventos de diseminación en Argentina, Brasil Chile y Colombia de los resultados obtenidos del Proyecto CAP4CITY

- Se dictaron los cursos de postgrado que forman parte de la curricula del proyecto CAP4CITY

- Puesta en marcha de la Maestría en Gestión y Tecnología de Ciudades Inteligentes, en proceso de acreditación CONEAU.

3. FORMACION DE RECURSOS HUMANOS

- El proyecto cuenta con integrantes realizando la Maestría Ingeniería de Software de la Facultad de Informática de la UNLP.

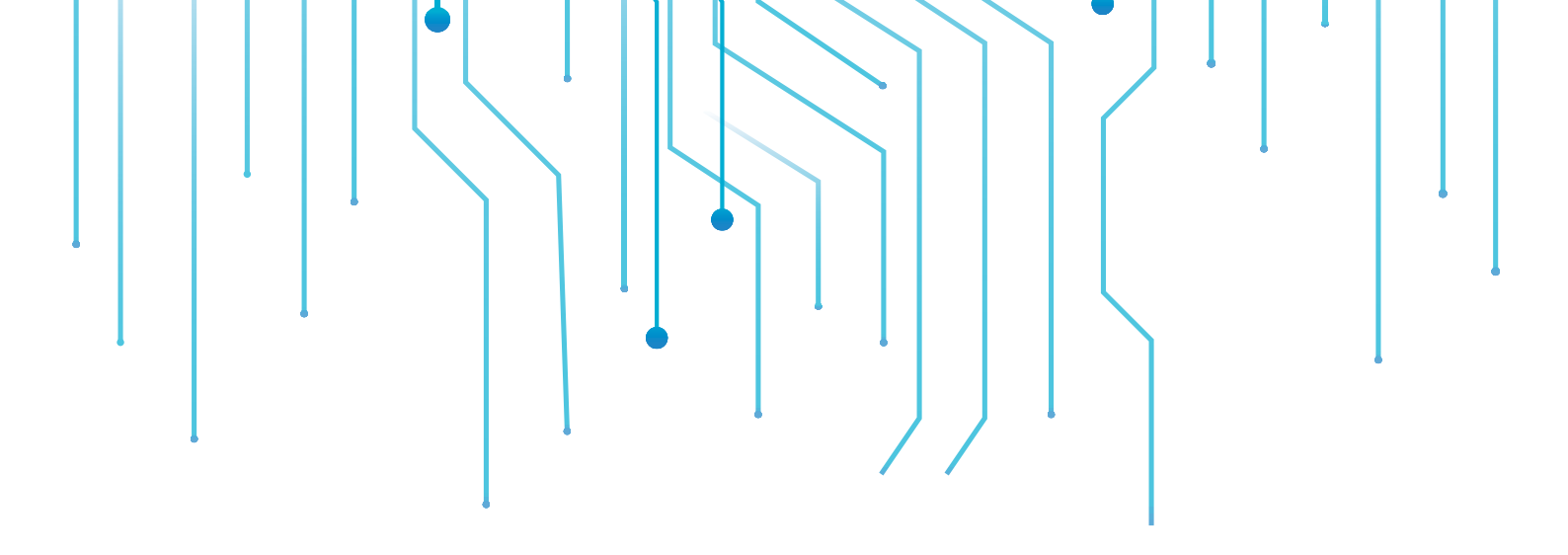
- Se dirigen tesis de postgrado, tesinas de grado y PPS en temas del área.

- Los integrantes de esta línea de Investigación participan en el dictado de asignaturas/cursos de grado/postgrado en la Facultad de Informática de la UNLP y en otras universidades del país. En particular, en la UNLP, se dicta la asignatura “Calidad de Sistemas de Software”

4. BIBLIOGRAFIA

- [1] E. Estevez, “Innovación en Servicios Públicos Digitales.”
- [2] ITU-T, “An overview of smart sustainable cities and the role of information and communication technologies,” 2014.
- [3] R. Muñoz, A. Pasini, and P. Pesado, “Innovación en el Sector Público para Ciudades Inteligentes Sostenibles.”
- [4] C. Sánchez, *Innovación pública*.

- [5] S. A. Chun, S. Shulman, R. Sandoval, and E. Hovy, "Government 2.0: Making connections between citizens, data and government," *Inf. Polity*, vol. 15, no. 1–2, pp. 1–9, 2010.
- [6] A. E. Millat, "Ciudades Resilientes Formando un Mundo Digital," 2020.
- [7] S. Esponda, A. Pasini, M. Boracchia, J. Calabrese, R. Muñoz, J. Santiago Preisegger, and P. Pesado, "Fortalecimiento de la calidad en procesos de software y Gobernanza Digital," *Sedici.Unlp.Edu.Ar*, pp. 574–578.
- [8] J. S. Preisegger, A. Greco, A. Pasini, M. Boracchia, and P. Pesado, "Marco de vinculación de datos abiertos aplicado al contexto de datos medioambientales," pp. 684–694.
- [9] A. Naser and A. Ramirez, "Plan de gobierno abierto. Una hoja de ruta para los Gobiernos de la Región," *CEPAL - Manuales*, vol. 81, p. 80, 2017.
- [10] A. Pasini, J. S. Preisegger, and P. Pesado, *Open Government Assessment Models Applied to Province's Capital Cities in Argentina and Municipalities in the Province of Buenos Aires*, vol. 995. 2019.
- [11] C. Calderón and S. Lorenzo, *Open Government. Gobierno Abierto*. 2010.
- [12] N. Noy, M. Burgess, and D. Brickley, "Google dataset search: Building a search engine for datasets in an open web ecosystem," in *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, 2019, pp. 1365–1375.
- [13] J. Calabrese, R. Muñoz, A. Pasini, S. Esponda, and M. Boracchia, "Asistente para la evaluación de características de calidad de producto de software propuestas por ISO / IEC 25010 basado en métricas definidas usando el enfoque GQM," pp. 660–671, 2017.
- [15] A. Pasini, E. Estévez, P. Pesado, and M. Boracchia, "Una metodología para evaluar la madurez de servicios universitarios," *Proc. Congr. XXII Congr. Argentino Ciencias la Comput. Congr. Argentino Ciencias la Comput.*, pp. 636–646, 2016.
- [16] A. Pasini, J. S. Preisegger, and P. Pesado, "Modelos de evaluación de gobiernos abiertos , aplicado a los municipios de la provincia de Buenos Aires," *XXIV Congr. Argentino Ciencias la Comput.*, vol. XXIV, pp. 0–10, 2018.



ISS

Innovación en Sistemas de Software

Estado actual de la industria 4.0 en las PyMES del partido de Hurlingham

Marisa Panizzi¹, Fernando Puricelli¹, Felipe Ortiz¹, Cristian Schiffino¹, Iris Sattolo¹, Antonio Bencardino¹.

¹ Instituto de Tecnología e Ingeniería. Universidad Nacional de Hurlingham.

Av. Vergara 2222 (B1688GEZ) - Villa Tesei - Bs. As. Argentina.

marisa.panizzi@unahur.edu.ar, fernando.puricelli@unahur.edu.ar, felipe.ortiz@unahur.edu.ar,
Cristian.schiffino@unahur.edu.ar, Iris.sattolo@unahur.edu.ar, antonio.bencardino@unahur.edu.ar

Resumen

Las soluciones tecnológicas 4.0 enfocadas en la interconectividad, la automatización y los datos en tiempo real contribuyen a incrementar la capacidad y calidad de los procesos productivos de las PyMES. La adopción de estas tecnologías por parte de este tipo de empresas es un gran desafío en cualquier sector industrial. En este artículo presentamos una síntesis del proceso investigativo para la realización de un estudio exploratorio del estado actual de la industria 4.0 en las PyMES del partido de Hurlingham mediante una encuesta. Además, se describen los primeros hallazgos en relación con el desarrollo de sistemas y a la Ciberseguridad de las PyMES que han participado en la encuesta hasta el momento.

Palabras clave: Desarrollo de sistemas, ciberseguridad, procesos industriales minería de datos, PyMES, partido de Hurlingham.

Contexto

La línea de investigación que se reporta en este artículo es financiada por la Secretaría de Investigación de la Universidad Nacional de Hurlingham (Resolución Consejo Superior 000382-21) en el marco de un proyecto de investigación titulado “*Estudio de la práctica actual de las tecnologías 4.0 en las PyMES del partido de Hurlingham mediante una encuesta*”

Introducción

De acuerdo con el Ministerio de Desarrollo Productivo en [1] la industria 4.0 se refiere a una nueva manera de producir mediante la adopción de tecnologías 4.0, es decir, de soluciones enfocadas en la interconectividad, la automatización y los datos en tiempo real.

Su primera mención formal con esta connotación data del año 2011, en la Feria de Hannover, Alemania, en la presentación del artículo Industria 4.0: Con el internet de las cosas camino de la 4° revolución industrial[2], donde se expuso cómo Alemania podría ser el próximo líder y proveedor del nuevo mercado en 2020 gracias al internet de las cosas en el entorno industrial.

En Argentina, los sectores industriales se componen mayoritariamente por Pequeñas y Medianas Empresas (PyMES), lo que constituye un eslabón fundamental para el país y esto refuerza la necesidad de llevar adelante iniciativas que contribuyan con el desarrollo y mejora de competitividad de dichas empresas. En la relación a la industria del software y servicios informáticos, las PyMES representan casi el 80% del sector [3]. Este sector presenta un alto potencial para generar valor agregado al ecosistema productivo del país, promoviendo la generación de empleo calificado y evidenciando un crecimiento exponencial en los últimos años. De hecho, el notable desarrollo queda demostrado en el último informe publicado por el Observatorio de la Economía del Conocimiento (OEC) de Argentina [4]. Esto constituye un eslabón fundamental, en el sector, para el país y refuerza la necesidad de llevar adelante iniciativas que contribuyan con el desarrollo y mejora de competitividad de dichas empresas. De acuerdo con Rainer Drath y Alexander Horch en [5], las hipótesis o fundamentos que deben ocurrir para que se den las condiciones para el desarrollo de la industria 4.0 son:

- La infraestructura de comunicación en los sistemas de producción será más asequible y por tanto será parte de todo.
- Los dispositivos en el campo, máquinas,

plantas y fábricas (incluso productos individuales) estarán más conectados a una red (la Internet o una red privada del fabricante).

- Los dispositivos en el campo, máquinas, plantas y fábricas (incluso productos individuales) serán capaces de almacenar documentos y conocimiento acerca de sí mismos fuera de su corporeidad en la red.

Para determinar la manera en que las PyMES del partido de Hurlingham pueden fortalecer sus procesos mediante la adopción de tecnologías 4.0 [6] se requiere recolectar evidencia sobre el estado actual de la aplicación de las tecnologías 4.0 en sus procesos productivos. Para el llevar a cabo el estudio exploratorio mediante una encuesta, decidimos utilizar las directrices para encuestas de Molléri *et al.* [7].

Dentro de la planificación de la encuesta, definimos el objetivo principal de la encuesta utilizando la plantilla GQM (*Goal-Question-Metric*) propuesta por Basili *et al.* [8] que se presenta en la Tabla 1.

Analizar	La adopción de tecnologías 4.0,
con el propósito de	conocer el estado actual,
con respecto a	descubrir las características de las áreas de conocimiento: el desarrollo de sistemas, la ciberseguridad, la ciencia de datos y la inteligencia artificial; y además el grado de consideración de la ley de género relacionada a la equidad de la distribución de los puestos de trabajo incluyendo los cargos de conducción por parte de estas empresas.
desde el punto de vista	de profesionales informáticos,
en el contexto de	de PyMES del partido de Hurlingham.

Tabla 1. Objetivo de la encuesta.

Las preguntas de investigación (PIs) que impulsaron este estudio son las siguientes:

PI1: ¿Qué características tienen las PyMES del partido de Hurlingham?

PI2: ¿Qué características tiene el área de desarrollo de sistemas?

PI3: ¿Cuáles son las características que tiene el área de ciberseguridad?

PI4: ¿Cuáles son las características que tiene el área de inteligencia artificial?

PI5: ¿Qué características tiene el área de ciencia de datos?

PI6: ¿Sería deseable la adopción de tecnologías 4.0?

Una vez definidas las preguntas de investigación, definimos el proceso de envío, se diseñó una planilla para el registro sistemático del seguimiento del envío y un texto de presentación que acompañe la encuesta. En este texto se menciona el propósito de la investigación, quiénes participan, tiempo estimado para responder la encuesta, un agradecimiento por la colaboración y una invitación a que el encuestado difunda la encuesta entre sus contactos.

La población a la cual se decidió enviar la encuesta son PyMES del partido de Hurlingham de cualquier sector industrial. Se utilizó la clasificación de PyMES de Argentina según la cantidad de empleados propuesta por la Secretaría de Emprendedores y PyMES del Ministerio de Desarrollo Productivo [6].

Se diseñó un cuestionario autoadministrado que se envió por correo electrónico a los encuestados y se utilizó la herramienta *Google Forms*.

Para el diseño del cuestionario se definieron cinco categorías de preguntas que junto con sus objetivos se presentan en la Tabla 2.

Para algunas preguntas demográficas de la encuesta para la categoría “Empresa” se utilizaron preguntas del estudio HELENA [9] por tratarse de un estudio internacionalmente reconocido.

Actualmente hemos diseñado el cuestionario para las categorías “Empresa”, “Desarrollo de sistemas” y “Ciberseguridad” que se encuentra en el siguiente enlace: <https://forms.gle/qJbUMsSWM27E77eT8>

Categorías	Objetivos
Empresa.	Determinar las características de las empresas que participan en la encuesta, su sector industrial, la cantidad de empleados y su ubicación geográfica e identificar el perfil del participante de la encuesta, su rol y su antigüedad en el rol. Además, conocer el grado de consideración de la ley de género relacionada a la equidad de la distribución de los puestos de trabajo incluyendo los cargos de conducción vinculados a la tecnología.
Área de desarrollo de sistemas.	Conocer las características de la práctica actual del desarrollo de sistemas en las PyMES de Hurlingham.
Área de ciberseguridad.	Conocer las características de la práctica actual de ciberseguridad en las PyMES de Hurlingham.
Área de inteligencia artificial.	Conocer las características de la práctica actual de la inteligencia artificial en las PyMES de Hurlingham.
Área de ciencia de datos.	Conocer las características de la práctica actual de ciencia de datos en las PyMES de Hurlingham.

Tabla 2. Categorías del cuestionario de la encuesta y sus objetivos.

Para el envío de la encuesta se trabaja de manera colaborativa con el Centro PyME-UNAHUR [10] y con la base de contactos de empresas provisto por el director de la carrera Ingeniería en Metalurgia de UNAHUR.

A la actualidad hemos recolectado las respuestas de 6 PyMES. En el gráfico 1 se presentan las PyMES de acuerdo con su sector industrial. En el gráfico 2 se presentan las respuestas en relación con el desarrollo de sistemas. Respecto al grado de interés de las empresas por recibir capacitación o asesoramiento en temas vinculados al desarrollo de sistemas, 4 PyMES se encuentran interesadas y los temas más demandados son diseño y arquitectura de software.

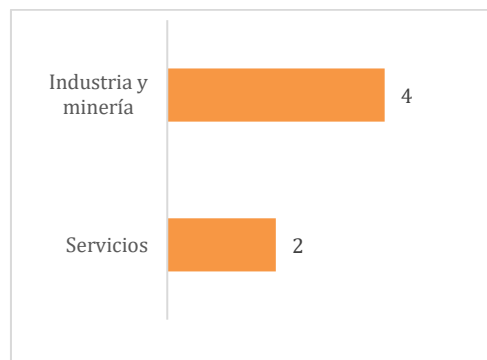


Gráfico 1. PyMES según sector industrial.

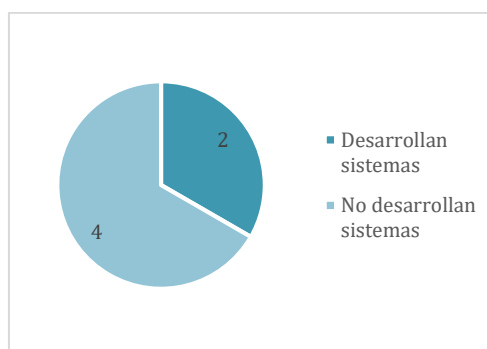


Gráfico 2. Desarrollo de sistemas.

En relación con la categoría “Ciberseguridad”, en la Tabla 3 se presenta si la empresa dispone de una estrategia de seguridad documentada o no. Las 6 PyMES que participaron no disponen de métricas, todas realizan copias de seguridad de la información, aunque solo en 3 de ellas los sistemas cuentan con software de protección contra malware. En el gráfico 3 se presenta el grado de interés por recibir capacitación o asesoramiento en esta temática.

Dispone de estrategia de seguridad	No dispone de estrategia de seguridad
1	5

Tabla 3. Estrategia de seguridad.

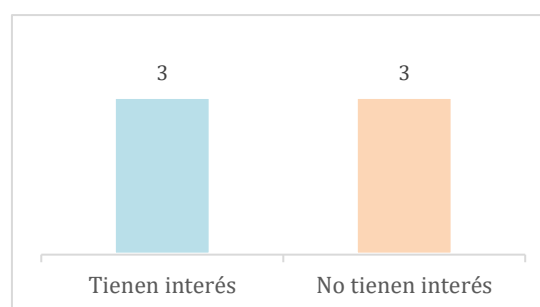


Gráfico 3. Grado de intereses de las PyMES por temas de Ciberseguridad.

Es importante aclarar que actualmente continuamos con la recolección de los datos y en paralelo realizamos el análisis de estos de manera manual. Además, nos encontramos diseñando las preguntas que corresponden a las categorías “Inteligencia artificial” y “Ciencia de datos” de la encuesta. Una vez lograda una muestra significativa realizaremos el análisis de los datos obtenidos mediante la aplicación de minería de datos.

Líneas de Investigación, Desarrollo e Innovación

El objetivo general de esta línea de investigación consiste en determinar el estado actual de la adopción de las tecnologías 4.0 por parte de las PyMES del partido de Hurlingham en sus procesos industriales. Y mediante la aplicación de procesos de minería de datos lograr obtener patrones de comportamiento de estas empresas con el propósito de que el Centro PyME-UNAHUR les facilite herramientas de capacitación y asesoramiento.

Resultados y Objetivos

Se lograron definir las preguntas que corresponden a las categorías “Empresa”, “Desarrollo de sistemas” y “Ciberseguridad” que componen la encuesta. Se comenzó con el envío y análisis manual de los datos obtenidos de estas categorías.

A continuación, se describen un conjunto de objetivos que se pretenden lograr:

- a) Académicos, dos tesinas de grado de la carrera Licenciatura en Informática.
- b) Producción Científica: se presentarán avances de la investigación en eventos científicos de alcance nacional (CACIC ¹ 2023) y en el ámbito internacional, CIACA 2023² e InGENIO³ 2023
- c) Formación en investigación: el grupo de investigación inicia un proceso de aprendizaje de métodos de investigación de ingeniería de software experimental como, por ejemplo, encuestas [7][11].

Para el desarrollo de este proyecto de investigación, se seguirá un enfoque de investigación clásico [12],[13]. Los métodos y materiales necesarios para desarrollo de la primera etapa del proyecto son los siguientes:

- **Métodos.**

Las encuestas son investigaciones que proporcionan una visión general, mediante la recolección de información estandarizada de una población específica o una muestra representativa de la misma (sujetos del estudio), por medio de un cuestionario o entrevista [11]. En la mayoría de los casos, los datos relativos a la encuesta provendrán de cuestionarios. Pero los cuestionarios, por sí solos no constituyen la encuesta. Para la construcción del cuestionario, se emplearán las directrices propuestas en [7]. Las encuestas son métodos de investigación primarios al igual que los experimentos, estudios de casos y entrevistas; permiten obtener evidencia empírica sobre algún tema de interés. encuesta y c) analizar los resultados obtenidos.

Prototipado Evolutivo Experimental (Método de la Ingeniería). El prototipado evolutivo experimental [14] consiste en desarrollar una solución inicial para un determinado problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución a casos de estudio (problemáticas) de complejidad creciente. El proceso de refinamiento concluye al estabilizarse el prototipo en evolución.

- **Materiales.**

Para la búsqueda de literatura existente se empleará el Sistema Nacional de Repositorios Digitales (SNRD) [15].

Formación de Recursos Humanos

El grupo se encuentra conformado por un Director, un Codirector, tres docentes-investigadores y tres alumnos de grado.

Se estima la formación de dos tesinas de grado de la carrera Licenciatura en Informática de UNAHUR.

¹ Congreso Argentino de Ciencias de la Computación.

² Conferencia Iberoamericana de Computación Aplicada.

³ Congreso Latinoamericano de Ingeniería

Referencias

- [1] “¿Qué es la Industria 4.0?”, *Argentina.gob.ar*, el 7 de abril de 2021. <https://www.argentina.gob.ar/produccion/plan-argentina40/industria-4-0> (consultado el 28 de junio de 2022).
- [2] “Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution - ingenieur.de”, *ingenieur.de - Jobbörse und Nachrichtenportal für Ingenieure*, el 1 de abril de 2011. <https://www.ingenieur.de/technik/fachbereiche/produktion/industrie-40-mit-internet-dinge-weg-4-industriellen-revolution/> (consultado el 20 de junio de 2022).
- [3] OPSSI. Reporte anual 2018 (2018). sobre el Sector de Software y Servicios Informáticos de la República Argentina. Disponible en <https://www.cessi.org.ar/opssi>.
- [4] Informe Argentina Productiva – Economía del Conocimiento (2019). Ministerio de Producción y Trabajo. Presidencia de la Nación. Secretaría de la Transformación Productiva. Diciembre de 2019. Disponible en: <https://biblioteca.produccion.gob.ar/buscar/?fid=16>.
- [5] R. Drath y A. Horch, “Industrie 4.0: Hit or Hype? [Industry Forum]”, *IEEE Ind. Electron. Mag.*, vol. 8, núm. 2, pp. 56–58, jun. 2014, doi: 10.1109/MIE.2014.2312079.
- [6] Ministerio de Desarrollo Productivo (2021). Plan de Desarrollo Productivo Argentina 4.0. Disponible en: <https://www.argentina.gob.ar/produccion/plan-argentina40/industria-4-0>.
- [7] Jefferson Seide Molléri, Kai Petersen, Emilia Mendes (2020). An empirically evaluated checklist for surveys in software engineering. *Information and Software Technology* 119, 106240.
- [8] Basili V., Rombach D. (1988). The TAME project: towards improvement-oriented software environments. *IEEE Transactions on Software Engineering*, 14(6), pp. 758-773.
- [9] Marco Kuhrmann, Paolo Tell, Jil Klünder, Regina Hebig, Sherlock Licorish, Stephen MacDonell (Eds.): *Complementing Materials for the HELENA Study (Stage 2)*. [online] DOI: 10.13140/RG.2.2.11032.65288, published: 2018-11-28.
- [10] Centro PyME-Unahur. Vinculación Tecnológica de la Universidad Nacional de Hurlingham. Disponible: <https://unahur.edu.ar/centro-pyme-unahur/>
- [11] Genero, M., Piattini, M., & Cruz Lemus, J. A. (2014). *Métodos de investigación en Ingeniería del Software*. Madrid: Ra-Ma S.A. Editorial y Publicaciones.
- [12] Riveros, H. y Rosas, L. (1985). *El Método Científico Aplicado a las Ciencias Experimentales*. Editorial Trillas. México. ISBN 96-8243-893-4. 2.
- [13] Creswell, J. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Prentice Hall. ISBN 10: 01-3613-550-1. 3.
- [14] Basili. *The Experimental Paradigm in Software Engineering*. En *Experimental Software Engineering Issues: Critical Assessment and Future Directions* (Ed. Rombach, H., Basili, V., Selby, R.). *Lecture Notes in Computer Science*, Vol. 706. (1993). ISBN 978-3-540-57092-9.
- [15] Ministerio de Ciencia, Tecnología e Innovación. Sistema Nacional de Repositorios Digitales. Disponible en: <https://repositoriosdigitales.mincyt.gob.ar/vufind/>

SISTEMA DE RECOMENDACIÓN UBICUO PARA CONTRIBUIR A LA DISMINUCIÓN DEL ESTRÉS ACADÉMICO EN ESTUDIANTES UNIVERSITARIOS

González, Eugenia Alejandra; Álvarez, Margarita María

Instituto de Investigaciones en Informática y Sistemas de Información
Departamento de Informática - Facultad de Ciencias Exactas y Tecnologías
Universidad Nacional de Santiago del Estero (UNSE)
e-mail: alegonzalez236@gmail.com.ar; alvarez@unse.edu.ar

RESUMEN

Actualmente, uno de los riesgos para nuestra salud física y mental es el estar expuesto constantemente a situaciones que causen estrés. El “estrés académico” es el particular caso en el que se considera, como sujeto central, al alumno que asiste a un instituto de educación. La importancia de su estudio radica en que la persona puede presentar síntomas físicos, psicológicos y/o comportamentales.

Por otro lado, los sistemas de recomendación son herramientas poderosas y populares que se encuentran implementados en diversas aplicaciones web. Su finalidad es sugerir alternativas según el perfil del usuario. Teniendo en cuenta su funcionalidad y las ventajas que ofrece la ubicuidad de los teléfonos móviles, se propone, en esta línea de investigación, el desarrollo de un prototipo funcional y móvil de un sistema de recomendación que indique diferentes alternativas posibles para que el estudiante pueda afrontar la situación académica estresante, de acuerdo al lugar donde se encuentre en ese momento.

El objetivo de esta propuesta es tratar de disminuir el estrés en los alumnos, y al mismo tiempo, mejorar su salud.

Palabras Clave: Estrés Académico, Sistemas de Recomendación, Computación Ubicua.

CONTEXTO

Esta línea de investigación se desarrollará en el marco del proyecto de investigación “*Modelos basados en Inteligencia Artificial y Computación Ubicua para la resolución de Problemas en Educación y otros dominios*”, correspondiente a la convocatoria 2021 de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SICYT - UNSE).

1. INTRODUCCION

Uno de los problemas más extendidos y debilitantes de nuestro bienestar subjetivo y de nuestra sociedad es el estrés. Éste puede definirse como la reacción que tienen las personas cuando están sujetas a exigencias y presiones que no corresponden a sus conocimientos y habilidades, y que pueden desafiar su capacidad de manejo. Si bien los niveles de estrés levemente elevados pueden ser funcionales para la productividad, la investigación ha demostrado que el estrés prolongado y severo, juega un papel en una amplia gama de condiciones físicas, psicológicas y conductuales (Bogomolov et al., 2014).

(Kloster Kantlen & Perrotta, 2019) mencionan que el estrés ha sido objeto de atención por parte de investigadores de diferentes ámbitos del conocimiento. En los últimos años, el interés se ha trasladado a estudiantes universitarios. La manera en que

cada estudiante recorra la universidad, estará regulada por su adaptación a los factores estresores (condiciones ambientales que producen sentimientos de tensión y/o se perciben como amenazantes o peligrosas). Por lo tanto, aquellos que logran asimilar favorablemente los cambios que el ámbito universitario conlleva, aprenderán de la experiencia y se enriquecerán personalmente tomando dichos cambios como un desafío. Sin embargo, en aquellas situaciones contrarias, los estudiantes que no puedan manejar adecuadamente estas exigencias, experimentarán insatisfacción personal y/o desgaste, produciéndose en ellos estrés académico, pudiendo manifestar síntomas físicos como la ansiedad, cansancio, insomnio o, con relación a lo académico, bajo rendimiento, desinterés profesional, ausentismo e incluso deserción.

Los estudiantes que experimentan estrés académico, en un intento de poder adaptarse a estas situaciones nuevas, ponen en movimiento una sucesión de estrategias de afrontamiento para superar, de manera exitosa, las exigencias que les demanda la universidad. En general, el afrontamiento actúa como regulador de un malestar emocional. Si éste es efectivo, no se presentará dicho malestar, pero de lo contrario podría verse afectada la salud de forma negativa (Kloster Kantlen & Perrotta, 2019).

En esta línea de investigación se propone un prototipo funcional y móvil de un sistema de recomendación (SR) para ayudar, a los alumnos que estén vivenciando una situación académica estresante, a conocer diferentes alternativas y recursos útiles para que puedan afrontar y superar exitosamente las exigencias que les demanda la universidad. Se considera que los SR son una herramienta apropiada para dar solución a dicho problema ya que “generan recomendaciones sobre un determinado objeto de estudio, a partir de las

preferencias y opiniones dadas por los usuarios” (Herrera-Viedma, Porcel & Hidalgo, 2004).

Este prototipo tomará como entrada un nivel de estrés y, teniendo en cuenta el perfil del individuo basado en sus preferencias, junto con su ubicación en la cual se encuentre y la hora actual, podrá recomendar alternativas para disminuirlo. Al considerar la ubicación actual del usuario, se pretenden aprovechar las características beneficiosas que ofrece la ubicuidad inherente de los dispositivos móviles. Según (Ding, Hui & Tingting, 2010), para que un sistema de computación ubicua sea capaz de asistir al usuario proactivamente, es necesario que posea la habilidad de sensor el contexto, razonar frente a cambios producidos en dicho contexto y reaccionar de manera acorde, de ser necesario.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Con esta línea de investigación se busca realizar contribuciones teóricas y prácticas en el campo de los sistemas de recomendación ubicuos para el afrontamiento del estrés académico. En consecuencia, el objetivo general es: “Contribuir a disminuir el estrés académico de los estudiantes universitarios mediante un sistema ubicuo de recomendaciones personalizadas”.

Además, se han definido los siguientes objetivos específicos:

1. Definir el modelo del estudiante que incluya las preferencias necesarias junto con su ubicación y horario para permitir las recomendaciones personalizadas.
2. Desarrollar un prototipo funcional móvil de un sistema de recomendación que dé sugerencias sobre cómo disminuir el estrés académico en los alumnos universitarios, teniendo en

cuenta el modelo del estudiante, su ubicación actual y horario.

Con el propósito de dar cumplimiento a los objetivos planteados, se realizarán las siguientes actividades:

1. *Exploración y Recolección de Información.* Se investigarán los temas relevantes en los que se centrará el trabajo, con el objetivo de seleccionar la información que se crea necesaria para su entendimiento y posterior realización, particularmente el marco referencial de éste. Se aplicará, principalmente, la búsqueda, el análisis y el estudio de documentación de diferentes fuentes de información.
2. *Determinación de las principales técnicas que se utilizan para afrontar situaciones académicas estresantes.* Se obtendrá información sobre si, una muestra de alumnos correspondiente a la Universidad Nacional de Santiago del Estero, aplican alguna técnica de afrontamiento en tales situaciones, y en caso de que sí lo hagan, cuáles son. Para ello, se diseñará y construirá una encuesta de carácter anónimo a través de herramientas de ofimática, luego del análisis y síntesis de la información recabada en la actividad previa.
3. *Recopilación, análisis y listado de las técnicas que se utilizan para afrontar situaciones académicas estresantes.* Se recolectarán y organizarán los resultados de las encuestas, obteniendo información y conclusiones acerca de cuáles son las situaciones académicas que causan estrés entre los estudiantes, y sus respectivas formas de afrontamiento. Se espera obtener una primera aproximación que relacione los niveles de estrés a partir de las

situaciones académicas indicadas por los estudiantes y sus técnicas de afrontamiento. Una vez ordenadas las respuestas, se comparará con la información obtenida de la investigación de la primera etapa, o si se considera necesario, se indagará más sobre el tema.

4. *Selección de las herramientas de desarrollo del sistema de recomendación y del algoritmo de recomendación.* Se investigarán y estudiarán herramientas software específicas que faciliten la creación del sistema de recomendación. Además, también se decidirá el algoritmo de recomendación que dicho sistema utilizará.
5. *Desarrollo del sistema de recomendación.* Teniendo en cuenta toda la información reunida, obtenida como resultado de las etapas anteriores, se diseñará y programará un prototipo funcional móvil, para el sistema operativo Android, de un sistema de recomendación que contará con una encuesta para obtener las preferencias del alumno y construir así su perfil de usuario. Con estos datos más la información de su ubicación y el horario actuales, le permitirá al sistema proponer, como recomendación, la aplicación de alguna técnica de afrontamiento según el nivel de estrés que presente el estudiante.
6. *Validación del sistema de recomendación construido.* Consiste en comprobar que, dados un nivel de estrés se obtienen las recomendaciones de afrontamiento correspondientes.

3. RESULTADOS ESPERADOS

El resultado que se espera obtener de esta investigación, conforme a los objetivos, es un prototipo de un sistema de recomendación funcional y móvil destinado para todos aquellos alumnos que sufran de estrés académico.

Este sistema considerará el nivel de estrés del alumno, su contexto geográfico, horario actual y sus preferencias para brindar las sugerencias más adecuadas, con el objetivo de ayudar a disminuir su estrés y, a largo plazo, permitiéndoles que puedan llevar una vida más saludable durante su etapa universitaria.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo de esta línea se integra por un estudiante, un profesor guía, dos asesores y el director del Proyecto de Investigación.

En esta línea de investigación, el estudiante desarrollará su trabajo final para optar por el título de Licenciado en Sistemas de Información de la Universidad Nacional de Santiago del Estero.

5. BIBLIOGRAFIA

Bogomolov, A., Lepri, B., Ferron, M., Pianesi, F. & Pentland, A. S. (2014, Noviembre). Daily Stress Recognition from Mobile Phone Data, Weather Conditions and Individual Traits. *Proceedings of the 22nd ACM international conference on Multimedia*. <https://doi.org/10.1145/2647868.2654933>

Kloster Kantlen, G. E. & Perrotta, F. D. (2019, Septiembre). *Estrés Académico en Estudiantes Universitarios de la Ciudad de Paraná* [Trabajo Final de Licenciatura en Psicología]. UNIVERSIDAD CATÓLICA ARGENTINA

Herrera-Viedma, E.; Porcel, C. & Hidalgo, L. (2004). Sistemas de recomendaciones: herramientas para el filtrado de información en Internet. *Hipertext.net*, n.º 2, <https://raco.cat/index.php/Hipertext/article/view/57818>

Ding, G. J., Hui L. & Tingting Z. (2010). A preliminary study of personal learning environment based on Ubiquitous Computing Model. *2010 3rd IEEE International Conference on Ubi-Media Computing*, pp. 350-354, doi: <https://doi.org/10.1109/umedia.2010.5544429>

Macías, A. B. (2010). Un modelo conceptual para el estudio del estrés académico. *Revista Electrónica de Psicología Iztacala; Vol 9, No 3, 9(3)*

EL AVANCE EN LA MADUREZ DIGITAL DE LAS PYMES A TRAVÉS DE LA TRANSFORMACIÓN DIGITAL

Lic. Abel Marcelo Del Do, mdeldo@uaco.unpa.edu.ar; Dra. Andrea Villagra, avillagra@uaco.unpa.edu.ar; Mg. Daniel Pandolfi, dpandolfi@uaco.unpa.edu.ar

Instituto de Tecnología Aplicada (ITA)
Unidad Académica Caleta Olivia - Universidad Nacional de la Patagonia Austral.
Ruta Nacional 3 Acceso Norte, Caleta Olivia, Santa Cruz, Argentina.

RESUMEN

La transformación digital es un proceso de innovación disruptivo en las PYMES y amenaza su existencia. El modelo comercial tradicional de las PYMES se ve desafiado por la presión de la digitalización impuesta por los competidores por un lado y por los cambios en los hábitos de compra de sus clientes por el otro. La mayoría de las iniciativas de transformación digital fracasan debido a que las organizaciones no cambiaron su mentalidad y procesos, o no construyeron culturas que fomentaran el cambio. Es por esta razón que, a pesar de todas las oportunidades habilitadas digitalmente para recuperar el poder competitivo, las PYMES todavía dudan en adoptar las tecnologías emergentes como herramientas y aplicaciones digitales.

El objetivo de esta línea de trabajo es examinar los impulsores y las barreras de la transformación digital de las PYMES para identificar los posibles desafíos desencadenantes que puedan promover su transformación digital. Además, se propone profundizar en el conocimiento de cómo las PYMES pueden avanzar en la transformación digital a través del desarrollo de capacidades organizacionales. Identificando la situación actual, analizando modelos de madurez existentes, su adaptabilidad a este tipo de empresas y la aplicabilidad para gestionar su evolución digital.

Palabras clave: *Transformación Digital; PYMES; Tecnologías Emergentes; Madurez Digital.*

CONTEXTO

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEem), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/B273 “Ciudades inteligentes y sostenibles: iniciativas y desafíos”. Y particularmente en el marco de la Beca de Investigación para Estudiantes de Postgrado de la UNPA, según Resolución N°0199/22-R-UNPA.

1. INTRODUCCIÓN

Los avances tecnológicos, cambios del entorno y el desarrollo de Internet en la última década ha permitido al mundo cambiar sistemáticamente hacia una nueva era industrial llamada Industria 4.0. Este concepto fue introducido en Alemania en la segunda década del siglo XXI hace referencia a la industria modernizada motorizada por uso de Internet que, conectada a través de sensores inteligentes, hace posible usar esta inteligencia en favor de las personas y ha abierto las puertas de la conexión global, el intercambio de información y análisis de datos (Lombardero, 2015). El término “transformación digital” se ha convertido en un término muy utilizado y responde al desarrollo de diversos campos de investigación (Schallmo et al., 2017).

El concepto “transformación” hace referencia a un cambio sustancial en la empresa, impactando directamente en la estrategia y la estructura organizacional (Matt et al., 2015) y

alude a un cambio disruptivo del modelo de negocio, de la organización empresarial y de las capacidades generadas por la aplicación de las tecnologías digitales. Por otro lado, el término “digital” hace referencia a la utilización de las nuevas tecnologías digitales asociadas directamente con la Cuarta Revolución Industrial, y con el desarrollo de tecnologías emergentes tales como Big Data, Inteligencia Artificial, Computación en la Nube, el Internet de las Cosas, Blockchain, entre otras.

Estas tecnologías han impulsado nuevas formas de hacer negocios y a construir mercados emergentes. Esta disrupción digital trae consigo las amenazas y oportunidades más importantes a las que se han enfrentado las empresas en los últimos años; la integración y explotación de las nuevas tecnologías digitales se ha convertido en uno de los mayores desafíos a los que se enfrentan. Ninguna es inmune a sus efectos y del éxito de su asimilación dependerá el desempeño futuro de las mismas (Hess et al., 2016; Kane et al., 2015) ya que tienen la capacidad para transformar los productos, servicios, operaciones, e incluso los modelos de negocio de las empresas, así como su entorno competitivo.

Como la incorporación de nuevas tecnologías es fundamental para mantener la competitividad de las organizaciones y su desarrollo futuro, es de gran importancia estudiar el potencial de las capacidades organizacionales de transformación digital para asimilarlas y ponerlas en valor. Las empresas necesitan establecer prácticas de gestión para dirigir estas transformaciones complejas. Según Morgan (2019), cerca de un 70% de las iniciativas de transformación digital fracasan debido a que las organizaciones no cambiaron su mentalidad y procesos, o no construyeron culturas que fomentaran el cambio.

Si bien la transformación digital afecta a empresas de cualquier tamaño, las PYMES son de particular interés a este respecto debido a su importancia en la economía, ya que representan más del 99% de las empresas en la Unión Europea (Muller et al., 2018) y el 99,5% de las empresas en América Latina y el Caribe. Según

el estudio de la OECD/CAF (2019) las PYMES juegan un papel clave en la generación de empleo y de valor añadido en la economía mundial, y se están viendo afectadas por la digitalización imparable que está transformando el entorno competitivo en el que se desenvuelven.

Aunque el concepto de transformación digital está más desarrollado actualmente, aún no existen modelos o marcos de buenas prácticas que ayuden a los directivos de las PYMES (Peter et al., 2020), esto provoca que la transformación digital se aborde con demasiada frecuencia de manera general, resultando muy abstracta para gerentes y empleados. Los directivos siguen sin tener los conocimientos y experiencia necesarios para el desarrollo e implementación de estrategias digitales, y esta carencia ha sido identificada como una de las razones principales que ha impedido a un gran número de empresas implementar sus planes de transformación digital (Hess et al., 2016; Matt et al., 2015). Además, el número de investigaciones realizadas sobre transformación digital centradas en capacidades organizacionales son reducidas y no suelen centrarse en PYMES (Evans, 2017).

Investigadores como Chanias et al. (2019) proponen que contar con modelos de madurez puede ayudar a definir la estrategia digital, y demuestra que la elaboración de una estrategia de transformación digital es diferente a la planificación estratégica inicial de una empresa; y a su vez es un proceso altamente dinámico que implica iterar entre aprender y hacer. Específicamente, demuestra que una estrategia de transformación digital está continuamente en desarrollo, sin un final previsible.

Prepararse para la transformación digital no es tarea fácil, es preciso desarrollar capacidades digitales en las que las actividades, las personas, la cultura y la estructura de la organización estén sincronizadas y alineadas con un conjunto de objetivos organizacionales (Kane et al., 2016). De esta manera las transformaciones digitales exitosas requieren que las empresas cultiven nuevas capacidades

organizacionales para sobrevivir y prosperar (Li et al., 2018).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Con la intención de cubrir la brecha de conocimiento identificada y contribuir a ampliar la investigación sobre transformación digital en las PYMES. En esta línea de investigación abordaremos las siguientes áreas de interés: (a) Profundizar en el conocimiento de cómo las PYMES pueden avanzar en la transformación digital a través del desarrollo de capacidades organizacionales. (b) Identificar cuál es la situación actual de las PYMES en relación a la transformación digital. (c) Determinar cuáles son los habilitadores e inhibidores fundamentales para avanzar en la transformación digital y los desafíos más importantes a los que se enfrentan para transformarse digitalmente. (d) Analizar la adaptabilidad de los modelos de madurez de transformación digital existentes, su relación con la estrategia empresarial de las PYMES y cómo se alinean con la estrategia digital.

Para lograr esto analizaremos la transformación digital dentro del contexto de la Cuarta Revolución Industrial que está viviendo la sociedad actual, identificando las nuevas tecnologías digitales que fomentan la transformación digital de las PYMES y las impulsan a avanzar en la madurez digital. Uno de los objetivos de este estudio es examinar los impulsores y las barreras de la transformación digital de las PYMES para identificar los posibles desafíos desencadenantes que puedan promoverla. Por lo tanto, inicialmente realizamos una revisión sistemática de la literatura sobre las investigaciones existentes hasta el momento en la literatura académica, con el fin de identificar los habilitadores e inhibidores que encuentran las PYMES para llevar a cabo la transformación digital, y los desafíos que estas enfrentan en el proceso de avanzar en su madurez digital.

Luego comprobaremos la utilidad de los modelos de madurez de transformación digital existentes, que en su mayoría fueron diseñados

para grandes empresas, y analizaremos su adaptabilidad en el estudio de la madurez digital de las PYMES. Llevaremos adelante un Estudio de Caso implementando un proyecto de Transformación Digital. Identificando las adaptaciones necesarias para realizar este tipo de proyectos de innovación en PYMES, cuáles son las consideraciones a tener en cuenta para implementarlos en este tipo de empresas y las capacidades digitales necesarias para que las PYMES puedan afrontarla exitosamente.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Tres características fundamentales nos indican que estamos ante una nueva revolución industrial: La velocidad del cambio actualmente es exponencial, de una magnitud superior en comparación a las revoluciones industriales anteriores. No hay ninguna industria que esté al margen de los cambios producidos por las tecnologías digitales. La complejidad de los cambios, su amplitud y profundidad predicen una transformación completa de los sistemas de producción, gestión y gobernanza de las organizaciones.

La Cuarta Revolución Industrial afectará fundamentalmente a las expectativas de los clientes, la mejora del producto, la innovación colaborativa y las formas organizativas. La velocidad a la que se están produciendo innovaciones disruptivas, los nuevos servicios basados en datos y su análisis, la mejora de las experiencias del cliente, etc. requieren de nuevas formas de colaboración, en una nueva realidad en la que surgen nuevos modelos de negocio y plataformas digitales globales.

Nuevas Tecnologías: El acrónimo **SMACiT** (*Social, Mobile, Analytics, Cloud Computing e Internet of Things*) se refiere a un conjunto de tecnologías digitales potentes, de fácil acceso para las organizaciones y alto grado de complementariedad; por eso son ideales para las PYMES y nos focalizamos en estas. La integración de estas tecnologías está generando una digitalización progresiva y sin precedentes que fomenta la innovación y transformación de

las organizaciones. Esta transformación supone uno de los mayores retos a los que se enfrentan en la actualidad, ya que ninguna es inmune a sus efectos y del éxito de su asimilación dependerá el desempeño futuro de las mismas (Hess et al., 2016; Kane et al., 2015). Existen otras tecnologías digitales, como la inteligencia artificial, blockchain, robótica, realidad virtual y otras que, aunque afectarán a las organizaciones, no son consideradas SMACiT.

Las PYMES tienen unas características diferentes a las grandes empresas: recursos limitados y capacidades de especialización restringidas. Por eso son más prudentes en la toma de riesgos, y en decidirse a incorporar nuevas tecnologías que actúen como facilitadores de la transformación digital. Sin embargo, también cuentan con ventajas. Presentan un nivel de jerarquía más simple que permite una toma de decisiones más rápida (North y Varvakis, 2016); y los ejecutivos pueden participar en todas las decisiones importantes, lo cual garantiza el compromiso de la dirección en este tipo de proyectos.

En Argentina las PYMES fueron enfrentadas a dos pandemias: Covid-19 y crisis económica. Las PYMES se vieron forzadas por la pandemia COVID-19 a encarar un proceso de transformación digital acelerado para poder subsistir. Entendiendo el impacto que provocó la pandemia a este tipo de empresas, la sustentabilidad y supervivencia de las mismas no sólo dependerá de su voluntad para innovar y sus ganas de seguir siendo productivas, sino de las decisiones gubernamentales, reglas previsibles y consensos básicos económicos y sociales que incluyan a este sector entre los prioritarios al momento de tomar decisiones. A pesar de que la pandemia aceleró los procesos de digitalización de muchas empresas, que debieron actualizarse para seguir funcionando, la falta de madurez organizacional se presenta como el mayor inhibidor para implementarla.

Uno de los principales desafíos a los que se enfrentan los directivos de las PYMES es cómo asimilar las nuevas tecnologías digitales e integrarlas al modelo de negocio para aprovechar todo su potencial. Un liderazgo

digital eficaz facilitará la implementación exitosa de los proyectos de transformación digital.

El avance en la madurez digital, a través de la transformación digital, permitirá a las PYMES adaptarse y competir en un entorno cada vez más digital. Las organizaciones utilizan la transformación digital para mejorar su relación con el cliente, aunque aquellas que han alcanzado un mayor nivel de madurez digital están transformando su negocio para obtener ventajas competitivas dentro del mercado.

La madurez digital es un concepto holístico que comprende tanto un aspecto tecnológico, como un aspecto de gestión. Aquellas organizaciones que abarquen el desarrollo de ambos aspectos alcanzarán un mayor nivel de madurez digital (Shahiduzzaman et al., 2017). El concepto de madurez digital es un fenómeno relativamente nuevo en el cual ninguna organización ha alcanzado en la actualidad el estado final de madurez. Por esta razón, es conveniente utilizar el término "madurando digitalmente" para describir las empresas más avanzadas en su proceso de transformación digital (Kane et al., 2015). Las empresas focalizadas en la integración de las tecnologías digitales, en su conjunto, y en la transformación de sus negocios alcanzan mayores niveles de madurez digital que aquellas focalizadas en resolver problemas concretos con tecnologías digitales individuales.

Entonces, podemos afirmar que la madurez digital está más relacionada con tener una estrategia digital clara que con una implementación mayor o menor de las nuevas tecnologías. Como indican Kane et al. (2017), la estrategia es el diferenciador más fuerte de las empresas que maduran digitalmente. Por tanto, uno de los mayores desafíos de las organizaciones, para avanzar en la madurez digital, será desarrollar una estrategia eficaz y enlazarla con los objetivos generales de la empresa.

4. FORMACIÓN DE RECURSOS HUMANOS

Actualmente el autor está desarrollando su tesis de Maestría en Informática y Sistemas de la UNPA y además cuenta con una beca de Investigación de Postgrado.

5. BIBLIOGRAFÍA

- Chanias, S., Myers, M.D., Hess, T., 2019. Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. *J. Strateg. Inf. Syst.* 28, 17-33. <https://doi.org/10.1016/j.jsis.2018.11.003>
- Evans, N.D., 2017. Mastering digital business : how powerful combinations of disruptive technologies are enabling the next wave of digital transformation. BCS Learning & Development.
- Hess, T., Benlian, A., Matt, C., Wiesböck, F., 2016. Options for Formulating a Digital Transformation Strategy. *MIS Q. Exec.* 15, 103-119.
- Kane, G.C., Palmer, D., Phillips, A.N., Kiron, D., Buckley, N., 2015. Strategy, Not Technology, Drives Digital Transformation. *MIT Sloan Manag. Rev.* Deloitte Univ. Press 27.
- Kane, G.C., Palmer, D., Phillips, A.N., Kiron, D., Buckley, N., 2016. Aligning the Organization for Its Digital Future. *MIT Sloan Manag. Rev.* Deloitte Univ. Press.
- Kane, G.C., 2017. Digital maturity, not digital transformation. *MIT Sloan Manag. Rev.* URL <https://sloanreview.mit.edu/article/digital-maturity-not-digital-transformation/> (Consultado el 23.2.23).
- Li, L., Su, F., Zhang, W., Mao, J.-Y., 2018. Digital transformation by SME entrepreneurs: A capability perspective. *Inf. Syst. J.* 28, 1129-1157. <https://doi.org/10.1111/isj.12153>
- Lombardero, J. 2015. Problemas y retos de la gestión empresarial en la economía global: Estudio comparado y sistémico de competencias directivas. Madrid: Universidad Camilo José Cela.
- Matt, C., Hess, T., Benlian, A., 2015. Digital Transformation Strategies. *Bus. Inf. Syst. Eng.* 57, 339-343. <https://doi.org/10.1007/s12599-015-0401-5>
- Morgan, B., 2019. Companies that failed at digital transformation and what we can learn from them. *Forbes*. URL: <https://www.forbes.com/sites/blakemorgan/2019/09/30/companies-that-failed-at-digital-transformation-and-what-we-can-learn-from-them/?sh=1322ba20603c> (Consultado el 20/2/2023)
- Muller, P., Mattes, A., Klitou, D., Lonkeu, O., Ramada, P., Ruiz, F., Shaan Devnani, J., Farrenkopf, A., Makowska, N., Mankovska, N.R., Steigertahl, L., 2018. Annual report on European SMEs 2017/2018: SMEs growing beyond borders. <https://doi.org/10.2873/248745>
- North, K., Varvakis, G., 2016. Competitive Strategies for Small and Medium Enterprises, *Competitive Strategies for Small and Medium Enterprises*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-27303-7>
- OECD/CAF, 2019. América Latina y el Caribe 2019: Políticas para PYMES competitivas en la Alianza del Pacífico y países participantes de América del Sur, OECD Publishing, Paris, <https://doi.org/10.1787/60745031-es>.
- Peter, M.K., Kraft, C., Lindeque, J., 2020. Strategic action fields of digital transformation. *J. Strateg. Manag.* 13, 160-180. <https://doi.org/10.1108/JSMA-05-2019-0070>
- Shahiduzzaman, M., Kowalkiewicz, M., Barrett, R., McNaughton, M., 2017. Digital Business Towards a Value-Centric Maturity Model, PWC Report Chair in Digital Economy. Brisbane.
- Schallmo, D., Williams, C.A., Boardman, L., 2017. Digital transformation of business models - Best practice, enablers, and roadmap. *Int. J. Innov. Manag.* 21, 1740014. <https://doi.org/10.1142/S136391961740014X>

Tecnologías de la información facilitadoras para la interoperabilidad de software en Gobierno Abierto: Segunda Parte

Roxana Martínez, Ezequiel Ricciardi, Gastón Lacuesta,
Nicolás Mellino, Francisco Bourse

Instituto de Tecnología (INTEC)
Universidad Argentina de la Empresa (UADE)
Lima 775, Ciudad Autónoma de Buenos Aires, Argentina

{mariarmartinez, ericciardi, glacuesta, frbourse, nmellino}@uade.edu.ar

RESUMEN

Conocer aspectos relacionados con el tratamiento de la interoperabilidad, permite comprender cómo las plataformas digitales logran procesar e interactuar datos a través de recursos informáticos, con el fin de brindar una gestión colaborativa entre distintos organismos públicos en aspectos del contexto de gobierno abierto. Para llevar a cabo una gestión colaborativa de datos, de forma eficiente y estandarizada, es necesario comprender que los procedimientos para realizar dichas tareas deberán implicar intercambios simples y seguros de datos, fomentando así, una gran base de conocimiento público y accesible.

Este trabajo permitirá conocer el estado y situación actual de las tecnologías utilizadas en el contexto de Gobierno Abierto, las cuales se encuentren enfocadas en el eje de la interoperabilidad de datos y las tecnologías, orientadas a las herramientas que permiten brindar servicios colaborativos entre los organismos estatales y los ciudadanos. Esto es con el fin de compartir recursos de información, y poder mejorar los procesos involucrados en las actividades diarias de las entidades gubernamentales. Por otra parte, este trabajo ofrece un prototipo de validación de la metadata de los conjuntos de datos abiertos públicos, que son disponibilizados en los portales web, y, además, en este proyecto se presentan cuáles son las mejores prácticas en aspectos de validaciones y tratamiento de datos ofrecidos, para que sean manipulados por cualquier persona que desee reutilizar estos en otros softwares o bien alimentar

otras fuentes diversas de bases de datos.

Palabras clave: Datos Abiertos, Gobierno Abierto, Interoperabilidad en informática, Diseño y Arquitectura de componentes para intercambio de datos.

CONTEXTO

El presente trabajo es la segunda parte del proyecto denominado “Tecnologías de la Información facilitadoras para la Interoperabilidad en Gobierno Abierto”, que tuvo inicio en el mes de octubre 2021 y que tiene fecha de culminación para fines del mes de septiembre 2023. Este proyecto pertenece a la línea de investigación de Ingeniería de Software (IS) del Instituto de Tecnología (INTEC). Las actividades del Instituto están directamente relacionadas con las carreras grado y posgrado de la Facultad de Ingeniería y Ciencias Exactas de la Universidad Argentina de la Empresa (UADE). Algunas de las tareas de este proyecto, se realizan en los laboratorios disponibles en los UADE Labs, edificio tecnológico inaugurado en 2010. El proyecto es financiado y evaluado por la Coordinación de Investigaciones, unidad que depende de la Secretaría Académica de la Universidad, tiene una duración de 2 años, y cuenta con la participación de docentes y estudiantes de grado y posgrado.

1. INTRODUCCIÓN

El concepto de gobierno abierto promueve que varios gobiernos de diversos Estados

Nacionales se orienten al uso de una metodología que fomente el enfoque de transparencia para las actividades realizadas en el Estado. Uno de los objetivos principales de Gobierno Abierto, es inculcar este concepto a través de la publicación de tareas y actividades a los ciudadanos u organismos estatales para lograr una rendición de cuentas de lo realizado, tanto en aspectos de gestión gubernamental, como en la administración de los recursos de un país. Este paradigma, tiene como una de sus finalidades, el diseño de políticas públicas transparentes y la concientización sobre los beneficios que pueden traer aparejados el trabajo colaborativo entre Gobierno y el ciudadano.

Desde enfoques tecnológicos, Gobierno Abierto “es una doctrina política que surge a partir de la adopción de la filosofía del movimiento del software libre a los principios de la democracia. Este paradigma tiene como objetivo que la ciudadanía colabore en la creación y mejora de servicios públicos y en el robustecimiento de la transparencia y la rendición de cuentas” [1]. La transparencia consiste en asegurar el derecho de todos los ciudadanos, y se orienta al libre acceso a la información pública gubernamental. De esta manera, un gobierno puede mostrar de manera sencilla y clara, la gestión llevada a cabo, y así promover a la gestión activa [2]. Este derecho a la información pública permite conocer y utilizar datos que producen o tienen los tres poderes del Estado, la Ley que ampara este acceso es la Ley 27.275 [3].

Por todo lo anteriormente dicho, los datos abiertos públicos, cobran un interés fundamental en la disponibilización y utilización de estos, tanto a los organismos como a los ciudadanos, ya que estos datos, permiten la implementación de la interoperabilidad entre entidades públicas, es decir, intercambiar información y el uso de esta [4], en las que se incluye la utilización de técnicas, semánticas, aspectos legales y políticas [5] que facilitarían este nuevo modelo. Todo esto conlleva a un gran beneficio a través de los protocolos de la interoperabilidad, porque permite que las áreas puedan trabajar en forma

interconectada, evitando así, la duplicación de tareas y el espacio de almacenamiento, entre otras ventajas [6]. Uno de los puntos importantes para tener en cuenta, es la diferencia entre la interoperabilidad y la integrabilidad.

Según [6], la interoperabilidad permite un espacio colaborativo en el que los organismos exponen y consumen servicios de forma constante, es decir, sin un limitante, básicamente pueden compartir entre todas las entidades del estado que se encuentren asociadas. El éxito de la plataforma reside en la retroalimentación constante, esto tiene 3 ejes centrales siendo éstos: el espacio colaborativo, la transformación digital y los beneficios al ciudadano. Por otro lado, la integrabilidad, es la capacidad de integración entre distintos productos/tecnologías para intercambiar datos y servicios de manera segura, esto logra que se genere un ecosistema de interoperabilidad. Tiene como ejes, la integración y la transparencia.

Para aplicar dichos aspectos, es importante mejorar la calidad de los datos abiertos públicos [7] que se disponibilizan y, a su vez, mejorar técnicas de interoperabilidad entre éstos. Un punto importante a tener en cuenta puede ser el análisis de la experiencia de los países líderes en este contexto, en los que se basan de buenas en guías de mejores prácticas, y, además, contar con un modelo de calidad estándar como la ISO 25012 [8], para generar valor y aprovechamiento por parte de los usuarios [9], tanto a nivel datos públicos, como la metadata que es necesaria para el intercambio e indexación entre estos datos. También, el enfoque de características necesarias para la compatibilidad con otras plataformas gubernamentales teniendo en cuenta algunos recursos, tales como: estándares de la W3C [10], estándares abiertos para APIs estilo REST [11], Pautas Técnicas de Interoperabilidad de Sistemas (Anexo II RES 19 E-2018) [12], Estándares de Servicios Digitales y Herramientas de la Plataforma Digital del Sector Público Nacional [13] [14], Herramientas para la construcción colaborativa de contenidos de la Plataforma Digital del Sector Público

Nacional [15], entre otras.

A modo de resumen, es necesario que el diseño y construcción de los servicios digitales se base en soluciones que logren una adecuada comunicación e intercambio [16]. Para que este intercambio sea posible, previamente deben abordarse aspectos de las cuatro dimensiones de la interoperabilidad [17]: Legal o Jurídica; Semántica; Organizativa; y Técnica.

Finalmente, es preciso concientizar a los organismos gubernamentales, que la implementación de la interoperabilidad conlleva a grandes beneficios a la sociedad, en aspectos de [18]: lograr una gestión pública de calidad que posibilite la provisión eficiente de bienes y servicios públicos a los ciudadanos, además de alcanzar una mayor integración y desarrollo de la sociedad, impulsando la utilización de sistemas y el uso intensivo de las tecnologías de la información y las comunicaciones (TIC) por parte del Estado.

2. LÍNEAS DE INVESTIGACIÓN DESARROLLO

Este proyecto pertenece a la línea de investigación de Ingeniería de Software (IS) del Instituto de Tecnología (INTEC). Para esta segunda parte del proyecto, los ejes principales del tema que se están investigando en referencia a las actividades de I+D+i (Investigación, Desarrollo e Innovación) son:

- Continuar con el estudio de análisis de las distintas tecnologías desde el nivel técnico de utilización de interoperabilidad de los datos, y compararlos con los ya relevados.
- Continuar con el desarrollo de guías de las mejores prácticas en aspectos de niveles técnicos para el tratamiento de los datos públicos que son asociados (relacionados entre sí) y visualizados en los portales estatales nacionales.
- En base a las falencias de las herramientas de IT detectadas, continuar con el desarrollo de un prototipo para la validación y detección del estado de los datos disponibilizados en los portales gubernamentales, tanto nacionales como

internacionales en aspectos de interoperabilidad.

- Análisis y estudio de casos para efectuar mejoras en los procedimientos utilizados en interoperabilidad.

Fundamentación del proyecto:

Desde el enfoque en Audiencia Científica:

Los resultados (parciales y finales) de este proyecto serán transferidos para contribuir con nuevo conocimiento a la comunidad académica y, a la sociedad en general en el ámbito nacional e internacional.

Desde el enfoque en Audiencia como Institución:

La audiencia interna la constituyen los estudiantes y docentes de la carrera de Ingeniería Informática y afines, pudiendo también incorporarse integrantes de diferentes carreras interesados en las cuestiones del gobierno abierto en esta temática.

Desde el enfoque Social: El proyecto corresponde a un abordaje desde la disciplina del Software a aspectos del ámbito de la gestión pública, tratándose de cuestiones de creciente preocupación. Su base fundamental surge por la necesidad de tener un mejor acercamiento tanto con el ciudadano como con las organizaciones a través del uso de las TIC.

3. RESULTADOS OBTENIDOS/ESPERADOS

Enfoque del proyecto:

Se buscará brindar nuevos enfoques especialmente en el diseño y construcción de tecnologías para la interoperabilidad, entendida como la capacidad de comunicar y transferir datos entre sistemas de información para posibilitar el intercambio de información y conocimiento entre ellos, como así también en el tratamiento de los datos. Por las características del proyecto guarda fuerte vínculo con conceptos de Ingeniería de Software, Diseño de Software y Arquitectura de Software.

Objetivos principales:

Identificar los principales aspectos tecnológicos que son facilitadores para llevar a cabo la interoperabilidad en el contexto de Gobierno Abierto y brindar buenas prácticas a niveles técnicos en lo que concierne a esta temática, con el fin de facilitar un adecuado manejo de interoperabilidad de datos colaborativos entre organizaciones. Esto conlleva a comprender que elementos son fundamentales en este nuevo paradigma gubernamental desde un aspecto de diseño y arquitectura de software.

Metodología y Técnicas:

En los que respecta a las técnicas de investigación para esta segunda etapa del proyecto, se continúa con el desarrollo de enfoques cualitativos que permiten obtener estudios enfocados a:

- Investigación bibliográfica: Con el fin de apoyar y sustentar el trabajo investigativo. Releva los estudios previos y contrastarlos a la actualidad.
- Estudios de Casos tanto en la República Argentina, como países internacionales en aspectos prácticos del tratamiento de los datos.
- Estudio de Técnicas de interoperabilidad y conexión interna de datos abiertos en organismos públicos nacionales desde el diseño y arquitectura de software.
- Análisis y encuadre del framework de la metadata en datos abiertos públicos gubernamentales.
- Herramientas de visualización (dashboard) interoperables.

Resultados obtenidos:

En lo que fue la primera parte del proyecto (primer año), se relevaron los conceptos fundamentales en esta temática que se relacionan con el marco de interoperabilidad, además, se estudiaron algunas de las herramientas de IT en este contexto, con el fin de conocer el grado de utilización de las mismas, como así también, se analizaron las falencias en cuestiones tecnológicas actuales de interoperabilidad. Por otra parte, se llevó a cabo un estudio general del marco regulatorio

legal de la interoperabilidad en la Administración Pública Argentina, para comprender el alcance del tratamiento de datos (privacidad, sensibilidad del dato, etc.). Se elaboró un estudio sobre el estado situación de la interoperabilidad en contexto de datos abiertos públicos actuales del país, y su historia a lo largo de los años, comprendiendo su crecimiento y adaptación en determinados períodos. En base a los resultados encontrados, se realizó el desarrollo de un pequeño prototipo para la validación de la metadata de los datasets disponibilizados en los sitios web gubernamentales, con el fin de conocer el estado de “salud” en aspectos de interoperabilidad de estos. Finalmente, se desarrolló una breve guía de las mejores prácticas en aspectos de niveles técnicos para el tratamiento de los datos públicos, la cual se continúa con el desarrollo de esta.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto se compone por 2 (dos) docentes investigadores de grado y posgrado que cuentan con varios años de experiencia en el dictado de clases en varias Universidades de Argentina y más de 18 años de experiencia en el ambiente tanto laboral como académico en tecnología informática. Uno de ellos es docente investigadora, Doctora en Ciencias Informáticas de la Universidad Nacional de La Plata (UNLP), actual directora del Proyecto de Investigación y Magister en Tecnología Informática de la Universidad Abierta Interamericana (UAI), además, el proyecto cuenta con otro docente investigador Magister en Procesos Educativos Mediados por Tecnologías de la Universidad Nacional de Córdoba y actual Doctorando en Ciencias Sociales de la Universidad Nacional de Mar del Plata.

Por otro lado, se cuenta con la participación de estudiantes de grado de carreras de la rama de Informática y Sistemas de la UADE (Universidad Argentina de la Empresa) en la que su participación es en calidad de colaboradores de la carrera de Ingeniería en

Informática, como así también, en el desarrollo de su tesis final de carrera, 2 (dos) estudiantes.

Por lo que a modo de resumen y en relación directa con la línea de I+D+i presentada para el proyecto, los miembros del equipo se encuentran en realización de: 2 trabajos finales de carreras afines en la UADE.

5. BIBLIOGRAFÍA

[1] Sosteniblepedia.org. “Gobierno Abierto”. Disponible en:

https://www.sosteniblepedia.org/index.php?title=Gobierno_abierto

[2] Martínez, R. (2022). “*Métricas de calidad para validar los conjuntos de datos abiertos públicos gubernamentales*” (Doctoral dissertation, Universidad Nacional de La Plata).

[3] Infoleg - Información Legislativa. “Derecho de Acceso a la Información Pública – Ley 27275”. Sancionada por el Congreso el 14 de septiembre de 2016. Disponible en: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/265000-269999/265949/norma.htm>

[4] IEEE Standard Glossary of Software Engineering Terminology. (1990). IEEE Std 610.12. *IEEE Computer Science*, 42.

[5] Jiménez, C. E., Solanas, A., & Falcone, F. (2014). E-government interoperability: Linking Open and Smart Government. *IEEE Computer Society*, 22-24.

[6] Argentina.gob.ar. “Interoperabilidad”. Jefatura de Gabinete de Ministros. Disponible en: <https://www.argentina.gob.ar/jefatura/innovacion-publica/innovacion-administrativa/interoperabilidad>

[7] Martínez, R. et al. (2021). Metrics proposal to measure the quality of governmental datasets. *IEEE Latin America Transactions*, Vol. 100. ISSN 1548-0992.

[8] ISO 25012 (2008). “*Ingeniería de software - Requisitos de calidad y evaluación de productos de software (SQuaRE) - Modelo de calidad de datos*”. Disponible en: <https://www.iso.org/obp/ui/es/#iso:std:iso-iec:25012:ed-1:v1:en>

[9] de Colombia, G. (2020). Calidad e Interoperabilidad de los datos abiertos del Gobierno de Colombia. Disponible en: https://herramientas.datos.gov.co/sites/default/files/2020-11/A_guia_de_estandares_final_0.pdf

[10] W3C. “Standards”. Disponible en:

<https://www.w3.org/standards/>

[11] OpenAPI (2022). “OpenAPI Initiative”. Disponible en: <https://www.openapis.org/>

[12] Resolución, Secretaría de Modernización Administrativa (2018). “*Pautas Técnicas de Interoperabilidad de Sistemas*”. Disponible en: https://www.argentina.gob.ar/sites/default/files/ont/ont/res_19_2018_anexo_ii_if_2018_09344892_apn_ssga_mm.pdf

[13] Argentina.gob.ar (2022). “Contenidos Digitales”. Disponible en:

<https://www.argentina.gob.ar/contenidosdigitales>

[14] HM Government (2012). “Open Standards Principles. For software interoperability, data and document formats in government IT specifications”. Disponible en: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/459074/Open-Standards-Principles-2012.pdf

[15] BID (2021). “Reutiliza estas herramientas”. Disponible en:

<https://code.iadb.org/es/funcionalidad/intecambio-e-interoperabilidad-de-datos>

[16] Araujo, S., & Vargas, M. P. (2020). La interoperabilidad en el marco del Gobierno Digital. Conocimiento Libre y Licenciamiento (CLIC), (22).

[17] CEPAL - Comisión Económica para América Latina y el Caribe (2023). “*Desde el gobierno digital hacia un gobierno inteligente*”. Bibliogúas - Biblioteca de la CEPAL. Disponible en:

<https://biblioguias.cepal.org/gobierno-digital/interoperabilidad>

[18] Gobierno de la Provincia de Buenos Aires. “Subsecretaría de Gobierno Digital”. Gobierno Digital. Disponible en: <https://www.gba.gob.ar/gobiernodigital/interoperabilidad>

Desarrollo de una app móvil de Nueva Generación

*Adriana Elizabeth Martín¹ ; Susana Beatriz Chavez²; Sergio Rafael Flores³ ; A. Sara Zogbe⁴,
Nelson Rodríguez⁵; María A. Murazzo⁶,*

*Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
Complejo Islas Malvinas. Ignacio de la Roza y Meglioli.
C.P. 5402. Rivadavia. San Juan, 0264 4234129*

*¹ arianamartinsj@gmail.com ² schavez@iinfo.unsj.edu.ar; ³ sergior@gmail.com;
⁴ sarazogbe@yahoo.com.ar; ⁵ nelson@iinfo.unsj.edu.ar; ⁶ marite@unsj-cuim.edu.ar,*

Resumen

La computación móvil ha avanzado notablemente en los últimos años, surgiendo a nivel de desarrollo millones de apps de la más variada utilidad, área de aplicación y costos. Muchas empresas y organizaciones de gobierno ofrecen soluciones para sus usuarios que son potenciadas en nuevas versiones. La pandemia aceleró el uso de los teléfonos inteligentes o smartphones, que pasaron a ser el aula de clase virtual, la biblioteca de PDFs, el cuaderno donde se hacen las actividades, el dispositivo para sacar turno para algún trámite, donde reproducir videos de estudio o complemento de recreación y muchos otros usos. Esto potenció el uso de la red y los dispositivos, permitiendo el surgimiento de nuevas aplicaciones a partir también de tecnologías emergentes. Las redes 5G y 6G, las superapps, la computación inmersiva y aplicaciones móviles basadas en AI, van a tener un rol significativo en los próximos años y es el objeto de estudio de esta línea de investigación.

Palabras claves: *Mobile Computing, SuperApps, Inteligencia Artificial, Computación Inmersiva.*

Contexto

El presente trabajo se encuadra dentro del área de I/D Innovación en Sistemas de Software y es una de las líneas de investigación internas, del proyecto: Soporte Serverless para aplicaciones móviles de nueva generación, cuya propuesta está en etapa de evaluación para el período 2023-2024. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil y distribuida desde hace más de 22 años. Como continuación de los proyectos anteriores: Computación Serverless para tratamiento de datos provenientes de dispositivos de IoT, y Modelo de sistema de comunicación en programación reactiva. Se continúa el trabajo con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

En los últimos años, la tecnología móvil ha avanzado rápidamente, y algunas de las tendencias más notables incluyen la adopción de redes móviles más rápidas, el surgimiento de las superapps, la computación inmersiva y las aplicaciones móviles basadas en inteligencia artificial (IA). La quinta generación de redes móviles (5G), adoptada en muchos países, permite velocidades de descarga mucho más rápidas y menor latencia, esto significa que los dispositivos pueden conectarse y comunicarse entre sí en tiempo real.

Superapps

Las superapps ofrecen una amplia gama de servicios y funciones en una sola plataforma, lo que las convierte en una opción conveniente para los usuarios que buscan simplificar sus vidas digitales.

Una superapp es una aplicación móvil que ofrece múltiples servicios y funciones en una sola plataforma. Estas aplicaciones son muy populares en países como China y algunos otros de Asia, y están ganando terreno en otros lugares del mundo. En lugar de tener varias aplicaciones para realizar diferentes tareas, como pedir comida, hacer una reserva de hotel o comprar productos en línea, reservar taxis hasta comprar entradas de cine y enviar dinero a amigos y familiares. Una superapp ofrece todo esto y más en una sola aplicación. Además, las superapps suelen tener una función de pago integrada, lo que significa que los usuarios pueden realizar pagos sin tener que salir de la aplicación.

Algunos ejemplos de superapps populares incluyen WeChat y Alipay en China, Grab en el sudeste asiático y Gojek en Indonesia.

En resumen, las superapps son aplicaciones móviles que ofrecen una amplia gama de servicios y funciones en una sola plataforma, lo que las convierte en una opción conveniente para los usuarios que buscan simplificar sus vidas digitales.

Computación inmersiva

La Computación Inmersiva es un concepto que se refiere a una experiencia informática en la que el usuario se sumerge en un entorno generado computacionalmente a través de interfaces de usuario avanzadas y dispositivos de entrada/salida como pantallas 3D, gafas de realidad virtual, guantes con sensores, y otros dispositivos de interacción.

La idea detrás de la computación inmersiva es crear un entorno informático que permita a los

usuarios interactuar de manera natural y fluida con el mundo virtual, sintiéndose como si estuvieran dentro de él. Esto se logra mediante el uso de tecnologías avanzadas, como la realidad virtual (VR), la realidad aumentada (AR) y la realidad mixta (MR).

La computación inmersiva tiene aplicaciones en una variedad de campos, como los videojuegos, la educación, la simulación, la medicina, el arte, el entretenimiento y muchos otros. Por ejemplo, en la medicina, se está utilizando la realidad virtual para la formación de cirujanos, la terapia de rehabilitación y la simulación de procedimientos médicos complejos. En la educación, se está utilizando para mejorar la experiencia de aprendizaje y enseñar a los estudiantes de manera más interactiva e inmersiva.

El entorno de la tecnología inmersiva presenta diferentes características como fruto de un progreso del continuo avance del hardware, lo cual permite que dichas tecnologías se vuelvan accesibles. Concretamente, se presentan dos características principales en la tecnología inmersiva, como son la unión de diferentes tecnologías y la adaptación y modificación sensorial.

En resumen, la Computación Inmersiva es una experiencia informática que utiliza tecnologías avanzadas como la realidad virtual, la realidad aumentada y la realidad mixta para sumergir al usuario en un entorno generado por ordenador y permitir una interacción natural y fluida con ese mundo virtual.

Servicio Serverless

Serverless es una arquitectura de computación en la nube que está ganando popularidad en el mundo de las aplicaciones móviles de nueva generación. En lugar de tener que gestionar y mantener infraestructuras complejas y costosas, el desarrollo de aplicaciones serverless permite que los desarrolladores se centren en escribir código de aplicación sin preocuparse por la infraestructura subyacente. Esto se logra mediante la externalización de la

gestión de servidores a los proveedores de servicios en la nube, como Amazon Web Services (AWS).

AWS Amplify es un conjunto de herramientas y servicios que permite a los desarrolladores crear y escalar aplicaciones móviles sin tener que preocuparse por la infraestructura subyacente. Amplify proporciona una plataforma completa para la construcción de aplicaciones móviles que incluye funcionalidades como la autenticación de usuarios, el almacenamiento de datos, el análisis y la integración con otras herramientas y servicios de AWS. Además, Amplify es compatible con una amplia gama de tecnologías y marcos de trabajo, lo que lo hace flexible y escalable. Actualmente, es compatible con iOS, Android y JavaScript (web y React Native).

En este trabajo se propone como plataforma de desarrollo a Flutter por ser un conjunto de herramientas que permite crear aplicaciones móviles, web y de escritorio, compiladas de forma nativa a partir de una única base de código. Ahora bien, Flutter y AWS Amplify se reúnen en Amplify Flutter, dando lugar a un espacio integrado de trabajo con todo el poder de AWS.

En resumen, la arquitectura serverless es una forma innovadora y rentable de construir aplicaciones móviles de nueva generación, y AWS Amplify es una solución completa que ofrece a los desarrolladores una manera fácil de construir, escalar y administrar aplicaciones móviles. Con el aumento de la demanda de aplicaciones móviles de alta calidad y con un rendimiento rápido y escalable, la adopción de la arquitectura serverless y las herramientas de Amplify es una opción inteligente para los desarrolladores móviles que buscan impulsar sus aplicaciones al siguiente nivel.

Objetivos

El objetivo del grupo de investigación en esta línea de conocimiento es analizar e investigar

la plataforma AWS Amplify y los servicios de Flutter para desarrollar un prototipo de Superapps que pueda ser utilizada con múltiples propósitos.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de seis investigadores que figuran en este trabajo de las universidades Nacional de San Juan y dos alumnos de grado. Además, el proyecto marco donde se está desarrollando esta propuesta ha establecido vínculos con investigadores de la Universidad Nacional de San Luis, la Universidad Champagnat y la Universidad Nacional de Salta y dos alumnos de grado. Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis de maestría en áreas afines y dos tesis de grado en el área de Serverless computing, Concurrencia y Computación distribuida. Además se espera aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.

Referencias

- [1] Jonas E, Schleier-Smith J, Sreekanti V, Tsai C-C, Khandelwal A, Pu Q, Shankar V, Carreira J, Krauth K, Yadwadkar N, Gonzalez JE, Popa RA, Stoica I, Patterson DA (2019) Cloud Programming Simplified: A Berkeley View on Serverless Computing. <http://arxiv.org/abs/1902.03383>. Accessed 6 Jan 2021
- [2] S. Eismann et al., ‘Serverless Applications: Why, When, and How?’ (2021), IEEE Software, vol. 38, no. 1, pp. 32–39, Jan. 2021, doi: 10.1109/MS.2020.3023302.
- problems. In Research Advances in Cloud Computing (pp. 1-20). Springer, Singapore.
- [6] Nastic, S., Rausch, T., Scekkic, O., Dustdar, S., Gusev, M., Koteska, B. & Prodan, R. (2017). A serverless real-time data analytics

platform for edge computing. IEEE Internet Computing, 21(4), 64-71.

[7] Mohanty, S. K., PremSankar, G., & Di Francesco, M. (2018). An Evaluation of Open Source Serverless Computing Frameworks. In CloudCom (pp. 115-120).

[8] Gottlieb, N. (2016). State of the Serverless Community Survey Results.

<https://serverless.com/blog/state-of-serverless-community/>.

[9] Jonas, E., Pu, Q., Venkataraman, S., Stoica, I., & Recht, B. (2017). Occupy the cloud: Distributed computing for the 99%. In Proceedings of the 2017 Symposium on Cloud Computing (pp. 445-451). ACM.

[10] Fromm, K. (2012). <https://readwrite.com/2012/10/15/why-the-future-of-software-and-apps-is-serverless/>

[11] Van Eyk, E., Iosup, A., Seif, S., & Thömmes, M. (2017). The SPEC cloud group's research vision on FaaS and serverless architectures. In Proceedings of the 2nd International Workshop on Serverless Computing (pp. 1-4). ACM.

[12]https://es.wikipedia.org/wiki/Tecnologia_inmersiva

[13] Desarrollo de pila completa: AWS Amplify. <https://aws.amazon.com/es/amplify/>

Soluciones IoT desde Smart Cities a Blockchain

Javier Díaz, Laura Fava, Agustín Candia, Emanuel Borda, Diego Vilches, Matías Pagano
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata, calle 50 esq. 120, 2^{do} Piso.
Tel: +54 221 4223528
{jdiaz, lfava, acandia, eborda, dvilches, mpagano}@linti.unlp.edu.ar

RESUMEN

La transformación digital viene avanzando rápidamente trayendo nuevas soluciones tecnológicas para múltiples escenarios. Esta transformación tiene como objetivos generar una mayor eficiencia operacional y una mejor experiencia en el uso de productos y servicios, además de ofrecer cada vez más inteligencia y seguridad.

Muchas de estas tecnologías emergentes, incluso, son aún más eficientes cuando se las aplica conjuntamente, como en el caso de las tecnologías de *Blockchain* e *Internet de las Cosas*, que juntas pueden ofrecer grandes soluciones a muchas problemáticas actuales.

La tecnología de Internet de las cosas (Internet of Things o IoT) ha crecido exponencialmente con la implementación de proyectos que la utilizan para monitorear dispositivos -en el campo, en los hogares, en las ciudades, etc.- lo que la ha colocado en la cima de la transformación digital brindando muchas soluciones a problemas actuales. Sin embargo, este crecimiento acelerado ha generado importantes desafíos como la administración de una enorme cantidad de dispositivos conectados y el manejo de una gran cantidad de datos e información. Es muy desafiante para la infraestructura actual y las arquitecturas existentes administrar eficientemente estos emergentes ecosistemas IoT.

Por otro lado, blockchain es una

tecnología que hace uso de sistemas descentralizados y criptografía; introduce un cambio de paradigma por el hecho de ser una red distribuida donde cada uno de los nodos participantes son los responsables de la validación y registro de las transacciones.

En los sistemas IoT tradicionales los dispositivos dependen de un servidor central en la nube para identificar y autenticar dispositivos individuales. Por el contrario, con la blockchain cada dispositivo firma criptográficamente sus mensajes descentralizando la tarea de identificación y autenticación, además de ayudar a mitigar el downtime debido a su naturaleza distribuida.

En este artículo se describen líneas de investigación y desarrollo que forman parte de un proyecto integral destinado a proveer soluciones tecnológicas que aporten a diferentes escenarios de smart cities, áreas rurales, marítimas o alejadas de lo urbano donde la conectividad 3G o 4G resulta ineficaz o inaccesible. Se presenta a LoRaWAN como red de soporte, y se introduce una nueva línea, blockchain, para trabajar conjuntamente con IoT.

Palabras claves: Internet de las Cosas (IoT), sensores, protocolos para IoT, Lora, LoRa WAN, blockchain.

CONTEXTO

El Laboratorio de Investigación de

Nuevas Tecnologías Informáticas LINTI de la Facultad de Informática, viene trabajando en proyectos relacionados con Internet de las Cosas en áreas como domótica, estacionamiento inteligentes, horticultura, deporte, etc., desde hace varios años.

Uno de los primeros proyectos de IoT vinculado a ciudades inteligentes fue *rParking: un sistema de plazas de estacionamiento reservadas* (Boccalari, E., González, F., 2016) actualmente en funcionamiento en la playa de estacionamiento del edificio de Rectorado de la UNLP y en la Ciudad de Las Condes, Chile. Actualmente se está investigando para desarrollar una tacha de estacionamiento de bajo costo que pueda ser fabricada en Pymes locales. Este nuevo nodo permitiría detectar el estado (ocupado/desocupado) de una plaza de estacionamiento, e informar este estado vía LoRaWAN. También se prevé integrarlas a Sistemas de estacionamiento medido existentes.

Asimismo se han propuesto soluciones de IoT para ayudar a personas con discapacidad como el calzado háptico, una solución tecnológica integral en forma de zapato que va acompañado por una aplicación móvil que guía a los disminuidos visuales a un destino específico en tiempo real y el desarrollo de un kit de bastón y anteojos con sensores que detectan y avisan la presencia de objetos en espacios bajos y altos para personas ciegas (Harari, I., Fava, L., Díaz, J., 2018).

Las líneas de trabajo que se describirán a continuación se desarrollan en el LINTI y están enmarcadas en el proyecto: *“De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región”*, , acreditado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

1. INTRODUCCIÓN

El término Internet de las Cosas o Internet of Things (IoT) se refiere generalmente a escenarios donde la capacidad de cómputo y la conectividad de las redes se extienden a objetos, sensores y elementos cotidianos -no computadoras personales-, permitiendo que estos dispositivos generen, intercambien y consuman datos.

Los proyectos antes mencionados, dan cuenta del trabajo que se viene desarrollado con IoT en diferentes escenarios, mayormente utilizando wi-fi y bluetooth, sin embargo, cuando se necesitan comunicaciones sobre distancias largas esos protocolos no son adecuados. En este sentido, han surgido nuevas tecnologías que permiten comunicación eficiente en distancias muy largas, donde no existe conectividad celular (GPRS, 3G o 4G), y también en escenarios donde las capacidades de cómputo y energía del nodo sean claramente limitadas. Dentro de estas tecnologías conocidas como LPWAN tenemos LoraWAN y Sigfox.

En este artículo se presentan avances en diferentes proyectos IoT que utilizan LoraWAN para brindar soluciones a problemáticas de diferentes escenarios y se introduce una nueva línea, blockchain, y su impacto en IoT.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Las líneas de investigación, desarrollo e innovación que se llevan a cabo en este proyecto están vinculadas al desarrollo de soluciones para ciudades inteligentes, el agro y otros escenarios, utilizando el protocolo LoRaWAN.

Los ejes principales de I+D+i son:

- Análisis de protocolos de comunicación

de largo alcance y bajo consumo. Despliegue de la red, instalación y prueba.

- Análisis de tecnologías de vanguardia para la construcción de nuevos dispositivos inteligentes para smart cities.
- Diseño y construcción de dispositivos basados en sensores para el monitoreo y análisis de la salud de los animales.
- Análisis de sistemas de posicionamiento para tracking de animales: GPS y localización sin GPS usando LoRaWAN.
- Diseño e implementación de plataformas horizontales que gerencien los dispositivos mencionados, escalen en cantidad y diversidad de nodos y permitan trabajar con los datos recolectados.
- Análisis de tipos de blockchains, estudio de los algoritmos de consenso, blockchain y su impacto en la seguridad de IoT.

3. RESULTADOS OBTENIDO/ESPERADOS

Para la línea de IoT planteada en este artículo se ha trabajado por un lado en la red de soporte LoRaWAN y por otro en la creación de dispositivos inteligentes que serán conectados utilizando esa red y aplicaciones de gestión desarrolladas específicamente para cada tipo de dispositivo.

En cuanto a los proyectos que están en funcionamiento podemos mencionar:

En el contexto de una tesina y en cooperación con un proveedor del rubro semaforización, se lleva a cabo el desarrollo de un prototipo funcional de red de *semáforos inteligentes*. Se incorporan funcionalidades tales como monitoreo del estado lumínico del artefacto,

sincronización de onda verde, sensores de posicionamiento e inclinación, habilitación de canales para circulación ante situaciones de emergencia, entre otras. La solución pensada de extremo a extremo, incluye el diseño de un nodo controlador, facilidades de comunicación vía LoRaWAN, monitoreo en tiempo real y herramientas de configuración y analítica.

Otro de los proyectos finalizados y deployados en varios lugares es la construcción de nodos para *monitorear la calidad del aire*, proyecto que ha sido de mucha utilidad en la pandemia de COVID-19.



Fig 1.: Nodo de medición de calidad de aire

Este nodo de monitoreo de calidad de aire, basado en el uso de sensores portátiles. Se analizó la performance de 3 modelos diferentes de sensores para la medición de material particulado, uno de los contaminantes primarios determinado por la Organización Mundial de la Salud (OMS, 2006).

Como parte del análisis exploratorio, se realizaron campañas de monitoreo en conjunto con un equipo de referencia certificado, a fines de analizar desvíos en

las mediciones y correlación entre los diferentes sensores. La tarea se desarrolló en conjunto con profesionales del ámbito de la tecnología ambiental pertenecientes a la UNLP (Candia, A. , 2018). La Fig. 2 muestra una imagen de un nodo de medición de Co2.



Fig 2.: Nodo de medición de Co2

Como ampliación del tema de calidad de aire y en consonancia con la pandemia de COVID-19, se desarrolló una línea de investigación sobre las condiciones de ventilación en ambientes de trabajo cerrados. Como resultado se implementaron dispositivos para medir las concentraciones de Co2, temperatura y humedad, con el fin de alertar cuando se alcanzaran umbrales que pudieran poner en riesgo de contagio a los ocupantes de un espacio de trabajo cerrado.

Otro proyecto en el que se ha trabajado y que se encuentra operativo es el **sistema de balizamiento del canal de entrada al Puerto La Plata**, partido de Ensenada. Esta red de sensores desplegada sobre las boyas -actualmente 20-, demarcan el canal de ingreso de embarcaciones comerciales, y permite acceder a información estratégica y operativa en forma oportuna

y veraz para la toma de decisiones a nivel operativo. Entre los objetivos alcanzados, se puede mencionar que el sistema permite un monitoreo remoto de los paneles solares y baterías instalados, una sincronización exacta en las señales de balizamiento. La Fig. 3 (a) muestra un nodo de sensores desplegado sobre una baliza y la (b) muestra una imagen de la torre donde se despliegan.



Fig 3.: (a) Nodo de sensores - (b) Despliegue

Actualmente se está trabajando en la integración con el sensado de mareógrafos y correntómetros, el envío de comandos para programación de tareas y la gestión de alarmas para dar aviso ante situaciones anómalas.

Otro de los proyectos está relacionado con **Tokenización y Tracking de Ganado** usando tecnologías IoT y blockchain. La sensorización de animales, la recolección de datos y su análisis es otra línea de trabajo. En este marco se está desarrollando un prototipo de collar para monitorear ganado. Este proyecto se lleva a cabo en conjunto con la Facultad de Veterinaria de la UNLP y se usará como escenario de despliegue el Establecimiento Don Joaquín, perteneciente a la UNLP, ubicado en Bavio. La tecnología blockchain permitirá identificar de manera segura y unívoca -a través de la

criptografía- y descentralizada -a través de una arquitectura de nodos blockchain- cada uno de los animales que forman parte del campo.

Todos estos proyectos en sus respectivos escenarios y con sus particularidades comparten un eje común, se trabaja desde el sensor a la aplicación incluyendo sistemas de alerta, monitoreo y visualización de datos.

Finalmente, cabe destacar que para dar soporte a todos estos proyectos se desplegó una red LoRaWAN usando la infraestructura de torres de la UNLP y el backbone de fibra óptica que las conecta. La arquitectura se implementa en una topología de estrellas en la que dispositivos repetidores (gateways) retransmiten mensajes entre los dispositivos finales (sensores) y un servidor de red central

Los gateways están conectados al servidor de red a través de conexiones IP estándar y actúan convirtiendo paquetes de RF en paquetes IP y viceversa. El despliegue de gateways en las antenas seleccionadas, cubrió de forma satisfactoria el área de interés, incluso utilizando una baja densidad de equipos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I+D+i presentada en este artículo se encuentra formado por docentes investigadores categorizados del LINTI y alumnos avanzados de la Licenciatura en Informática, Licenciatura en Sistemas e Ingeniería en Computación perteneciente a la Facultad de Informática y a la Facultad de Ingeniería.

En relación a las tesinas de grado vinculadas con esta línea de investigación, se está dirigiendo una tesina de grado vinculada a semaforización inteligente, una tesina relacionada con blockchain y dos PPS relacionadas con IoT y deporte. Asimismo se está participando con la Facultad de Ciencias Exactas de un proyecto de calidad de aire, financiado por el Ministerio de Ciencia y Tecnología.

5. REFERENCIAS

Fava, L., Vilches, D., Candia, A. Diaz, J. *Soluciones IoT con tecnología LORA*, XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019), ISBN: 978-987-3984-85-3. Marzo, 2019.

Harari, I, Fava, L, Diaz, J. *Rampas Digitales Innovativas para Personas con Discapacidad*. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018), ISBN: 978-987-3619-27-4. Abril, 2018.

Boccalari, E., González, F. *rParking: Sistema de plazas de estacionamiento reservadas*. Tesina de Grado accesible en <http://sedici.unlp.edu.ar/handle/10915/59485>. Octubre, 2016.

Fava, L., Vilches Antao, D., Díaz, J., Pagano, M., Romero Dapozo, R., *Tecnología aplicada al deporte de alto rendimiento*, XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018), ISBN: 978-987-3619-27-4. Abril, 2018.

Kashoash, H, Kemp, A. *Comparison of 6LoWPAN and LPWAN for the Internet of Things*. Australian Journal of Electrical and Electronics Engineering. Diciembre, 2017.

Candia, A., Luengo, M. Represa, S., Porta, A. Marrone, L. *Solutions for SmartCities: proposal of a monitoring system of air quality based on a LoRaWAN network with low-cost sensors*. ISBN: 978-1-5386-5447-7, 2018.

Joseph Thachil G., *Introducing Blockchain Applications: Understand and Develop Blockchain Applications Through Distributed Systems*. ISBN: 9781484274804, 2021.

Antonopoulos, A. *Mastering Bitcoin, Programming the Open Blockchain, 2nd Edition*. ISBN: 9781491954386, 2017.

Joseph Thachil G., *Introducing Blockchain Applications: Understand and Develop Blockchain Applications Through Distributed Systems*. ISBN: 9781484274804, 2021.

Elad Elrom, *The Blockchain Developer: A Practical Guide for Designing, Implementing, Publishing, Testing, and Securing Distributed Blockchain-based Projects.*, ISBN 9781484248478, 2019.

RECURSOS/DESARROLLOS EXISTENTES

Ciudades Sustentables,
<https://www.argentina.gob.ar/ambiente/ciudadessustentables>

LORA Alliance, <https://www.lora-alliance.org>

Design Support Technical resources and documents,
<https://www.semtech.com/design-support>

Diseñar una solución de identidad auto-gestionada para acceso a servicios de calidad con redes Blockchain multipropósito en la Universidad Nacional de Río Negro

Mauro Cambarieri¹, Alejandra Viadana¹, Nicolás García Martínez¹, Luis Vivas¹, Carlos Lugani¹

¹ Universidad Nacional de Río Negro. Sede Atlántica
Laboratorio de Informática Aplicada
{mcambarieri,caviadana,ngarciam,lvivas,clugani}@unrn.edu.ar

RESUMEN

El objetivo de este proyecto es diseñar la Identidad Digital Autogestionada en la UNRN adoptando el nuevo paradigma de producción de software, basado en tecnologías descentralizadas.

Se utilizarán tecnologías blockchain aportando a la transformación digital en el ámbito universitario y desarrollo de servicios públicos digitales innovadores. Se desarrollarán componentes de software para optimizar procesos y servicios utilizando tecnologías descentralizadas, es decir, las redes de web 3, ofrecen una alternativa al deterioro del status quo digital. La introducción de Blockchain, plantea una revolución tecnológica que repercute directamente en cambios organizacionales, económicos y políticos.

La Identidad Digital le da la posibilidad a cada individuo la administración de sus datos, como los certificados analíticos y la forma en que serían presentados a terceros a través de billeteras digitales.

El reto de las redes descentralizadas es que son bienes públicos. Al no haber una entidad central que controle las decisiones y obtenga beneficios, es difícil incentivar su mantenimiento y desarrollo. La criptografía ayuda a resolver este problema mediante la coordinación descentralizada y la provisión de incentivos económicos para el desarrollo.

Como resultado se espera la sensibilización a la comunidad, autoridades, entre otros, además de, divulgar los resultados, formación de recursos humanos y transferencia tecnológica.

Palabras clave: Web 3, Blockchain, LACChain, Wallet, descentralización

Inteligentes” desarrollado en el Laboratorio de Informática Aplicada, Sede Atlántica, Universidad Nacional del Río Negro (UNRN). Se pondrá especial énfasis en la utilización en tecnologías blockchain aportando a la transformación digital en el ámbito universitario y desarrollo de servicios públicos digitales innovadores. Se enfocará en la necesidad de instituciones universitarias nacionales e internacionales. Se desarrollarán componentes de software para optimizar los procesos, políticas y servicios utilizando tecnologías descentralizadas, teniendo en cuenta la relevancia de la tercera era de internet, es decir, las redes descentralizadas de web 3.

Contamos con una fortaleza para el desarrollo de este proyecto, ya que en el año 2022 el LIA se sumó a un proyecto liderado por el BID Lab con el objetivo del desarrollo del ecosistema blockchain en América Latina y el Caribe. Como resultado nace la alianza LACChain que se encuentra compuesta por un grupo de organizaciones que están activamente participando en el desarrollo de aplicaciones blockchain y su uso real por la sociedad, cuyo objetivo se centra en materializar las oportunidades que representa la tecnología blockchain para la región, haciéndola viable.

Esta Alianza Global, permitió realizar el despliegue de un nodo escritor en los servidores del LIA, en el cual es posible realizar transacciones de información sensible sobre la Blockchain, permitiendo sumar integrabilidad al proyecto de investigación 40-C-875, además de contar con un entorno de prueba para el despliegue y desarrollo de aplicaciones descentralizadas que se pretende llevar a cabo en la presente propuesta.

CONTEXTO

El presente trabajo se enmarca en el proyecto de investigación PI 40-C-875 “Herramientas Informáticas de Dominio Específico para el Desarrollo de Servicios Digitales Innovadores para Comunidades Urbanas y Rurales en el Marco de Ciudades y Regiones”

1. INTRODUCCIÓN

En los años '90 nació lo que se denominó Web 1, Tim Berners-Lee y Robert Cailliau juntos crearon la World Wide Web, en la que un sitio web típico de un usuario (autoalojado) que estaría formado por texto hipervínculo de, archivo, imágenes, aplicaciones y otros objetos digitales que pueden ser leídos y/o

descargados por los navegadores, es lo que se conoce como "Web 1.0" o "web de sólo lectura" [1]; luego llegaron las interacciones del usuario y las redes sociales, esta web marcó el agrupamiento, ordenamiento, la gran centralización y monopolio de algunos proveedores, La Web 2.0 (término acuñado por Tim O'Reilly en 2007 [2]) o "web de lectura y escritura" (término acuñado por Richard McManus en 2003) empezó a desarrollarse en la década del 2000, cuando surgieron plataformas como Facebook, Amazon, eBay, etc. Sin embargo, no refleja la idea inicial de que la Web como medio para el intercambio seguro y descentralizado de datos públicos y privados [3]. La Web 3 o internet del valor es la evolución de la Web 2 y lo que se espera de esta es que sea una red completamente descentralizada, sin "censura", donde los usuarios puedan compartir información de forma segura sin temor a que sea borrada o modificada [4]. Esto es posible gracias a blockchain, esta tecnología de registro distribuido que facilita una lista ampliada de registros transaccionales irrevocables ordenados cronológicamente y firmados criptográficamente que comparten todos los participantes de una red, además permite eliminar la necesidad de intermediarios y también garantiza la integridad de los datos al registrar el historial de todas las transacciones. Esto significa que la Web 3 será una internet más segura que la actual [5], proveyendo una plataforma libre, abierta, democrática.

La Tecnología blockchain, ha ido creciendo su interés de forma exponencial como resultado de las nuevas aplicaciones de uso potencial que ya se está viendo no sólo en el ámbito financiero donde tuvo su primer y más exitoso campo de aplicación con las criptomonedas, sino en el gran interés por parte de múltiples sectores económico-financieros y sociales - tales como la Academia, la Industria, el Agro-, por su potencial para ofrecer soluciones a gran escala [6]. Esta tecnología de registro distribuido (DLT- por sus siglas en inglés Distributed Ledger Technology) es una tecnología que facilita una lista ampliada y ordenada cronológicamente de registros transaccionales irrevocables y firmados criptográficamente que comparten todos los participantes de una red. Cualquier participante con los derechos de acceso adecuados puede rastrear un evento transaccional, en cualquier momento de su historia, perteneciente a cualquier actor de la red. La tecnología almacena las transacciones de forma descentralizada. Las transacciones de intercambio de valor se ejecutan directamente entre pares conectados y se verifican de forma consensuada mediante algoritmos a través de la red. La introducción de Blockchain, plantea una revolución tecnológica que repercute directamente en cambios organizacionales, económicos y políticos. Esta nueva era del internet del valor, blockchain y la Web 3, implicarán un gran desafío de adaptación y una gran oportunidad hacia la transición digital, económica, social y política de nuestras sociedades. La sociedad hoy cuenta con

herramientas que permiten auto-organizarse como nunca antes y expresarse sin la necesidad de intermediarios [17]. La contribución de la web 3 en las primeras etapas de desarrollo, permitirá que las comunidades sean incentivadas y recompensadas por mantener y desarrollar la infraestructura de base(blockchain) [7].

La importancia de la tercera era de internet, las redes descentralizadas de web 3, ofrecen una alternativa al deterioro del status quo digital. Aunque la centralización ha ayudado a millones de personas a acceder a la tecnología, muchas de estas de uso gratuito, también ha "ahogado" a la innovación. En la actualidad, las empresas o corporaciones propietarias de las redes tienen un gran poder unilateral sobre cuestiones importantes como, por ejemplo: acceso a la red, datos de los usuarios, etc. Las plataformas centralizadas siguen un ciclo de vida predecible. Cuando empiezan, hacen todo lo posible por captar usuarios y terceras partes, como desarrolladores, empresas y medios de comunicación. Lo hacen para que sus servicios sean más valiosos, ya que las plataformas (por definición) son sistemas con múltiples intereses en la red. A medida que las plataformas ascienden en la curva de adopción, su poder sobre los usuarios y las terceras partes aumenta constantemente. Cuando llegan a la cima de la curva, sus relaciones con los participantes de la red pasan de ser de suma "positiva" a "cero". La forma y estrategia de seguir creciendo consiste en extraer datos de los usuarios y competir por la audiencia y los beneficios: Ejemplos históricos de ello son Microsoft contra Netscape, Google contra Yelp, Facebook contra Zynga y Twitter contra clientes de terceros. Esto dificulta que las empresas emergentes, los desarrolladores y otros grupos aumenten su presencia en Internet, ya que deben preocuparse de que las plataformas centralizadas cambien las reglas y puedan llevarse sus audiencias o beneficios [4][18].

El reto clásico de las redes descentralizadas es que son bienes públicos. Al no haber una entidad central que controle las decisiones y obtenga beneficios, es difícil incentivar su mantenimiento y desarrollo. La criptografía ayuda a resolver este problema mediante la coordinación descentralizada y la provisión de incentivos económicos para el desarrollo [8].

La Web 3 pondrá el poder en manos de las comunidades y no de las empresas. A continuación, se presentan algunas de las características claves que se incluyen en esta nueva internet: 1- la identidad digital autogestionada (según Sovrin, "identidad autogestionada (IAG) es un término utilizado para describir el movimiento digital que reconoce que un individuo debe poseer y controlar su identidad sin la intervención de las autoridades administrativas. La IAG permite a las personas interactuar en el mundo digital con la misma libertad y capacidad de confianza que en el

mundo físico”) [9] 2- Contratos inteligentes(método automatizado para realizar transacciones comerciales en línea); 3- Dapps (aplicaciones descentralizadas que se ejecutan en Web 3, ejecutan códigos basados en contratos inteligentes que permiten a los usuarios acceder a sus servicios a través del sistema) y 4- DAOs (Organizaciones descentralizadas, donde los participantes de un proyecto que se ejecuta en la Web 3 están a cargo del destino de su negocio, ya que les da más poder y más posibilidad de votación en cada decisión, que en las estructuras verticales tradicionales).

En el presente proyecto, se enfocará en uno de los aspectos claves antes mencionados. La Identidad Digital Autogestionada da la posibilidad a cada individuo la administración de sus datos y la forma en que serían presentados a terceros.

Entre esos datos (credenciales) a los cuales los usuarios tendrán acceso a su administración soberana podríamos enumerar: títulos académicos, certificados analíticos, certificados de alumnos regulares, rendimiento académico, etc. Para que esto sea posible se recurre al desarrollo de billeteras digitales que podrán disponer en sus dispositivos en forma de aplicaciones móviles.

Concepto de identidad digital

“La identidad es un conjunto de atributos relacionados con una entidad.” – ISO/IEC 24760-1 [11]. Para definir la identidad digital se aclara que el dueño de dicha identificación tiene la posibilidad de elegir cuáles atributos quiere hacer públicos mostrando esta capacidad para mostrar el control que genera la condición de Descentralizada de esta nueva Identidad Digital (ID).

Otra característica importante de la ID es la inclusión de los ciudadanos, convirtiéndolos en ciudadanos digitales ya que a partir de esta tecnología permite usar estos servicios de forma remota y en tiempo real, incluyendo a pobladores de lugares a los que su ubicación y dificultad de acceso a la misma pueden acceder a sus credenciales que luego podrán administrar de forma personal sin la necesidad de la intervención de las entidades otorgantes.

Según el Banco Mundial, las identidades digitales se crean y utilizan como parte de un ciclo vital que se compone de cuatro etapas: (a) registro, incluyendo inscripción y validación, (b) emisión de documentos o credenciales, (c) autenticación de identidad y (d) autenticación para la prestación de servicios o transacciones. [10] [11].

La literatura sobre Identidad Digital Autogestionada ha aceptado como válidos los 10 principios establecidos por Christopher Allen en 2016 [12]. Los mismos hacen referencia al Acceso, Consentimiento para el uso de datos por terceros de ser necesario

Control de sus identidades, Existencia independiente de los usuarios, Interoperabilidad, Minimización de los reclamos de los usuarios y la difusión de los mismos, Persistencia en tanto las identidades deben ser sostenidas en el tiempo, Protección de los datos y derechos de los usuarios, Portabilidad tanto de la información y los servicios utilizados. Transparencia en los algoritmos utilizados.

De acuerdo a Marcos Allende López [13]: “Consideraremos que la identidad auto-gestionada es un modelo de identidad digital siempre que cumpla con los 16 principios siguientes.

- Las personas pueden generar sus propios identificadores únicos (control, existencia).
- Las personas tienen el control de sus autenticadores (acceso, control, existencia).
- Las personas tienen el control de sus credenciales y certificados digitales (acceso, control, existencia).
- Las personas pueden recuperar las credenciales y certificados en caso de pérdida o robo de sus autenticadores (acceso, control, existencia, persistencia y protección).
- Las personas administran y controlan los datos asociados con su identidad digital (acceso, control).
- Las personas pueden hacer divulgaciones selectivas de datos (consentimiento, control, minimización, protección).
- La información de identificación personal (IIP) de los individuos se minimiza (minimización, protección).
- Las pruebas criptográficas de la propiedad de los identificadores se pueden encontrar en una red pública descentralizada (interoperabilidad, persistencia, transparencia).
- Las pruebas criptográficas de la propiedad y la validez de las credenciales se pueden encontrar en una red pública descentralizada (interoperabilidad, persistencia, transparencia).
- El derecho al olvido está garantizado (protección).
- Las unidades de gestión de identidad (billeteras digitales) son portables (portabilidad).
- Los proveedores de billeteras digitales no tienen acceso a la información sobre el acceso de los individuos a los servicios o las interacciones con otros (acceso, control, protección).
- Las copias de seguridad garantizan los niveles máximos de seguridad y privacidad (persistencia, protección).
- Las implementaciones cumplen con las políticas regulatorias (protección).”.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Es objetivo general de este proyecto:

Diseñar la identidad autogestionada en la UNRN adoptando el nuevo paradigma de producción de software, basado en tecnologías descentralizadas.

Sus objetivos específicos son:

1. Especificar los requerimientos funcionales utilizando un modelo metodológico y enfoque de construcción.
2. Diseñar una arquitectura de referencia que permita la adopción para la ID autogestionada.
3. Validar el diseño, pasible de ser implementado y transferido al medio.

Las principales actividades a desarrollar son:

A1- Investigar el estado del arte de las tecnologías descentralizadas y el paradigma de la Web 3 para la entrega e implementación de productos de software.

A2- Identificar herramientas, enfoques, metodologías y soluciones innovadoras emergentes en el área.

A3- Analizar el dominio de aplicación, en particular se estudiará el contexto de la Universidad Nacional de Río Negro, y otros gobiernos (municipal, provincial).

A4- Especificar requerimientos funcionales sobre el dominio planteado.

A5- Diseñar la solución en función de los requerimientos funcionales sobre el dominio planteado.

A6- Comunicar los resultados obtenidos.

Tareas a desarrollar:

Estudiar el estado del tema – conducir la revisión de bibliografía de trabajos publicados en congresos nacionales e internacionales y revistas de prestigio a fin de estudiar enfoques de desarrollo, marcos de trabajo, metodologías y soluciones tecnológicas adoptadas para el desarrollo de las tecnologías descentralizadas en el marco de ciudades y regiones inteligentes. (A1, A2)

Estudiar el dominio de aplicación – Conducir un relevamiento de los casos prácticos de desarrollo de soluciones de software con tecnologías descentralizadas y web 3, a fin evaluar las diferentes implementaciones. Comprender a través de un caso de estudio específico como se puede aplicar la tecnología blockchain y la Web 3 revisadas. (A2)

Especificar requerimientos- definir los aspectos funcionales y no funcionales para el diseño de la solución. (A3, A4)

Implementar prototipo(s) de solución – en base a los requerimientos especificados, seleccionar las tecnologías más adecuadas a utilizar para el desarrollo

de las soluciones que mejor resuelvan las necesidades de los distintos actores o roles existentes en el contexto del caso de estudio. En esta tarea, las soluciones se implementarán bajo el nuevo paradigma Web 3 y blockchain. (A4, A5).

Identificar el gap entre investigación y práctica – analizar los desarrollos teóricos que no se aplican en la práctica y/o los desarrollos prácticos que no cuentan con sustento teórico que sean relevantes para el desarrollo de las tecnologías descentralizadas en el marco del desarrollo de ciudades inteligentes en la región, a fin de determinar un espacio de contribución científica y definir requerimientos específicos para las soluciones tecnológicas a proponer(A6)

Validar la implementación y adopción de la web 3 y blockchain - Se validará mediante pruebas de campo a coordinar, como caso testigo pasible de replicarlos en otros contextos. (A5, A6)

Divulgación de Resultados – publicar en congresos nacionales e internacionales y en revistas con referato los resultados de investigación producidos durante el proyecto(A6).

3. RESULTADOS OBTENIDOS/ ESPERADOS

Como resultado de este proyecto, se espera:

- Sensibilización a la comunidad sobre tecnologías descentralizadas.
- Elaboración de material para cursos de grado/posgrado
- Dictado de seminarios y/o cursos para desarrolladores de software.
- Definición de la Arquitectura para el desarrollo de las aplicaciones adoptando las tecnologías descentralizadas.
- Definición y selección de herramientas, enfoques, metodologías y soluciones innovadoras emergentes en el área.
- Se obtendrá mediante un Caso de estudio, una prueba de concepto (PoC, por sus ingles, Proof of Concept) para el diseño de una plataforma de servicios de ID autogestionada.
- Se transferirá la elaboración de este prototipo funcional a la Universidad Nacional de Río Negro.
- Se desarrollará estrategias para ingresar al Comité Nacional de Blockchain.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se encuentra formado por tres investigadores formados, dos investigadores en formación y tres alumnos avanzados de la carrera Licenciatura en Sistemas. En su marco se desarrolla

una tesis de Maestría en Ciencias de la Computación y se producirán tres trabajos finales de carrera de grado.

5. BIBLIOGRAFÍA

- [1] Gaurish Korpai and Drew Scott: Decentralization and web3 technologies. The University of Arizona
- [2] O'Reilly, Tim, What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Communications & Strategies, No. 1, p. 17, First Quarter 2007, Available at SSRN: <https://ssrn.com/abstract=1008839>.
- [3] T. Berners-Lee, "SolidProject origin." disponible en: <https://solidproject.org/origin07>.
- [4] Web 3.0 y blockchain, un cambio de paradigma para hacer negocios con los propios datos personales. Disponible en: <https://www.cronista.com/columnistas/web-3-0-y-blockchain-un-cambio-de-paradigma-para-hacer-negocios-con-los-propios-datos-personales/>
- [5] Liguori, Walter. Web 3 -The Decentralized Future. October 2022 disponible en: DOI: 10.13140/RG.2.2.20599.09129 Practices and Patterns. Addison-Wesley (2001).
- [6] VEGA MAZA, Marina (2019). «El auge de blockchain y sus posibilidades reales de aplicación en los registros de las Administraciones Públicas». IDP. Revista de Internet, Derecho y Política. N.º 28, págs. 109-126. UOC Consultado: 20/12/2022. <http://dx.doi.org/10.7238/idp.v0i28.3154>.
- [7] The web3 Landscape October 2021 disponible en: <https://a16z.com/wp-content/uploads/2021/10/The-web3-ReadIng-List.pdf>.
- [8] Globant. Blockchain un cambio de juego para todas las industrias.Disponible en: <https://reports.globant.com/es/sentinel-report/blockchain/> Consultado el 01-02-2023
- [9] SOVRIN FOUNDATION(2020). disponible en: <https://sovrin.org/wp-content/uploads/Principles-of-SSI-V1.01-Spanish-v01.pdf>. Consultado el 20-02-2023.
- [10] World Bank. (2018). Technical standards for digital identification systems. Obtenido de <http://documents.worldbank.org/curated/en/707151536126464867/pdf/129743-WP-PUBLIC-ID4D-Catalog-of-Technical-Standards.pdf>
- [11] International Organization for Standardization. (2019). IT Security and Privacy — A framework for identity management - Part 1: Terminology and concepts. (ISO/IEC Standard No. 24760-1) Obtenido de <https://www.iso.org/standard/77582.html>
- [12] Christopher Allen. The Path to Self-Sovereign Identity” Disponible en: <http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html>
- [13] Allende Marcos. El futuro de la identidad digital: auto-gestión, billeteras digitales y blockchain. Disponible en: <https://publications.iadb.org/publications/spanish/viewer/Identidad-digital-auto-soberana-El-futuro-de-la-identidad-digital-Auto-soberania-billeteras-digitales-y-blockchain.pdf>
- [14] Allende Marcos. Blockchain:Cómo desarrollar confianza en entornos complejos para generar valor de impacto social. BID disponible en: <https://publications.iadb.org/publications/spanish/document/Blockchain-C%C3%B3mo-desarrollar-confianza-en-entornos-complejos-para-generar-valor-de-impacto-social.pdf>.
- [17] Jolías, L (2022): ""Identidad Digital Descentralizada: una guía de implementación de blockchain en gobierno """. Editorial OS City +. Disponible en <https://plus.os.city/publicaciones/identidad-descentralizada>.
- [18] Chris Dixon (2018). ¿Por qué importa la descentralización? Disponible en: <https://onezero.medium.com/why-decentralization-matters-5e3f79f7638e>.

APLICACIÓN DE GEOFENCING PARA EL CUIDADO DE ADULTOS MAYORES MEDIANTE UN HARDWARE DEDICADO

Pablo M. Vera, Rocío A. Rodríguez, Esteban A. Carnuccio

Centro de Altos Estudios en Tecnología Informática (CAETI)
Facultad de Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{pablomartin.vera, rocioandrea.rodriguez} @uai.edu.ar
{estebanandres.carnuccio} @alumnos.uai.edu.ar

RESUMEN

Los adultos mayores que tienen la posibilidad de tener una vida independiente y manejarse solos, pero con condiciones de salud que complican su movilidad o actividades diarias, pueden contar del apoyo de la tecnología para poder disparar una alarma en forma automática, en caso de que requieran ser asistidos. Esto permite que una o más personas de contacto reciban en sus dispositivos móviles una alerta y puedan acudir para atenderlos. Para ello se considera la realización de un dispositivo electrónico de bajo costo. Este dispositivo permite detectar caídas y también por medio de geolocalización saber si ha salido de una zona segura (se pueden configurar zonas lo que permite detectar que una persona que tenga problemas de orientación se encuentre perdida). Este trabajo introduce también el concepto de IoT (Internet de las Cosas) en el hardware dedicado y también el de tecnologías vestibles (wearable) incorporando el dispositivo en un accesorio de vestir, por ejemplo un cinturón.

Palabras clave: Android, IoT, Geolocalización, Wearable

CONTEXTO

Esta línea de investigación y desarrollo (I+D) forma parte de los proyectos de la Facultad de Tecnología Informática radicados en el Centro de

Altos Estudios en Tecnología Informática (CAETI) sede Ciudad Autónoma de Buenos Aires de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado.

1. INTRODUCCIÓN

Las TIC (Tecnologías de la Información y Comunicación) pueden contribuir a mejorar la calidad de vida de los adultos mayores que desean vivir en la forma más independiente posible, pero con alguna condición de salud que se lo dificulta. El cuidado de adultos mayores puede afectar significativamente su independencia, estado de ánimo, privacidad y calidad de vida [1]. Dado que a veces resulta necesario que este se encuentre acompañado de un tercero para hacer sus actividades diarias, al estar dependiendo de otra persona. Tal como un familiar, allegado o un cuidador [2]. Por ese motivo el índice de Barthel es utilizado para clasificar el nivel de dependencia de una persona, según la dificultad que tenga para hacer determinadas actividades, según [3]. Por otro lado, si el individuo es independiente, puede ser que sus familiares estén intranquilos, ya que no saben las actividades que realiza el adulto mayor, tales como ir al médico, salir a caminar, entre otros. Por lo que puede llegar a ser necesario un control sobre estos [4]. Que una persona tenga independencia para realizar sus actividades diarias, es muy importante para su calidad de vida. Pero al mismo tiempo genera intranquilidad de sus seres queridos, dado que querrán saber que

no tengan ningún problema durante su rutina diaria, de acuerdo a [5]. Así se espera que pueda llevar a cabo su rutinas laborales y sociales sin sentirse observado en todo momento. Al mismo tiempo se procura que una solución basada en las TIC resulte de costo accesible, para ser asequible por diferentes sectores de la sociedad.

La posibilidad de disponer de tecnología que se encuentre de forma ubicua integrada a las propias prendas del usuario da lugar al concepto de wearable. “La palabra wearable posee una raíz inglesa cuya traducción significa “llevable” o “vestible”, en el argot tecnológico hace referencia a pequeñas computadoras que van siempre con el usuario. Bajo esta concepción, el PC deja de ser un dispositivo extraño para el usuario que solo lo usaba en un espacio definido pasando a ser un factor que se incorpora e interactúa de forma continua con él, además de acompañarlo a todas y cada una de las partes” [6].

Para poder sensor los movimientos bruscos que puedan ser indicativos de una posible caída, es necesario contar con sensores en el dispositivo, al igual que para tomar en consideración la posición del usuario o la zona en la que se encuentra. Dando origen a la necesidad de contar con distintos sensores: acelerómetro, barómetro, giroscopio, GPS (Sistema Global de Posicionamiento).

Para establecer zonas seguras se recurre a Geofencing que permite definir un área geográfica reducida que se considera como zona de referencia para activar un evento en el momento que un usuario ingresa o sale de esa zona. Estas zonas se denominan geovallas [7]. Cada geovalla está conformada por un perímetro virtual que permite monitorear a un objetivo concreto con distintos fines por ejemplo ofrecer servicios o enviar notificaciones. Ese perímetro puede ser una simple línea (calle), un círculo (radio desde un punto de interés) ó tomar formas más complejas mediante polígonos [8].

Podría utilizarse un smartphone como hardware para esta solución a implementar, dado que ya tiene estos sensores disponibles (de hecho, un trabajo previo del grupo de investigación ha

tomado como dispositivo el teléfono celular para escolares de la ciudad autónoma de Buenos Aires [9]). Pero la solución debe enfocarse al tipo de usuario, y para las personas adultas mayores no es habitual estar prácticamente todo el tiempo con el teléfono a cuestas, como si sucede con las nuevas generaciones. Es por ello, que se optó por montar la solución en un hardware dedicado, que será el encargado de monitorear al adulto mayor (ver figura 1).



Figura 1. Hardware dedicado para el adulto mayor y Aplicación Móvil para las personas que lo asisten (configuración y alertas)

Luego se requiere del desarrollo de un software que permitirá realizar las configuraciones sobre lo que sea requerido monitorizar según las necesidades médicas del adulto en cuestión, ese mismo software será el que recepcione las alertas disparadas por el hardware dedicado mediante IoT. “IoT es considerado como la conexión en red de objetos físicos o dispositivos en una red abierta y llena de objetos inteligentes que tiene la capacidad de auto gestionarse, compartir información, datos y recursos, reaccionar y actuar frente a situaciones y cambios en el medio ambiente” [10]. Este software se planificó realizarlo para Android (con Android Studio, lenguaje Java), se elige este sistema operativo por ser el que más inserción tiene en el mercado. De este modo los familiares o personas cercanas pueden encargarse de configurar la aplicación.

“Android domina el mercado global, el 70% de los smartphones lo usan” [11].

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto de investigación busca:

- Hacer uso de la geolocalización y de las técnicas de geofencing para aplicarlos a aplicaciones móviles.
- Diseñar una aplicación de monitoreo no invasiva en Android
- Desarrollo de un hardware dedicado de costo accesible para la sociedad

3. RESULTADOS OBTENIDOS/ESPERADOS

Parte de hardware:

Se espera obtener un prototipo inicial que cumpla con 2 parámetros:

- Costo accesible: De forma de ser competitivo con otras soluciones del mercado
- Tamaño pequeño: Incorporado a accesorios de uso cotidiano, por ejemplo, un cinturón

El hardware deberá tener conectividad a internet y dar respuesta de forma rápido, esto será probado con el hardware ya construido.

Parte de software:

Se espera para la parte de software aprovechar el conocimiento de Geofencing aportado por trabajos previos en el área [12]. El sistema incluirá principalmente:

- Configuración de parámetros del usuario: En esta configuración se podrá indicar que se habilitará para el monitoreo. Así como los contactos que recibirán las alertas pudiendo establecerse para cada alerta usuarios distintos.

- Zonas seguras: Lugar de vivienda, Lugares visitados frecuentemente (pudiendo establecerse días y horarios). Esto se resuelve mediante geofencing.
- Restricciones: Eventos a detectar como por ejemplo: la persona no puede subir ó bajar escaleras.

Como ha comenzado en el presente año este prototipo por medio de un hardware dedicado aún no hay resultados obtenidos, se espera poder avanzar con el prototipo y realizar las pruebas para evaluar que cumpla con todas las características planteadas.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 5 personas, docentes de grado, postgrado y alumnos.

En el área de dispositivos móviles se encuentran en realización, 2 tesis de maestría y 1 tesina de grado en la UAI (Universidad Abierta Interamericana).

5. BIBLIOGRAFÍA

- [1] Ardila, A., Orejana, D., & Silva, L. (2018). Relación del nivel de dependencia y la depresión en los adultos mayores de un hogar vida de Florida Blanca, Santander en el año. Universidad Cooperativa de Colombia.
- [2] Rosa, R. V. (2014). Manual de cuidados generales para el adulto mayor disfuncional o dependiente (Literatura). México.
- [3] Pinto, D., Alarcón, S., Rosero, A., Zomora, J., & Landázuri, M. (2016). Actividad diaria según índice de Barthel en adultos mayores. Maldonado LMV/et Al/Enfermería Investiga, Investigación, Vinculación, Docencia y Gestión, 1(3).
- [4] Serrano, J. (2013). El Papel De La Familia En El Envejecimiento Activo. In The Family Watch. <https://fiapam.org/wp-content/uploads/2013/06/Informe20132.pdf>

- [5] Celeiro, T. M., & Galizzi, M. P. (2019). Calidad de vida en adultos mayores entre 70 a 85 años institucionalizados y no institucionalizados de la ciudad de Nogoyá. Universidad Católica Argentina.
- [6] Vera, P. M., Rodríguez, R. A., & Kaimakamian Carrau, M. (2020). Tecnologías vestibles aplicadas al cuidado de la salud: Construcción de un prototipo de monitoreo. In XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz).
- [7] K. Zuva, and T. Zuva, "Tracking of Customers using Geofencing Technology". 2019
- [8] Vera P., Rodríguez R., Viavattene H., y Martínez M. Diseño y desarrollo de aplicaciones móviles basadas en geofencing. In XXII Workshop de Investigadores en Ciencias de la Computación (WICC, El Calafate, Santa Cruz), 2020
- [9] Buenos Aires Ciudad, Buscador de Senderos Escolares. Disponible en:
<https://www.buenosaires.gob.ar/justiciayseguridad/senderos-escolares>
- [10] Pérez, M. R., Mendoza, M. A., & Suarez, M. J. (2019). Paradigma IoT: desde su conceptualización hacia su aplicación en la agricultura. Paradigma, 40(18), 1-8.
- [11] Curry David, Android Statistics (2023)
<https://www.businessofapps.com/data/android-statistics/>
- [12] P. Vera, R. Rodríguez, and C. Delgado, "Geofences Application Develoment for assisting people through monitoring", LAJC, vol. 9, no. 1, pp. 98-107, Jan. 2022.

Programa integrador de proyectos de Software con transferencia a la comunidad

Mg. Rodolfo Bertone¹ C.C. Marcelo E. Taruschio²
Lic. Martin Correa³ Lic. Diego Luna⁴ Lic. Franco Quintero⁵

Facultad de Ciencias Exactas e Ingeniería. Universidad Católica de La Plata.

¹ rodolfo.bertone@ucalp.edu.ar ² mtarus2012@gmail.com ³ martin.correa@ucalp.edu.ar
⁴ diegomartinluna@gmail.com ⁵ fpquintero@gmail.com

Resumen

El concepto “Usina de Ideas” define un espacio virtual y/o físico donde se generan, evalúan, proponen, discuten, acrecientan ideas de toda la comunidad de la facultad y/o universidad. Esta definición abarca las actividades que el grupo de investigación y transferencia de FACEI se propuso realizar desde la presentación inicial en WICC 2022. La actual presentación es una evolución de dicho proyecto. Respecto del mencionado proyecto aparecen algunas diferencias de cómo estructurar al mismo dentro del contexto de nuestra universidad y se describen, además, las tareas realizadas, las nuevas propuestas y los avances obtenidos.

Para adecuar los trabajos de investigación que se realizaron y se están realizando, fue necesario definir un programa que permita definir el marco de cada uno de los proyectos que posteriormente se fueron presentando. Esta definición se realizó para poder enmarcar las actividades dentro del contexto de trabajo de la UCALP.

Este proyecto tenía originalmente como finalidad primaria la gestión de alternativas de trabajo para el desarrollo profesional de los alumnos en el marco de sus tesinas de grado. Esta finalidad no fue alterada, pero asociado a la misma se generaron actividades de investigación desarrolladas

por profesores de la casa y que involucran alumnos, como participantes de la misma. El proyecto en cuestión se denomina “Control y monitoreo de servicios públicos municipales” y representa un nuevo emprendimiento que se liga a las actividades del presente programa y que será descrito en esta presentación.

Palabras clave:

Gestión de proyectos de software. Automatización y Control. Internet de las cosas. Gobernanza electrónica

Contexto

Esta línea de Investigación continua la presentada 2022 y se muestran resultados obtenidos y nuevas propuestas de trabajo. Hasta la definición de este programa toda la tarea de investigación se desarrolla ad-hoc a partir de gestión de docentes y alumnos de esta facultad. La formalización viene como parte de un plan rector de nuestra Universidad donde se continúan volcando mayores recursos a la investigación en general y a la transferencia de conocimiento en particular.

El contexto informático actual representa un enorme desafío que entre

otras características busca incentivar a los alumnos para realizar actividades de investigación con transferencia. Es conocida la enorme demanda laboral, pero también es ampliamente sabido que es importante que se formen nuevos investigadores en la disciplina. Con este tipo de proyectos se busca incentivar en el desarrollo de aplicaciones novedosas donde se genere un grupo multidisciplinario que permita realizar actividades de investigación.

Introducción

Dentro de esta presentación se engloban una serie de subproyectos cuyo contexto de investigación está basado principalmente en analizar y discutir soluciones específicas a problemas planteados por organizaciones públicas privadas o mixtas, como así también inquietudes de la comunidad que son acercadas a nuestro ámbito de trabajo.

Como se consignara en la presentación del año anterior, desde 2016 se han presentado problemas relacionados con la Licenciatura en Sistemas y cuyo desarrollo necesitó de investigación, análisis e implementación de soluciones concretas. En varios de estos casos se obtuvieron soluciones innovadoras que generaron transferencia a la comunidad. Además, los resultados fueron presentados, aprobados y expuestos en congresos de la especialidad.

Líneas de Investigación. Objetivos.

El objetivo general del programa es generar un entorno que permita definir subproyectos con transferencia al medio y que resuelvan problemas específicos, donde el software sea el elemento constitutivo esencial (y que, eventualmente podrían incluir desarrollos específicos de hardware

utilizando varias tecnologías, como por ejemplo Arduino).

Los Objetivos Específicos son: 1) Establecer reuniones de brainstorming donde participen alumnos y docentes, 2) Generar casos de estudio donde se analice la viabilidad de los subproyectos presentados, 3) Presentar el sub-proyecto de acuerdo a la modalidad definida en la normativa FACEI, 4) Desarrollar el sub-proyecto y generar informes de avances, informes técnicos, monografías y/o publicaciones para congresos.

La metodología de trabajo involucra diversas actividades, que pueden ser interdisciplinarias y serán desarrolladas durante 2023, estableciendo el plazo del programa en 10 meses a partir del primero de marzo.

Cada subproyecto tiene su planificación temporal y su plazo de desarrollo que se establece en cuatro meses. Es posible que simultáneamente se tengan varios subproyectos en ejecución y los mismos sean evaluados por un grupo de expertos que emiten su opinión vinculante respecto de la validez del subproyecto presentado.

Cada subproyecto tendrá un director, el cual lo pondrá a consideración de los participantes para su aprobación. En caso de tratarse de proyectos definidos como Tesinas de Grado, deberá contar con el aval del director de carreras de sistemas.

Desarrollo e Innovación

En la presentación 2022 se definió el subproyecto AIS Signal Detector. El mismo fue pensado, desarrollado y presentado durante dicho año. Los resultados obtenidos fueron presentados en el Congreso CACIC de 2022 realizado en La Rioja. [1]

Un resumen de dicho proyecto puede ser definido como que, en el ámbito portuario y de navegación, la tecnología de sistemas de identificación automática ha mejorado notablemente a la navegación como seguridad, eficiencia, protección del ambiente e identificación de buques. Un Sistema de Identificación Automática (AIS, por su sigla en inglés) se utiliza en todos los buques, puertos y artefactos navales (ej. las boyas). Cada puerto tiene asociado un conjunto de artefactos navales que sirven para realizar marcas específicas que ayudan a la navegación. El sistema desarrollado permite analizar el recorrido de los buques y poder determinar si alguno de ellos por fallos, errores y omisiones, colisionan con los artefactos disponibles. Es factible que acciones como estas ocurran, por lo tanto, contar con sistemas de monitoreo que permite detectar situaciones anómalas y demostrar el hecho ocurrido es muy importante en particular debido al alto costo de reemplazo de los elementos electrónicos dañados. El sistema se encuentra operativo en el puerto de la ciudad de La Plata, lugar de trabajo del desarrollador del problema. Asimismo, durante 2023 se le están realizando mejoras y ampliaciones.

Actualmente dentro del marco de trabajo de esta presentación se están desarrollando dos subproyectos, en tanto que un tercero está en fase de definición. El primero de ellos se denomina Bot de Trading, en tanto que el segundo, como se indicó en el resumen, se denomina “Control y monitoreo de servicios públicos municipales”. Ambos subproyectos están presentados y serán brevemente descriptos a continuación. La tercera actividad en fase de definición y presentación tiene que ver con el desarrollo de un prototipo de aplicación para la renta temporaria de espacios de coworking.

El proyecto “Bot de Trading” tiene

como objetivo la automatización de las operaciones que se realizan en el mercado de las criptomonedas. Para lograrlo, se piensa en un sistema que sea capaz de obtener datos del mercado, procesarlos y luego, calcular ciertos indicadores para permitir realizar un análisis del contexto y ejecutar operaciones según corresponda. Esto es posible gracias a la comunicación que se lleva a cabo entre la aplicación y la Api de Exchange que se utiliza para acceder al mercado. [2][3][4][5][6]

Se piensa desarrollar, además, un segundo Bot, que actúe como complemento para el usuario y le ofrezca diferentes opciones de consulta a quienes lo utilicen, como por ejemplo un balance de cuenta con la que se está operando. Se espera presentar los resultados del mismo en el congreso CACIC 2023.

El segundo subproyecto tiene que ver con el control y monitoreo de los servicios municipales. Los municipios de la provincia de Buenos Aires en particular y de la República Argentina, en general, no cuentan con un servicio de control y monitoreo de sus actividades básicas. Como actividades básicas, se contempla el seguimiento y reporte de problemas asociados con la dinámica municipal: control de estado de luminarias, control de estado de espacios públicos, control de estado de semáforos y redes viales, por ejemplo. Cada uno de estos puntos son ejemplos posibles donde a partir del uso de aplicaciones en dispositivos móviles pueden servir para mejorar la interacción entre los vecinos y el municipio, posibilitando así la incorporación de los habitantes al control municipal, permitiendo administrar mayor información, de manera más precisa y eficiente.

Hay aplicaciones disponibles para teléfonos celulares, como por ejemplo

“Cuando llega mi micro”, ¿“Cuando llega?”, “Como llego”, “Tu bondi” o “Cuando Subo” entre otras, que pone a disposición de los usuarios la información en vivo recopilada por el sistema SUBE y el ministerio de Transporte de la Nación, sobre unas 294 líneas de colectivos con 18 mil vehículos monitoreados que circulan por AMBA [7]. Al elegir una parada se muestra una tabla con los colectivos que están en camino hacia ella y cuantos minutos tardarán en llegar. Asimismo, se puede ver el desplazamiento de los mismos en un mapa.

Esta misma acción puede ocurrir con el sistema de recolección de residuos. En la actualidad cada municipio establece un rango horario para sacar la basura a la vía pública. Si se pudiera administrar el recorrido de los recolectores e informar su ubicación en tiempo real, el vecino podría sacar su basura instantes antes de la pasada del camión recolector, o, en su defecto, no sacarla debido a que el mismo ya paso ese día. Como corolario, es posible, además, elevar quejas por situaciones, como, por ejemplo, el camión paso y mi basura no fue recogida o la recolección se hizo de manera inadecuada dejando bolsas tiradas en la vía pública.

Existen varias funcionalidades que se pueden agregar a una aplicación desarrollada para un municipio. Como parte del trabajo propuesto se pretende investigar el estado del arte y generar una solución simple y de bajo costo que consiga soluciones concretas para el vecino y que para el municipio sea importante incorporarla para brindar mejores servicios.

Resultados y Objetivos

Los resultados obtenidos hasta el momento se plasmaron en una publicación

en CACIC 2022 [1]. La misma fue presentada y defendida en el congreso y se obtuvo una muy buena experiencia respecto de los resultados.

Finalmente, el tercer proyecto del programa busca ser un prototipo de aplicación para facilitar la renta temporaria de espacios de trabajo a distancia en oficinas de coworking, el mismo esta como se menciono en etapa de definición.

Formación de Recursos Humanos

Los proyectos presentados tienen un director del mismo y en cada uno de ellos al menos hay un alumno incorporado al mismo

La formación de recursos humanos es una actividad fundamental del proyecto. Cada subproyecto generado tiene y tendrá la participación de alumnos en su fase de terminación de carrera.

Asimismo, profesores de la casa están trabajando en sus postgrados.

Referencias

- [1] AIS-Signal Detector. Control de balizas de acceso a puertos. Martin Andres Cachile, Facundo Ferro, Marcelo Taruschio, Rodolfo Bertone. Anales Cacic 2023, pág. 536 a 545. ISBN 978-978-1364-31-2
- [2] <https://www.coinbase.com>
- [3] <https://bitcoin.org>
- [4] <https://guiadetrading.com>
- [5] <https://python-binance.readthedocs.io>
- [6] Caballero Belasco, MA y Cilleros Serrano Ciberseguridad y transformación digital (2020) Ediciones Anaya multimedia
- [7] https://www.clarin.com-servicios-6-mejopres-apps-saber-pasara-colectivo-datos-tiempo-real_0_mIVcUtsVb.html

Innovación en las Ciudades mediante las TIC

Análisis de implementaciones en Ciudades Inteligentes

Rocío A. Rodríguez¹, Pablo M. Vera¹, Dora Mendoza^{1,2}, Mariano G. Dogliotti¹, Claudia G. Alderete¹

¹Centro de Altos Estudios en Tecnología Informática (CAETI)
Facultad de Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

²Laboratorio de Investigación y Desarrollo Tecnológico en Energías Renovables. Sede Regional
Orán. Universidad Nacional de Salta, Argentina

{rocioandrea.rodriguez, pablomartin.vera, mariano.dogliotti, claudia.alderete}@uai.edu.ar
dmendoza412@oran.unsa.edu.ar

RESUMEN

Actualmente se necesita disponer de la implementación de las TIC (Tecnologías de la Información y Comunicación) en las urbes para beneficio de la ciudadanía. Las soluciones tecnológicas pueden contribuir a mejorar distintos aspectos, por ejemplo: la estructura, el reciclado, la movilidad. Todas estas mejoras contribuyen de forma directa a la calidad de vida de la ciudadanía. Las ciudades necesitan ser sustentables y en esa dirección se construye el concepto de Ciudades Inteligentes (Smart Cities). En este artículo se presenta un proyecto de investigación y desarrollo, cuyo objetivo es analizar cómo se hace uso de la tecnología con el fin de crear soluciones para ciudades inteligentes.

Palabras clave: Ciudades Inteligentes, TIC, Sustentabilidad

CONTEXTO

Esta línea de investigación y desarrollo (I+D) forma parte de los proyectos radicados en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). En este proyecto participan docentes y alumnos, tanto de sede Centro como de la Castelar (ambas en la provincia de Buenos Aires). El proyecto

cuenta con financiamiento asignado y una duración de 2 años.

1. INTRODUCCIÓN

La implementación de las TIC (Tecnologías de la Información y Comunicación) es indispensable para la sostenibilidad en las grandes ciudades, que se ven impactadas por la sobrepoblación. En la región de América Latina más del 80% de la población está viviendo en áreas urbanizadas [1]. “Los usos y potencialidades de las TIC, hoy, son un elemento fundamental en la sociedad, sin embargo, el reto de estas tecnologías sigue siendo que se conviertan en instrumentos de cambio y transformación que garanticen mejores niveles de vida, equidad social y cuidado al medio ambiente. Entre los aspectos centrales de las ciudades es que ofrecen espacios generadores de oportunidades, pero al mismo tiempo, es en ellas donde se concentran y magnifican diversos problemas, en este sentido, es que las TIC podrían jugar un papel central para afrontar los diversos desafíos y a la vez aprovechar o mejor aún potencializar sus beneficios” [2].

El desafío es emplear las TIC para conseguir una ciudad inteligente. “Una Ciudad Inteligente es aquella que coloca a las personas en el centro del desarrollo, incorpora

Tecnologías de la Información y Comunicación en la gestión urbana y usa estos elementos como herramientas para estimular la formación de un gobierno eficiente que incluya procesos de planificación colaborativa y participación ciudadana. Al promover un desarrollo integrado y sostenible, las Smart Cities se tornan más innovadoras, competitivas, atractivas y resilientes, mejorando así las vidas” [3].

“Para que una ciudad sea inteligente, se requiere del uso integrado de diversas Tecnologías de la Información y las Comunicaciones, dentro de las cuales se destacan las redes IP móviles, clouding, Big Data y el Internet de las Cosas” [4]. “No obstante, las ciudades no son sólo las que aplican nuevas tecnologías para mejorar los sistemas de gestión urbana sino aquellas que fomentan una economía ecológica, movilidad sostenible, un medio ambiente sano, modos de comportamiento saludables y una mayor calidad de vida en favor de la sostenibilidad urbana en su conjunto” [5].

Si bien los gobiernos locales son los que ofrecen la mayor cantidad de servicios, los cuales con el tiempo se fueron actualizando y desplegándose en distintos canales con el fin que sean accesibles para la ciudadanía [6] es indispensable que los gobiernos nacionales lo consideren como línea de acción a nivel país. Los gobiernos nacionales impulsan por medio de líneas de acción, marcos y modelos, así como regulación por leyes la concreción de mejoras en el cuidado del medio ambiente, la infraestructura y la sostenibilidad de las ciudades, lo que impacta directamente en la calidad de vida de los ciudadanos. Distintas terminologías recibieron las ciudades que han implementado las TIC: “Ciudades Digitales (Digital Cities), Ciudades Inteligentes (Intelligent Cities), y Ciudades Inteligentes Sostenibles (Smart Cities), que se incorporan de acuerdo con el grado y la naturaleza de la capacidad de la tecnología digital de la ciudad” [7].

“Las Ciudades Digitales integran la tecnología digital en los sistemas básicos de infraestructura de la ciudad, mientras que las Ciudades Inteligentes se basan en la infraestructura de Ciudad Digital para construir edificios inteligentes, sistemas de transporte, escuelas, espacios y servicios públicos, integrados en sistemas urbanos inteligentes. Las Ciudades Inteligentes Sostenibles despliegan sistemas urbanos inteligentes que sirven para el desarrollo socio-económico y ecológico, para mejorar la calidad de vida de sus habitantes y para abordar los orígenes de la inestabilidad social en las ciudades. El modelo más evolucionado es el de Ciudad Inteligente Sostenible” [7].

En Argentina, la Secretaría de Modernización de la Nación ha establecido un modelo de ciudad inteligente [8], el cual está formado por 5 dimensiones, las cuales a su vez tienen dentro ejes de interés: (1) Gobernanza: Infraestructura y Capacidades; Plataforma de Servicios; Gobierno Abierto; (2) Ambiente: Calidad Ambiental; Gestión de Recursos; Ecología Urbana y Resiliencia; (3) Desarrollo Humano: Sociedad; Seguridad; Salud; Educación; (4) Planeamiento Urbano: Transporte; Movilidad; Espacio Urbano; (5) Competitividad: Innovación; Productividad.

Existen modelos de Ciudades Inteligentes en distintos países, esto impulsa fuertemente como política de estado, el mejoramiento de las ciudades y fortalece tanto los servicios como la infraestructura que ponen a disposición los gobiernos locales.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Analizar las implementaciones existentes para Smart Cities (SC) y clasificarlas
- Comparar los planes de acción de los gobiernos nacionales de América Latina con respecto a las SC

- Analizar formas de evaluar el grado en que las ciudades han implementado características de las SC
- Desarrollar estrategias e implementar recursos innovadores que permitan mejorar distintos aspectos propios de las SC

3. RESULTADOS OBTENIDOS/ESPERADOS

Se ha realizado un mapeo sistemático de la literatura, en el cual se pudo evidenciar distintas implementaciones para las SC en la dimensión ambiental, entre los cuales se destacan publicaciones de implementaciones en los siguientes países:

- Argentina (Bahía Blanca, Buenos Aires, Resistencia (Chaco), San Martín (Mendoza), Río Negro) [9], [10], [11], [12], [13]
- Bélgica (ciudad: Ghent) [14]
- España (con una Red de Ciudades Inteligentes compuesta por 75 ciudades) [15], [16], [17], [18]
- Israel (Tel Aviv, su ciudad más poblada) [19]
- India (ciudades: Coimbatore, Gift City y Lavasa) [20], [21]
- México (ciudad: Nezahualcóyotl) [22]

Actualmente nos encontramos comparando los modelos de Ciudades Inteligentes propuestos a nivel nacional en los distintos países de habla hispana. Con el fin de analizar las innovaciones realizadas por medio de recursos tecnológicos y el impacto de estas en la sociedad.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está formado por 5 docentes, 2 de ellos doctores en Ciencias Informática graduados en la Universidad Nacional de La Plata (UNLP), 1 realizando el Doctorado en Informática en la Universidad Abierta

Interamericana (UAI) y 1 realizando una maestría (UAI). Esto implica que 4 de los 5 docentes que componen el grupo tienen estudios de posgrados finalizados o en progreso.

Este proyecto también cuenta con la participación de alumnos de grado y posgrado de la UAI que no están vinculados con actividades docentes (actualmente en el proyecto colaboran 3 alumnos).

En esta área encuentran en realización una tesis de doctorado (UNLP) y una tesina de grado (UAI), siendo directores de dichas tesis miembros del equipo de investigación.

5. BIBLIOGRAFIA

- [1] Naciones Unidas, Departamento de Asuntos Económicos y Sociales, División de Población (2014). World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)
- [2] Alvarado-López, R. A. (2020). Ciudades inteligentes y sostenibles: una medición a cinco ciudades de México. Estudios sociales. Revista de alimentación contemporánea y desarrollo regional, 30(55).
- [3] Bouskela, M., Casseb, M., Bassi, S., De Luca, C., & Facchina, M. (2016). La ruta hacia las smart cities: Migrando de una gestión tradicional a la ciudad inteligente (Vol. 454). Inter-American Development Bank.
- [4] Zona-Ortiz, A. T., Fajardo-Toro, C. H., & Pirachicán, C. M. A. (2020). Propuesta de un marco general para el despliegue de ciudades inteligentes apoyado en el desarrollo de Iot en Colombia. Revista Ibérica de Sistemas e Tecnologías de Informação, (E28), 894-907.
- [5] Jimenez Herrero L. (2016). Hacia ciudades y territorios inteligentes, resilientes y sostenibles. Gestión y gobernanza para la gran transición urbana.

Asociación para la Sostenibilidad y el Progreso de las Sociedades (ASYPS)

Investigadores en Ciencias de la Computación (Salta, 2015).

- [6] Rodríguez, R., Vera, P., Marko, I., Alderete C. & Conca A (2015). El gobierno electrónico y la implementación de las TIC para brindar nuevos canales de comunicación. *Revista Latinoamericana de Ingeniería de Software*, 3(5), 187-196.
- [7] Estévez, E. C., & Janowski, T. (2016). Gobierno digital, ciudadanos y ciudades inteligentes. *Bit & Byte*, 2.
- [8] Castiella L., Palacio Catalina & Rueda M. (2016). La Importancia de un Modelo de Planificación Estratégica para el Desarrollo de Ciudades Inteligentes – Secretaria de Modernización, Presidencia de la Nación.
- [9] Macri & Ibarra A. (2015). Buenos Aires para Todos los Argentinos. Buenos Aires. Ciudad inteligente que construye futuro. Ed. Galt SA. Buenos Aires, Argentina.
- [10] Quartucci E., Wirsky S. y Mazzuchelli A. (2020). Bahía Blanca como Ciudad Inteligente. XIV Simposio de Informática en el Estado (SIE 2020) - JAIIO 49. Páginas: 44-57.
- [11] Piasentini C.I., Statkiewicz, J., Bolatti J., Scappini R. J. y Gramajo S.D. (2019). Smartcities con LoRaWAN: el caso de monitoreo de condiciones ambientales de lagunas en tiempo real. XXV Congreso Argentino de Ciencias de la Computación (CACIC). ISBN: 978-987-688-377-1. Páginas: 910-918.
- [12] Mercado G., Da Peña J.M., Stasi R., López G., Burlot A., Vivone G.C., Amstutz C., Barnabo L., De Paolo J., Ledda M., Caceres R., Taffernaberry C., Pérez S., Alvarez L., Fernandez J., Arena A.P (2015), SG-SM – Smart Grid San Martín. Red de Distribución y Generación de Energía Inteligente en ciudad Gral. San Martín – Mendoza. XVII Workshop de
- [13] Difabio L.A., Vivas H. L. & Muñoz Abbate H. (2016). Internet de las Cosas aplicada a la trazabilidad de la recolección de residuos en Ciudades Inteligentes. SIE 2016, 10º Simposio de Informática en el Estado. 45 JAIIO. ISSN: 2451-7534.
- [14] Wijnants M., Lamotte W., Letor N., Blondia C., De Poorter E., Naudts D., Verstichel S., Lannoo B., Moerman I., Matthys N., Huygens C. (2012). An Eco-friendly Hybrid Urban Computing Network Combining Community-based Wireless LAN Access and Wireless Sensor Networking. IEEE International Conference on Green Computing and Communications, Conference on Internet of Things, and Conference on Cyber, Physical and Social Computing.
- [15] González A., Villazón-Terrazas B. y Gómez J.M. (2014). A Linked Data Lifecycle for Smart Cities in Spain.
- [16] De la Serna I. (2016). Medio ambiente y tecnología se dan de la mano en la Smart City. Secretaría General Técnica, Ministerio de Agricultura, Alimentación y Medio Ambiente. *Revista Ambienta*, N° 115. ISSN 1577-9491.
- [17] Pérez-González D. y Díaz-Díaz R. (2015). Public Services Provided with ICT in the Smart City Environment: The Case of Spanish Cities. *Journal of Universal Computer Science*, vol. 21, no. 2 (2015), 248-267.
- [18] Barroso S., Sánchez A., Núñez P., García P. & Bustos P. (2019). Smartpolitech: Estimación del consumo de agua basado en modelo de mezclas de Gaussianas. *Actas del X International Greencities Congress 2019: 10º Foro de Inteligencia y Sostenibilidad Urbana*. ISBN 978-84-09-09960-3, págs. 99-110.

- [19] Toch E. & Feder E. (2016). Estudios de casos internacionales de ciudades inteligentes. Investigación conjunta BID (Banco Interamericano de Desarrollo) – KRIHS (Korea Research Institute for Human Settlements). Tel Aviv, Israel. Documento para discusión N° IDB-DP-444.
- [20] Tiwari A. y Jain K. (2014). GIS Steering Smart Future for Smart Indian Cities. International Journal of Scientific and Research Publications, pp 4-8. ISSN 2250-3153.
- [21] Mohammed y Bagavathi (2016). Framework for a Smart Water Management System in the Context of Smart City Initiatives in India. Procedia Computer Science, Vol.92, Año 2016, pp. 142-147.
- [22] Hernández Mar R., Morales Calderón J. R., Rózga Luter R.E. (2021). Gestión y recolección de residuos sólidos urbanos (RSU) desde la perspectiva de la Ciudad Inteligente (CI): el caso de recolección de basura en el municipio de Nezahualcóyotl, Estado de México. Repositorio universitario de la Universidad Nacional Autónoma de México.

APLICACIÓN DE TÉCNICAS DE INTELIGENCIA ARTIFICIAL EN LA PREVENCIÓN DE PROBLEMAS DE SALUD MENTAL

Di Felice, M.¹; Lamas, C.¹; Maleh, F.¹; Sabelli, A.¹; Norscini, J.¹; Ramón, H.¹; Deroche, A.¹; Montenegro Aguilar, G.¹; Pytel, P.¹; Vegega, C.¹; Chatterjee, P.¹; Pollo-Cattaneo, Maria F.¹

¹Grupo GEMIS - Programa Maestría en Ingeniería en Sistemas de Información. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.

Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires

cvegega@frba.utn.edu.ar, clamas@frba.utn.edu.ar, flo.pollo@gmail.com

Resumen

La Organización Mundial de la Salud define a la salud mental como un “estado de bienestar en el que las personas son conscientes de sus propias capacidades, pueden hacer frente a las tensiones normales de la vida, pueden trabajar de manera productiva y fructífera, y son capaces de contribuir a su comunidad”. Los trastornos mentales y los trastornos relacionados con sustancias psicoactivas son muy prevalentes en todo el mundo y contribuyen de manera importante a la morbilidad, la discapacidad y la mortalidad prematura. Por otro lado, en los avances recientes de la medicina moderna, la Inteligencia Artificial (IA) se destaca como uno de los principales actores, principalmente para la atención médica predictiva y preventiva. En adición, con el advenimiento de los enfoques digitales para la salud mental, la IA moderna, en particular el aprendizaje automático, se utiliza en el desarrollo de soluciones de predicción, detección y tratamiento para la atención de la salud mental. Aunque ha habido un progreso considerable en la salud digital y la aplicación de la IA a la salud física en general, la adopción de la IA en la salud mental es relativamente incipiente. Durante la pandemia de COVID-19, los casos reportados sobre

trastornos de ansiedad y depresión han aumentado un 74% según una encuesta de la Asociación Estadounidense de Psicología. Específicamente en América Latina, la brecha de tratamiento para la depresión es del 73,9%. Una parte importante de los trastornos de salud mental es prevenible si se detecta a tiempo y se trata adecuadamente. Este trabajo apunta precisamente a la incorporación de la IA en la salud mental, hacia una atención médica predictiva, preventiva, personalizada y de precisión. El objetivo principal es desarrollar algoritmos predictivos para el diagnóstico temprano de trastornos de salud mental, incluida la depresión, seguido del uso de aprendizaje automático para realizar análisis predictivos de los datos clínicos.

Palabras clave: Inteligencia Artificial, Sistemas Inteligentes, Aprendizaje Automático, salud mental, análisis predictivo

Contexto

La Universidad Tecnológica Nacional ha articulado, a través de los años, diversos Programas de Investigación, Desarrollo e Innovación, incluyendo entre ellos el Programa de Sistemas de Información e

Informática. Éste cuenta como objetivo "intensificar y focalizar las acciones tendientes a fortalecer y promover el crecimiento de las temáticas de investigación en Sistemas de Información e Informática, y promover la interacción con la industria informática en general y la industria de sistemas de información y desarrollo de software en particular" [1].

De la misma manera, en el Departamento de Ingeniería en Sistemas de Información de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional (UTN-FRBA), surge, en el año 2009, el Grupo GEMIS. Este equipo está conformado por un plantel de docentes, graduados y estudiantes interesados en la sistematización del conocimiento y su promoción en el campo de la Ingeniería del Software y la Ingeniería en Sistemas de Información, comprendiendo sus aplicaciones y enfoques metodológicos en múltiples escenarios. En este contexto, se vienen desarrollando de manera ininterrumpida actividades científicas articulando carreras de grado y posgrado afines a la disciplina, con la convicción que los resultados obtenidos mejoran la formación del plantel docente y el alumnado (despertando en éste vocaciones nóveles en el ámbito científico).

1. Introducción

La Inteligencia Artificial (IA) es una rama de las ciencias de la computación que se encarga de resolver problemas complejos, no lineales, que usualmente necesitan la interacción humana. La IA, así, busca emular el comportamiento humano para poder automatizar tareas de forma de resolverlas con una eficiencia similar pero de forma más veloz [2].

Para poder emular ese comportamiento humano los algoritmos basados en IA

utilizan grandes conjuntos de datos, llamados datasets [3]. Cuanto más grandes, completos y heterogéneos son estos datasets, estos algoritmos pueden inferir de mejor manera las relaciones entre estos datos y así generar reglas que ante la aparición de nuevos datos puedan predecir cómo se van a comportar los mismos [4].

Por otro lado, debido al incremento del poder computacional y a la disponibilidad de cada vez más datos, la IA ha aumentado su participación en distintos campos durante los últimos años [5,6]. Uno de los campos donde cada vez tiene mayor participación es el campo de la salud [7]. También ha incursionado en el campo de la salud mental, aunque a un ritmo menor [8]. Si bien los aspectos éticos de su uso aún están en debate [9], los beneficios que devienen de su aplicación parecen prometedores [7,10], entre los cuales se destacan la velocidad a la hora de realizar un diagnóstico y el hecho de eliminar la subjetividad del experto cambiándola por un método objetivo con base científica.

Aunque ya existen técnicas objetivas y parametrizadas para el diagnóstico de las enfermedades mentales, y en particular de la depresión [11-17], el uso de los métodos de IA no sólo ofrece el proceso de realizar el diagnóstico transformando los datos correspondientes a síntomas en una salida correspondiente a una enfermedad, sino que también puede ayudar a encontrar esos síntomas, transformando expresiones coloquiales en síntomas objetivos, como así también descubrir relaciones entre los distintos tipos de síntomas.

Se sabe que, la salud mental es un componente fundamental para el individuo y la sociedad en la que habita ya que le permite pensar, manifestar sentimientos, interactuar con los demás, ganar el sustento y disfrutar de la vida. Por lo cual se considera que la promoción, la protección y el restablecimiento de la

salud mental son preocupaciones vitales de las personas, las comunidades y las sociedades de todo el mundo [18].

A partir de ello, a pesar de la cautela que prevalece en la medicina, incluida la psiquiatría, respecto al uso de la Inteligencia Artificial [19, 20], las aplicaciones de IA en medicina están aumentando constantemente [21, 22]. Es por ello que, los profesionales de la salud mental deben familiarizarse con la IA, comprender sus usos actuales y futuros, y estar para trabajar en conjunto con sus técnicas.

Al mismo tiempo, el uso de la IA en la medicina para aplicaciones de salud física es más frecuente que en la salud mental [23], ya que los profesionales de la salud mental confían en habilidades "soft" y en datos clínicos subjetivos y cualitativos. Sin embargo, hay oportunidades emergentes para incorporar la IA en las intervenciones digitales y mejorar la experiencia del usuario y la atención personalizada de la salud mental [24-26]. Ésta tiene el potencial de redefinir el diagnóstico y la comprensión de las enfermedades mentales, identificando biomarcadores para definiciones más objetivas. La implementación de enfoques computacionales que se adapten a grandes datos permite desarrollar herramientas de detección previa al diagnóstico y modelos de riesgo para la predisposición a enfermedades mentales.

2. Líneas de Investigación, Desarrollo e Innovación

En el ámbito del grupo de investigación, las primeras investigaciones en torno a la Inteligencia Artificial se desarrollan con el Proyecto de Investigación y Desarrollo (PID) denominado "Implementación de sistemas

inteligentes para la asistencia a alumnos y docentes de la carrera de ingeniería en sistemas de información", que toma lugar desde el año 2016 hasta el 2019. Éste busca utilizar tecnologías de IA y desarrollos de software inteligente para la resolver problemas dentro del campo educativo.

De la misma manera, desde el año 2019 hasta el 2022, se lleva a cabo un PID llamado "Prácticas ingenieriles aplicadas para la implementación de Sistemas Inteligentes basados en Machine Learning". Este proyecto analiza el uso de buenas prácticas ingenieriles para el abordaje de desarrollo de software inteligente que hace uso Machine Learning.

Finalmente, a partir del presente año (2023), se da inicio al proyecto "Inteligencia Artificial para el análisis predictivo en salud mental". En este contexto, se busca proponer un conjunto de técnicas, tareas y métodos que aportan a la implementación de Modelos Predictivos basados en Inteligencia Artificial en el ámbito de la salud mental.

De esta manera, se articula dentro de los objetivos de GEMIS en el campo de la Informática, la generación de nuevos conocimientos en el área de la Ingeniería de Software y la Inteligencia Artificial.

3. Resultados esperados

Objetivo General

Proponer un conjunto de tareas, métodos, y técnicas que contribuyan a la implementación de Modelos Predictivos basados en Inteligencia Artificial en el campo de la salud mental.

Objetivos Específicos

Los objetivos específicos son:

- Identificar las tareas de descubrimiento, evaluación,

recolección y preparación de los datos para la construcción de un Modelo Predictivo en la salud mental teniendo en cuenta las características de las fuentes de datos previamente identificadas y estructurar posteriormente los datos en términos de datasets estandarizados e interoperables.

- Desarrollar modelos supervisados utilizando el aprendizaje automático para la predicción de la aparición de trastornos de salud mental, como la depresión, en función del análisis de datos clínicos y de comportamiento.
- Desarrollar modelos no supervisados utilizando el aprendizaje automático, como la agrupación para realizar análisis holísticos de cohortes y perfiles de riesgo predictivos para trastornos mentales.
- Desarrollar un nuevo conjunto de datos estructurados utilizando los conocimientos de los modelos de aprendizaje automático y utilizando los formatos estandarizados DSM-5.
- Aplicar los modelos supervisados para el diagnóstico precoz de los trastornos de salud mental comparando datos clínicos y conductuales retrospectivos y prospectivos de la cohorte de estudio.
- Diseñar un sistema de apoyo a la decisión clínica basado en modelos supervisados y no supervisados para diagnosticar trastornos de salud mental seguido de apoyo al profesional médico para un enfoque clínico personalizado.

Resultados esperados

Los resultados que se esperan del presente proyecto son los siguientes:

- Asistir a los médicos y profesionales de la salud mental a mejorar el proceso de diagnóstico de las enfermedades en busca de una intervención pronta y

personalizada de acuerdo con las características de los pacientes.

- Contribuir con los desarrolladores en la implementación de Modelos Predictivos en el ámbito de la salud.
- Contribuir con nuevo conocimiento a la comunidad académica y, a la sociedad en general en el ámbito nacional e internacional, transfiriendo los resultados (parciales y finales) de este proyecto.

4. Formación de Recursos Humanos

El equipo de trabajo se encuentra conformado por 4 investigadores formados, 1 tesista doctoral, 2 tesistas de maestría, 2 graduados de grado, 1 Becario BINID y 2 alumnos en la carrera de Ingeniería en Sistemas de Información. El proyecto busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para que asciendan dentro del escalafón de la carrera de investigadores.

Los recursos humanos deben poseer una firme vocación de trabajar en el área de la Ingeniería de Software, la Ingeniería del Conocimiento y Machine Learning, de manera de canalizar y proveer una base sustentable de aporte al proyecto. De esta manera, se logra un doble beneficio, el proyecto obtiene e incorpora el conocimiento tecnológico de los recursos humanos en el área de la especialidad, a la vez que plantea un esquema de formación de especialistas de punta en el proceso de gestión.

Asimismo, en el marco de este proyecto de investigación se prevé incorporar alumnos avanzados en la carrera de Ingeniería en Sistemas de Información con posibilidades de articular sus Trabajos Finales de Carrera de Grado, así como,

también tesistas de posgrado que desarrollarán sus Tesis de Maestría. De esta manera se espera generar un verdadero espacio integrado de investigación en el área de Sistemas de Información en el nivel de carreras de grado y posgrado.

5. Referencias

- [1] Universidad Tecnológica Nacional (2016). Resolución Nro. 2508/16 Referencias bibliográficas
- [2] M. A. Boden, *Artificial Intelligence*. Elsevier, 1996.
- [3] «Definitions of dataset in the scientific and technical literature - Renear - 2010 - Proceedings of the American Society for Information Science and Technology - Wiley Online Library». <https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/meet.14504701240> (accedido febrero 2023).
- [4] A. R. Ajiboye, R. Abdullah-Arshah, H. Qin, y H. Isah-Kebbe, «Evaluating the effect of dataset size on predictive model using supervised learning technique», *Int. J. Comput. Syst. Softw. Eng.*, vol. 1, n.o 1, pp. 75-84, feb. 2015, doi: 10.15282/ijsecs.1.2015.6.0006.
- [5] «5 Trends Emerge In Gartner Hype Cycle For Emerging Technologies 2018», Gartner. <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018> (febrero 2023).
- [6] R. L. Villars, M. Eastwood, y C. W. Olofson, «Big Data: What It Is and Why You Should Care», p. 14
- [7] S. Graham et al., «Artificial Intelligence for Mental Health and Mental Illnesses: an Overview», *Curr. Psychiatry Rep.*, vol. 21, n.o 11, p. 116, nov. 2019, doi: 10.1007/s11920-019-1094-0.
- [8] F. Jiang et al., «Artificial intelligence in healthcare: past, present and future», *Stroke Vasc. Neurol.*, vol. 2, n.o 4, dic. 2017, doi: 10.1136/svn-2017-000101.
- [9] J. Morley et al., «The ethics of AI in health care: A mapping review», *Soc. Sci. Med.* 1982, vol. 260, p. 113172, sep. 2020, doi: 10.1016/j.socscimed.2020.113172.
- [10] B. X. Tran et al., «The Current Research Landscape on the Artificial Intelligence Application in the Management of Depressive Disorders: A Bibliometric Analysis», *Int. J. Environ. Res. Public Health*, vol. 16, n.o 12, Art. n.o 12, ene. 2019, doi: 10.3390/ijerph16122150.
- [11] K. Kroenke, T. W. Strine, R. L. Spitzer, J. B. W. Williams, J. T. Berry, y A. H. Mokdad, «The PHQ-8 as a measure of current depression in the general population», *J. Affect. Disord.*, vol. 114, n.o 1, pp. 163-173, abr. 2009, doi: 10.1016/j.jad.2008.06.026.
- [12] K. Kroenke y R. L. Spitzer, «The PHQ-9: A New Depression Diagnostic and Severity Measure», *Psychiatr. Ann.*, vol. 32, n.o 9, pp. 509-515, sep. 2002, doi: 10.3928/0048-5713-20020901-06.
- [13] «Beck Depression Inventory (BDI)», <https://www.apa.org>. <https://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/beck-depression> (accedido febrero 2023).
- [14] D. J. A. Dozois, «Beck Depression Inventory-II», en *The Corsini Encyclopedia of Psychology*, John Wiley & Sons, Ltd, 2010, pp. 1-2. doi: 10.1002/9780470479216.corpsy0113.
- [15] «Children's Depression Inventory - PscNET». <https://doi.apa.org/doiLanding?doi=10.1037%2Ft00788-000> (accedido febrero 2023).
- [16] «APA - The Structured Clinical Interview for DSM-5®». <https://www.appi.org/products/structured-clinical-interview-for-dsm-5-scid-5> (febrero 2023).
- [17] S. Shiffman, A. A. Stone, y M. R. Hufford, «Ecological momentary assessment», *Annu. Rev. Clin. Psychol.*, vol. 4, pp. 1-32, 2008, doi: 10.1146/annurev.clinpsy.3.022806.091415.
- [18] «Mental health: strengthening our response». <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response> (accedido febrero 2023).
- [19] Metz C, Smith CS. «A.I. Can Be a Boon to Medicine That Could Easily Go Rogue». *The New York Times*; 2019. March 25;B5.
- [20] Kim JW, Jones KL, Angelo ED. How to Prepare Prospective Psychiatrists in the Era of Artificial Intelligence. *Acad Psychiatry*. 2019;1-3. DOI: 10.1007/s40596-019-01025-x
- [21] A Muehlematter, Urs J.; Daniore, Paola; Vokinger, Kerstin N. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *The Lancet Digital Health*, 2021, vol. 3, no 3, p. e195-e203. DOI: 10.1016/S2589-7500
- [22] Malik, Paras, et al. Overview of artificial intelligence in medicine. *Journal of family medicine and primary care*, 2019, vol. 8, no 7, p. 2328-2331. DOI: 10.4103/jfmpe.jfmpe_440_19
- [23] John McCarthy. Artificial intelligence, logic and formalizing common sense In *Philosophical logic and artificial intelligence* 1989. (pp. 161–190). Springer, Dordrecht.
- [24] Pang Z, Yuan H, Zhang Y-T, Packirisamy M. Guest Editorial Health Engineering Driven by the Industry 4.0 for Aging Society. *IEEE J Biomed Heal Informatics*. 2018;22(6):1709–10. DOI: 10.1109/JBHI.2018.2874081
- [25] Schwab K The fourth Industrial Revolution. first New York, NY: Currency; 2017. p. 192
- [26] Simon HA. Artificial Intelligence: Where Has It Been, and Where is it Going? *IEEE Trans Knowl Data Eng*. 1991;3(2):128–36. DOI: 10.1109/69.87993

SOLUCIONES DE IOT: ÁMBITOS DE APLICACIÓN Y DESAFÍOS

Sebastián U. Flores, Mario Berón, Daniel Riesco

Departamento de Informática - Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis - Ejército de Los Andes 950 - San Luis - Argentina
s.flores@outlook.com.ar, { mberon, driesco }@unsl.edu.ar

RESUMEN

El Internet of Things involucra un conjunto heterogéneo de tecnologías, posee diferentes ámbitos de aplicación y un gran número de riesgos asociados. Es por ello que se debe analizar minuciosamente cada componente de las soluciones a crear. Un componente esencial a tener en cuenta son las personas. Estas interactúan con el sistema, a veces de forma voluntaria y otras, desconociendo su presencia. Diferentes regiones geopolíticas han llevado adelante legislaciones de diferentes características, con el fin de proteger a las personas frente a vulneraciones de su privacidad, o a los posibles riesgos de seguridad físicos asociados a la interacción con soluciones tecnológicas, sea que estas involucren IoT o que no lo hagan. Un sistema de IoT Industrial posee diferentes riesgos y legislaciones a los que posee un sistema hogareño para el control de electrodomésticos, o una red de producción y distribución eléctrica inteligente.

No obstante, los riesgos van más allá de su relación con las personas. Están también relacionados con el negocio, y la posibilidad de que se detenga el funcionamiento del sistema, o que se tomen decisiones incorrectas a partir de la captura de datos erróneos.

Este artículo pretende presentar el estado del arte del IoT y abordar algunos de sus desafíos más importantes en la actualidad.

Palabras clave: *IoT, Dispositivo, Nube, Escalabilidad, Seguridad, Privacidad, Confiabilidad, Arquitectura, Internet.*

CONTEXTO

La presente línea de investigación se enmarca en el proyecto de investigación denominado

“Ingeniería del Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”. El proyecto, a su vez, es un *Proyecto de Investigación Consolidado - PROICO*, y posee el código *03-2020*. Tal proyecto es la continuación de diferentes proyectos de investigación, a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional.

1. INTRODUCCIÓN

De acuerdo con Oracle [1], el Internet de las cosas (IoT) describe la red de objetos físicos ("cosas") que llevan incorporados sensores, software y otras tecnologías con el fin de conectarse e intercambiar datos con otros dispositivos y sistemas a través de Internet.

A grandes rasgos, podrían identificarse varios ámbitos de aplicación de tecnologías IoT [2]:

- Hogares conectados. Los dispositivos de IoT pertenecientes a este ámbito incluyen una amplia gama de electrodomésticos creados con un enfoque de *diseño orientado al usuario* [3, 4], que permite analizar el comportamiento de este último y adaptarse a sus necesidades, para mejorar la experiencia de uso y la utilidad. Este ámbito también incluye dispositivos para el incremento de la seguridad hogareña, como sensores de gas/humo, cerraduras inteligentes y alarmas, entre otros.
- Vehículos conectados. Este ámbito todavía se encuentra en etapas tempranas, e incluye, entre otros beneficios, la posibilidad de formar una red de vehículos conectados a Internet que puedan comunicarse entre sí para realizar optimizaciones en el tráfico e incrementar la seguridad de los pasajeros.
- Ciudades inteligentes. Este ámbito

incluye la posibilidad de instalar sensores en diferentes partes de una ciudad, tanto en ámbitos públicos como privados, para tener más información que ayude a mejorar los servicios públicos, la calidad del aire, disminuir el ruido ambiental, optimizar el mantenimiento de los sistemas de iluminación y carreteras, detectar fugas de cañerías de agua potable, entre otros.

- La industria. Este ámbito es conocido como Internet Industrial de las Cosas (IIoT), y refiere al uso de dispositivos para optimizar e incrementar la seguridad en fábricas, servicios de logística, venta minorista, sistemas de producción de energía, entre otros.

En los últimos años, se produjeron grandes avances tecnológicos que propiciaron el surgimiento y masificación del IoT. Algunos de estos avances se presentan a continuación:

- Disminución de los costos en sensores de buena calidad, y en chips de alta potencia.
- Grandes progresos en los medios de comunicación, que incrementaron su ancho de banda y disminuyeron, tanto su latencia como sus costos, a nivel mundial.
- Expansión y perfeccionamiento de las plataformas en la nube para el almacenamiento y procesamiento de información.
- Avances en las tecnologías de análisis de datos y aprendizaje automático.
- Incremento de la accesibilidad a la tecnología, a partir de la expansión de inteligencias artificiales conversacionales y dispositivos móviles.

Cada sistema de IoT, dependiendo del ámbito y de las necesidades de negocio, estará conformado por diferentes componentes. No obstante, los siguientes componentes podrían considerarse comunes a la mayor parte de los sistemas de IoT [2]:

- Dispositivos de IoT. Son dispositivos tecnológicos conformados por sensores, actuadores, sistemas de conectividad y microcontroladores, cuyo fin es la recopilación

de datos del entorno, de los usuarios o de los patrones de uso; la comunicación de los datos recopilados a través de Internet; y la ejecución de acciones para transformar el ambiente físico donde se encuentran, por medio de sus actuadores.

- Servidores de almacenamiento y cómputo. Se encargan de integrar los datos publicados por los dispositivos IoT, almacenarlos, analizarlos por medio de aprendizaje automático y otras tecnologías, y tomar decisiones sobre los mismos. Estas decisiones incluyen informar a administradores de la solución o solicitar a los dispositivos la ejecución de diferentes acciones inteligentes. En su mayoría, se encuentran desplegados en la nube, por lo que cuentan con capacidades de almacenamiento y cómputo prácticamente sin límites.

- Puertas de enlace. Existen diversos casos en que los dispositivos no pueden comunicarse directamente con los servidores, por lo que se utilizan puertas de enlace, que son puntos de conexión que actúan como intermediarios entre los dispositivos y los servidores.

- Aplicaciones de usuario. Son aplicaciones de software que poseen interfaces visuales por medio de las cuales los diferentes usuarios de la solución IoT pueden visualizar los datos obtenidos por los dispositivos, controlar el estado de conexión, modificar configuraciones de la solución y solicitar la ejecución de acciones a los dispositivos.

- Computación de borde. Es un conjunto de tecnologías que permite dotar a los dispositivos de la capacidad de realizar un análisis y procesamiento sobre los datos obtenidos por medio de sus sensores, para tomar decisiones más rápidamente, evitando una dependencia total de los servidores.

En base a los párrafos anteriores, puede notarse que el espectro del IoT es muy amplio y que, si bien existen grandes avances, debe tenerse en cuenta una amplia variedad de factores al momento de diseñar e implementar un sistema de IoT [5, 6, 7]:

- Seguridad. Este es un factor muy importante que incluye múltiples facetas. En primer lugar, se encuentra la protección de datos en tránsito, que incluye a diferentes protocolos de comunicación entre los componentes del sistema, y diferentes formas de encriptación de los datos antes de ser transmitidos. En segundo lugar, la protección de datos en reposo incluye diferentes metodologías y algoritmos de encriptación para el almacenamiento seguro de los datos, una vez transmitidos hacia un servidor. En tercer lugar, la seguridad del hardware de los dispositivos es aquella que impide a personas sin autorización modificar los dispositivos, o bien, extraer y analizar los datos almacenados en ellos y su código fuente. En caso de usarse puertas de enlace, también deben incluirse medidas de seguridad en las mismas, ya que estas interactúan con los dispositivos y con los servidores. Adicionalmente, la seguridad debe implementarse en los servidores, para protegerlos de diferentes ataques maliciosos, como los ataques de denegación de servicio (DDoS). Para concluir, deben incluirse protecciones de seguridad complementarias para las aplicaciones de usuario, como la autenticación multifactor y la autenticación sin contraseña [8].

- Privacidad. Dependiendo del entorno en donde se despliegan los sistemas de IoT, podrían verse involucradas diferentes personas en las mediciones de los sensores, y muchas de ellas podrían desconocer la existencia de dichos sistemas, o bien, desconocer la clase de información que estos extraen y los potenciales riesgos asociados a las vulneraciones de seguridad. Es por ello que, sumado a las protecciones de seguridad establecidas, debe minimizarse (y de ser posible, evitarse) el almacenamiento de Información de Identificación Personal (PII) [9, 10]. Por otra parte, debe informarse claramente a todas las personas que puedan interactuar con el sistema de IoT, sobre los fines del sistema y los datos extraídos, además de solicitar su consentimiento, tal y como se ha establecido en diferentes legislaciones alrededor del mundo.

- Restricciones legales. Como se mencionó anteriormente, existen diferentes legislaciones alrededor del mundo, asociadas a la privacidad, la localidad, la pertenencia y los medios habilitados para la extracción, almacenamiento y procesamiento de los datos [11, 12, 13].

- Confiabilidad. Las soluciones IoT a menudo se implementan a gran escala y, pueden funcionar en redes poco seguras o de mala calidad, con una alta exposición a condiciones ambientales adversas, que podrían afectar los requisitos establecidos en su diseño inicial. No obstante, es fundamental que la solución funcione correctamente en los momentos críticos. Es muy importante que los sensores aporten datos confiables, para evitar que se tomen decisiones erróneas. Incluso, los actuadores deben funcionar correctamente, para que pueda actuarse de la manera adecuada y en el momento que se lo necesita. Por lo mencionado anteriormente, es esencial diseñar la arquitectura de IoT con la disponibilidad y resistencia en mente, realizando una cuidadosa selección del hardware de los dispositivos, los protocolos, los servicios en la nube y los medios de conexión. También debe considerarse la posibilidad de tener redundancia en los sectores críticos de la solución, para reducir los puntos únicos de falla que afecten al sistema en su totalidad.

- Resistencia al ambiente. Dependiendo del ámbito, los dispositivos de IoT podrían estar expuestos a diferentes condiciones ambientales, como cambios de temperatura, lluvia, humedad, polvo, ruido eléctrico, inconsistencias en el suministro eléctrico o manipulación errónea por parte de los usuarios.

- Escalabilidad. Un sistema de IoT podría crecer con el tiempo de forma horizontal (si se agregan más dispositivos de un mismo tipo) o vertical (si se incluyen dispositivos diferentes). Es así que la arquitectura del sistema debe planificarse adecuadamente, buscando prever estos posibles cambios en la escala del sistema, y teniendo en cuenta el incremento en los costos, en la complejidad de la infraestructura

y en la cantidad de mensajes transmitidos.

- **Flexibilidad.** Las soluciones de IoT se caracterizan por la integración de componentes propios o creados por terceros, los cuales podrían utilizar diferentes protocolos de comunicación, tecnologías de hardware y software, o tener diferentes requisitos de infraestructura y mantenimiento. A su vez, pueden escalar notablemente de acuerdo con cambios en los requisitos del usuario final o a la detección de puntos de mejora. La flexibilidad aborda la capacidad de la solución de adaptarse a diferentes necesidades en diferentes momentos del tiempo.

- **Eficiencia energética.** Los dispositivos podrían ser desplegados en entornos donde sea difícil o costoso acceder al suministro eléctrico, por lo que deberían tener un bajo consumo energético. Esto trae otras implicancias como la selección de hardware de menor potencia, el uso de protocolos de comunicación más ligeros (y quizás, menos seguros) o la necesidad de utilizar energías renovables.

- **Conectividad.** Dependiendo de los riesgos de los procesos, podría requerirse una menor latencia en las comunicaciones, para poder tomar decisiones en menor tiempo. En ámbitos controlados como los industriales, es posible establecer las configuraciones necesarias de conectividad de sistemas y dispositivos locales. En ámbitos no controlados, como aquellos correspondientes a instalaciones de consumidores finales, sistemas de logística o sistemas desplegados en zonas rurales, es muy probable que el nivel de conexión se vea afectado por condiciones climáticas o disponibilidad de proveedores de conexión locales. Es por ello que los sistemas de IoT deben estar preparados para funcionar correctamente, a pesar de tener largos períodos sin conexión. Otra alternativa es realizar un despliegue de sistemas de conexión propios, como la creación de redes LoRaWAN [14].

- **Presupuesto.** En todos los factores mencionados antes, los costos son una variable

común que influye en la selección de las configuraciones del sistema de IoT. En la medida que se busque incrementar la seguridad, mejorar la latencia o escalar el sistema, los costos se verán notablemente incrementados, tanto en los dispositivos de IoT como en el resto de la infraestructura utilizada para las comunicaciones, almacenamiento y procesamiento de datos. Debe analizarse a detalle los requisitos de negocio, los riesgos asociados al sistema y la posibilidad de que este cambie con el tiempo, para seleccionar los componentes a utilizar y conformar el presupuesto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para diseñar una solución IoT, es imprescindible identificar cada uno de los problemas involucrados y atacarlos con la mejor tecnología posible, comprendiendo las limitaciones inherentes al proyecto y a factores externos.

Con el fin de facilitar la búsqueda de soluciones óptimas a los problemas que se presenten en el diseño de una solución de IoT escalable y segura, esta línea de investigación propone:

- Investigar posibles soluciones a los desafíos presentados en la introducción.
- Estudiar las legislaciones y la matriz productiva nacional y regional, para determinar posibles ámbitos de aplicación de soluciones IoT.

3. RESULTADOS OBTENIDOS/ESPERADOS

Realizar informes que detallen el estado del IoT a nivel nacional y regional, en relación con el estado en otras regiones del mundo.

Realizar propuestas de soluciones a los diferentes desafíos mencionados en la Introducción, poniendo énfasis en aquellas con mayor viabilidad a nivel nacional y regional.

Implementar, o colaborar en la implementación de soluciones de IoT reales.

4. FORMACION DE RECURSOS HUMANOS

Los progresos obtenidos en esta línea de investigación sirven como base para el desarrollo de tesis de posgrado, ya sea de doctorado o maestrías en Ingeniería de Software y desarrollo de trabajos finales de las carreras Licenciatura en Ciencias de la Computación, Ingeniería en Informática e Ingeniería en Computación de la Universidad Nacional de San Luis, en el marco de los Proyectos de Investigación mencionados en la Sección *Contexto*.

5. BIBLIOGRAFÍA

- [1] Oracle, «¿Qué es el IoT?,» [En línea]. Available: <https://www.oracle.com/es/internet-of-things/what-is-iot/>.
- [2] Amazon, «¿Qué es el IoT?,» [En línea]. Available: <https://aws.amazon.com/es/what-is/iot/>.
- [3] M. Mezzenzana, «Internet-of-Things as an enabling factor for user-centered service engineering.,» 2019. [En línea]. Available: https://www.researchgate.net/publication/360256759_Internet-of-Things_as_an_enabling_factor_for_user-centered_service_engineering.
- [4] ThingsCon, «User Centred IoT-Design,» 12 06 2017. [En línea]. Available: <https://medium.com/the-state-of-responsible-internet-of-things-iot/andreakrajewski-aff52af1e065>.
- [5] Microsoft, «IoT Overview,» [En línea]. Available: <https://learn.microsoft.com/es-es/azure/architecture/framework/iot/iot-overview>.
- [6] Microsoft, «IoT Reliability,» [En línea]. Available: <https://learn.microsoft.com/es-es/azure/architecture/framework/iot/iot-reliability>.
- [7] S. N. C. Z. S. e. a. Moore, «IoT reliability: a review leading to 5 key research directions,» 2020. [En línea]. Available: <https://doi.org/10.1007/s42486-020-00037-z>.
- [8] Entrust, «¿Qué es la Autenticación Multifactor (MFA)?,» [En línea]. Available: <https://www.entrust.com/es/resources/faq/what-is-multi-factor-authentication-mfa>.
- [9] Investopedia, «What Is Personally Identifiable Information (PII)? Types and Examples,» [En línea]. Available: <https://www.investopedia.com/terms/p/personally-identifiable-information-pii.asp>.
- [10] V. J. Guareteguá, «Información de Identificación Personal (PII) / Personally Identifiable Information (PII),» 2020. [En línea]. Available: <https://www.linkedin.com/pulse/informaci%C3%B3n-de-identificaci%C3%B3n-personal-pii-personally-javier/?originalSubdomain=es>.
- [11] European Union, «General Data Protection Regulation,» [En línea]. Available: <https://gdpr-info.eu/>.
- [12] Gobierno Nacional de la República Argentina, «Ley 25.326 de la Protección de los Datos Personales,» 04 10 2000. [En línea]. Available: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/texact.htm>.
- [13] State of California Department of Justice, «California Consumer Privacy Act,» 02 2023. [En línea]. Available: <https://www.oag.ca.gov/privacy/ccpa>.
- [14] LoRa Alliance, «What is LoRaWAN® Specification,» [En línea]. Available: <https://loro-alliance.org/about-lorawan/>.

MACHINE LEARNING Y DEEP LEARNING EN LA INTERPRETACIÓN DEL LENGUAJE DE SEÑAS

Raúl Oscar Klenzi, Maria Isabel Masanet, Facundo Recabarren, Silvia Saez, Gustavo Conturso

Instituto de Informática / Departamento de Informática /Facultad de Ciencias Exactas Físicas y Naturales / Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas", Rivadavia, San Juan,
Teléfonos: 4260353, 4260355 Fax 0264-4234980, Sitio Web: <http://www.exactas.unsj.edu.ar>
e-mail: {rauloscarklenzi, mimasanet}@gmail.com

RESUMEN

La pandemia COVID19 puso en evidencia la gran dificultad de la comunidad hipoacúsica para comunicarse con el resto de la sociedad. El uso de barbijos, máscaras y barreras transparentes como alternativas atenuadoras de contagios generó una barrera casi infranqueable para esa comunidad cuando utilizaban la lectura de labios para establecer la comunicación. Por ello, la única alternativa que tenían era al uso de la Lengua de Señas (LS), herramienta que el resto de la sociedad, en su mayoría, desconoce; por lo que debían y deben recurrir a una tercera persona que no siendo hipoacúsica y comprendiendo la lengua de señas hiciera de interprete.

Los objetivos que persigue el presente trabajo se centran en la utilización de algoritmos derivados del Machine Learning (ML) y Deep Learning (DL) aplicado al reconocimiento de expresiones en LS a partir de una secuencia de imágenes (video), y lograr traducir a texto o audio estas expresiones, como así también el camino inverso, emulando con ello al interprete humano. Estos objetivos exigen la utilización de hardware veloz tipo GPU, gran capacidad de memoria, y algoritmos eficientes para el procesamiento de la información.

Palabras clave: Machine Learning, Deep Learning, LSA.

CONTEXTO

En el marco del proyecto “Evaluación de visualizaciones eficientes en Ciencia de Datos” concluido en diciembre de 2022 se comenzó la ejecución, bajo el asesoramiento de integrantes del proyecto, de un trabajo final de grado en la carrera Licenciatura en Ciencias de la Computación. El trabajo, “Sistema Web Intérprete de Lengua de Señas Argentina” permitió ir conociendo la punta de un iceberg evidenciado por las particularidades de la Lengua de Señas Argentina (LSA), el contexto de la comunidad sorda donde lo aplica, y la necesidad cierta y mundialmente explicitada de generar un proceso automático que permita la intercomunicación de la comunidad hipoacúsica con el resto de la sociedad. Esto habilitó que tras la incorporación de intérpretes de LSA al grupo de trabajo, se presentara el proyecto para el bienio 2023-2024 “MACHINE LEARNING Y DEEP LEARNING APLICADO A LENGUA DE SEÑAS ARGENTINA”, el cual próximamente será evaluado en el ámbito de la UNSJ por investigadores externos a nuestra Universidad. El background obtenido por los investigadores en el área del aprendizaje profundo aplicado al procesamiento de imágenes permite realizar la propuesta e intentar dar un primer paso en el objetivo de facilitar la comunicación entre ambas

comunidades, la hipoacúsica y la oyente, cuya necesidad y dura realidad se evidenció en forma acrecentada en la pandemia COVID19.

1. INTRODUCCIÓN

En las últimas dos décadas, el campo del Machine Learning (aprendizaje automático) ha jugado un papel relevante en el desarrollo de aplicaciones de software [1]. Se trata de un área de las ciencias de la computación y de la ciencia de datos (Data Science) que utiliza algoritmos y métodos estadísticos para aprender de los datos, extraer inferencias y reconocer patrones sin programación explícita. El Deep Learning (aprendizaje profundo) es una subárea del Machine Learning (Figura 1) que utiliza una red neuronal de varias capas para imitar el complejo procesamiento de información del cerebro humano [2].

La explosión de datos digitales en una amplia variedad de dominios, como la ciencia, la ingeniería, el internet de las cosas, la biomedicina, la atención médica y muchos sectores comerciales, ha declarado la era de los grandes volúmenes de datos (big data). Estos datos no pueden analizarse con las estadísticas clásicas, sino que requieren de las más modernas técnicas de ML y DL. Dado que el aprendizaje automático aprende de los datos en lugar de programar reglas de decisión codificadas, se está intentando utilizar el aprendizaje automático para crear computadoras que puedan resolver problemas como expertos humanos en el campo [1].

Las personas con discapacidad verbal y auditiva enfrentan dificultades en la comunicación, debido a esto, se sienten aisladas y dependientes [1]. Estas personas utilizan la lengua de señas para comunicarse, pero la interacción social se dificulta debido a que la mayoría de la población no puede interpretar la lengua de señas.

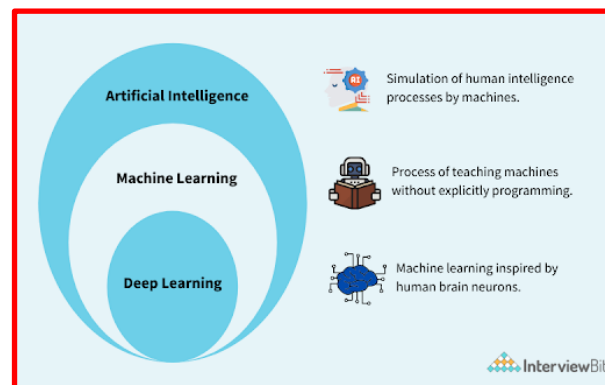


Figura 1 - Machine Learning y Deep Learning en la Inteligencia artificial.

Esta dificultad en la comunicación con el mundo que habla verbalmente, hace difícil obtener educación, trabajo y muchos otros requerimientos básicos. Para comunicarse con la sociedad, necesitan un intérprete de señas manual o automático. El primero puede convertir cualquier secuencia de gestos complicada a voz con todas las modulaciones emocionales requeridas para ello, pero reduce la privacidad de las personas involucradas en la conversación.

Un intérprete no manual o automático garantiza el factor de privacidad, pero aún se está investigando para desarrollar un sistema completo de reconocimiento de lengua de señas que pueda funcionar con la misma precisión que un intérprete humano. La complejidad del problema se evidencia en los desafíos involucrados en la interpretación automática de la lengua de señas, como ser:

- La lengua de señas es muy diversa; difiere de una geografía a otra, e incluso dentro de una región.
- La lengua de señas involucra gestos con una sola mano, gestos con dos manos, expresión facial y postura corporal.
- La conversación en lengua de señas contiene gestos estáticos y dinámicos realizados en una secuencia y, por lo tanto, encontrar el comienzo y el final de cada gesto en la

secuencia implica tareas de preprocesamiento en los datos (videos).

- Oclusión: Mientras se realiza el gesto, una mano puede ocluir a la otra, lo que dificulta la interpretación [1].

En este caso la propuesta contempla el desarrollo de un prototipo de aplicación web que interprete de forma automática expresiones de la Lengua de Señas Argentina (LSA) y emita, en formato escrito o de audio, la interpretación realizada. Dicha aplicación será testada por los intérpretes de LSA integrantes del proyecto como así también por otros miembros de la comunidad hipoacúsica de San Juan.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación consideradas son las atinentes al procesamiento de imágenes mediante la utilización de algoritmos de ML y DL, que permitan hacer una interpretación adecuada y automática de los movimientos expresados por el hipoacúsico obteniendo así el correspondiente fragmento de texto o audio.

La investigación se orientará a integrar el Machine Learning y Deep Learning con la Lengua de Señas Argentina, con el objetivo de generar conocimiento de interés para desarrollos tecnológicos que faciliten la comunicación e interacción entre personas que utilizan la LSA y personas oyentes.

Técnicas de Machine Learning y Deep Learning, Visión por computadora (Computer Vision) y la Lengua de Señas Argentina son los principales conceptos sobre los que se basará la investigación propuesta.

2.1. Machine Learning

Dentro del área de la Inteligencia Artificial (IA) se desprende una rama específica conocida como Machine Learning (ML) o Aprendizaje de Máquina. Es una rama de

algoritmos computacionales diseñados para emular la inteligencia humana mediante el aprendizaje del entorno [3].

2.2. Deep Learning

El Deep Learning es un área del Machine Learning que a través de una red neuronal trata de imitar el complejo procesamiento de información del cerebro humano.

Al igual que las neuronas biológicas en el cerebro humano, se utiliza un perceptrón como componente básico de una red neuronal artificial y realiza el procesamiento de la información [2].

El aprendizaje profundo está siendo ampliamente aceptado para tareas de visión por computadora, ya que ha demostrado capacidades casi humanas o incluso mejores para realizar numerosas tareas. Las redes profundas son efectivas en el reconocimiento visual, reconocimiento de objetos, seguimiento de movimiento, reconocimiento de acciones, estimación de la pose humana y la segmentación semántica [1].

2.2.1. Red Neuronal

Una red neuronal consiste en una gran cantidad de neuronas (o perceptrones) interconectadas entre sí en capas y es responsable del procesamiento paralelo masivo de datos de entrada para producir salidas (Figura 2). Consta de una capa de entrada, una capa de salida y varias capas ocultas.

En un sistema de aprendizaje profundo, las capas ocultas definen la profundidad de la red, la que depende de la naturaleza del problema en cuestión y del volumen de datos considerado para el procesamiento.

Las redes neuronales son potentes modelos de aprendizaje que logran resultados de vanguardia en una amplia gama de tareas de aprendizaje automático (ML) supervisado y no supervisado [4].

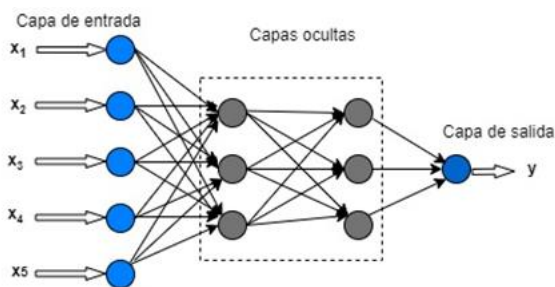


Figura 2 - - Arquitectura de una red neuronal con dos capas ocultas y una capa de salida [5].

2.3. Visión por computadora (Computer vision)

La visión por computadora se puede entender de distintas maneras, dependiendo del objetivo con el cual es usada. Puede ser utilizada para el procesamiento digital de imágenes, el análisis de imágenes, el reconocimiento de patrones o computación gráfica entre otras [6].

Implica interpretar y comprender el mundo a partir de sus imágenes o videos. Las tareas de la visión por computadora se pueden organizar jerárquicamente en niveles bajo, medio y alto. Mientras que las tareas de bajo nivel se centran en el procesamiento de imágenes para mejorarlas y extraer información, las tareas de nivel medio implican estimar las propiedades de los objetos en las imágenes, incluidas las propiedades geométricas, el movimiento y las categorías de objetos. Las tareas de alto nivel se centran en interpretar y comprender eventos o actividades en las imágenes/videos [7].

2.4. Lengua de señas

Existen muchas lenguas de señas en el mundo. La existencia de una o más lenguas en un mismo país o en más de un país depende de la historia de las comunidades Sordas de cada lugar [8]. La Lengua de Señas Argentina es la lengua de la comunidad sorda argentina. Es una lengua natural, con una estructura gramatical diferente a la del español. Esto significa que la LSA y el español son dos lenguas distintas.

Cada seña que compone a una lengua de señas está constituida por la combinación de 5 elementos: la configuración, la orientación, la ubicación, el movimiento y componentes no manuales (movimientos de hombros, cabeza, expresiones faciales).

La estructura básica entre el español y la Lengua de Señas Argentina difieren. En vez de la estructura del español ‘sujeto, verbo, objeto’ (Ejemplo: el perro corre la pelota), en la LSA el verbo va al final (Ejemplo: el perro a la pelota corre). Este orden es canónico, es decir, es el orden hacia el que la lengua tiende aunque en el discurso se viola por efectos semánticos y pragmáticos [9].

3. RESULTADOS OBTENIDOS Y ESPERADOS

Al momento, y fundamentalmente desde el desarrollo de un trabajo final de grado, se ha trabajado con el reconocimiento de letras expresadas en LSA como imágenes estáticas. Se han reconocido palabras en las que ya aparecen secuencias de imágenes (short video) y breves frases de entrenamiento realizadas por diferentes intérpretes; cuyos videos conforman el set de datos “LSA64” construido en [10]; y que, desde la representación realizada por diferentes personas hipoacúsicas se independizan del “quien” para centrarse en el “qué o cuál” es la secuencia analizada. Sin perder capacidad de reconocimiento.

Lo anterior llevó a tratar de reducir la tasa de frame que tiene cada video con el objetivo de disminuir el espacio de almacenamiento de las diferentes expresiones lingüísticas y acelerar el procesamiento sin perder capacidad de discernimiento y reconocimiento.

Para el preprocesamiento y desarrollo de los modelos de reconocimiento se usa el lenguaje de programación Python, el cual proporciona diferentes librerías como Matplotlib, Pandas, NumPy, Sci-Kit Learn, ScraPy, Keras, Tensor

Flow, entre otros, que ayudan a que los procesos de ciencia de datos y Machine Learning sean más manejables y eficientes. Por último, hay librerías que permiten el procesamiento de fotogramas, sin dudas necesarios para las tareas de preprocesamiento requeridas, como son MediaPipe, PoseNet, HandPose, OpenPose, entre otras. Estas librerías marcan los puntos del cuerpo, cara, brazos, manos, piernas en una secuencia de imágenes (video), necesarios para reconocer movimientos y con ello asociar a una expresión de la LSA.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto de investigación está integrado por docentes investigadores responsables de cátedras de las carreras del DI-FCEFYN y que son inherentes a las diferentes columnas que dan soporte al área de conocimiento de la Ciencia de Datos, el Machine Learning y Deep Learning. Estos Docentes investigadores fomentan constantemente la participación de alumnos que integran el espacio e incitan a aquellos que cursan esas cátedras a abrazar la temática e integrarse al proyecto. Durante el pasado 2022, se defendieron dos tesis de grado en Licenciatura en Ciencias de la Computación.

Actualmente se encuentra en instancia de evaluación por la Comisión de Licenciatura un trabajo final, desde donde el grupo publicó en la JAIIO 2022 particularmente en el Congreso Argentino de AgroInformática - CAI 2022. Así mismo, se lleva adelante y seguramente se defenderá en el primer semestre de 2023 el trabajo final de Licenciatura en Ciencias de la Computación.

A nivel de posgrado, se ha presentado el informe final de una tesis de maestría, en etapa de evaluación por el tribunal asignado para su ulterior defensa.

5. BIBLIOGRAFÍA

- [1] P. Singh, Ed., *Fundamentals and methods of machine and deep learning : algorithms, tools and applications*, 1st ed. USA, 2022.
- [2] S. Dash, S. K. Pani, J. Rodrigues, and B. Majhi, Eds., *Deep learning, machine learning and IoT in biomedical and health informatics : techniques and applications*, Fourth. Massachusetts, 2022.
- [3] I. El Naqa and M. J. Murphy, "What Is Machine Learning?," *Mach. Learn. Radiat. Oncol.*, pp. 3–11, 2015, doi: 10.1007/978-3-319-18305-3_1.
- [4] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning," 2015.
- [5] C. C. Aggarwal, *Neural networks and deep learning : a textbook*, 1st ed. Cham, Switzerland: Springer Nature Switzerland AG, 2018.
- [6] D. Mery, "Visión por Computador," Santiago de Chile, 2004.
- [7] Q. Ji, "Computer vision applications," *Probabilistic Graph. Model. Comput. Vis.*, pp. 191–297, Jan. 2020, doi: 10.1016/B978-0-12-803467-5.00010-1.
- [8] "Señario de términos y expresiones en Lengua de Señas Argentina – CAS." <https://cas.org.ar/senario-de-terminos-y-expresiones-en-lengua-de-senas-argentina/> (accessed Dec. 03, 2022).
- [9] M. I. Massone, "Lenguas de señas: 'cada comunidad desarrolló la propia por necesidad' | CONICET," 2012. Accessed: Nov. 11, 2022. [Online]. Available: <https://www.conicet.gov.ar/lenguas-de-senas-cada-comunidad-desarrollo-la-propia-por-necesidad/>.
- [10] F. Ronchetti, "Reconocimiento de gestos dinámicos y su aplicación al lenguaje de señas," Mar. 2017, doi: 10.35537/10915/59330.

Diseño de una estrategia para la gestión del tiempo en la comunicación de corrección de fallas en el mantenimiento industrial

Juan J. López. Mario M. Berón. Germán A. Montejano.

Universidad Nacional de San Luis

Ejército de los Andes 950 - <http://www.sel.unsl.edu.ar/lacis/index.html>

lopezjj@outlook.com {gmonte, mberon}@unsl.edu.ar

RESUMEN

Las organizaciones y los avances tecnológicos siguen demostrando que son un factor clave de la competitividad empresarial. La presente línea de investigación en Ingeniería del Software aplicado al sector industrial, es parte de una estrategia de alta disponibilidad gerencial bajo la filosofía de Industria 4.0.

En este artículo se describe una línea de investigación donde se diseña una estrategia para gestionar los tiempos de operación de la industria.

La tarea implica determinar la capacidad de adaptación de los componentes de la organización- personal, máquinas, activos, procesos, aplicaciones, señales y datos –, para abordar una transformación digital escritas en los Proyectos de Investigación de la UNSL.

Por medio de la estrategia diseñada y software, se va generando un impacto directo en la productividad de la empresa y en la eficiencia de sus flujos de trabajo.

Para la gestión de una prueba piloto se utilizó una metodología ágil para encontrar, especificar y modelar las entidades propias del dominio en UML que permitan brindar entendimiento del problema para el desarrollo de un sistema de información web online, donde visualizar la lógica del negocio mediante código de colores.

Palabras clave: Ingeniería del software aplicada, estrategia de mantenimiento, transformación digital, manufactura avanzada, fábrica inteligente o industria 4.0.

CONTEXTO

La presente línea de investigación se enmarca en el Proyecto de Investigación Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube. Director: Daniel Riesco y sus dos líneas de investigación: Línea 1: Desarrollo de Estrategias de Mantenimiento que Preserven la Calidad y Seguridad los Sistemas de Software que Ejecutan en la Nube. Director: Mario Berón y Línea 2: Desarrollo de Métodos Estrategias y Herramientas de Migración Preventivo de Sistemas de Software que Ejecutan en la Nube. Director: Germán Montejano. Pertenecientes a el Área de Programación y Metodologías de Desarrollo de Software y el Laboratorio de Calidad e Ingeniería de Software (LACIS) del Departamento de Informática de la Universidad Nacional de San Luis (UNSL): Proyecto N° P. 3-2020. CI N° 22/F022. Tal proyecto es la continuación de diferentes proyectos de investigación, a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además de una efectiva extensión al sector privado e industrial, donde la organización y sus necesidades de transformación serán caracterizadas en la presente línea de investigación.

Las pruebas de conceptos se implementaron físicamente en la Planta Industrial de Salsas y Aderezos (PSA) desde el área de mantenimiento del complejo Niza, S.A. del grupo AGD, Aceitera General Deheza. <https://www.agd.com.ar/es/niza-sa>.

1. INTRODUCCIÓN

El desarrollo industrial mundial ha incorporado metodologías, herramientas y técnicas informáticas en sus procesos con la finalidad de mejorar la competitividad.

Ha marcado etapas y avances en su historia, la segunda guerra mundial impulsó la industrialización, luego el consumo masivo de productos lo perfeccionó.

La generación de volúmenes de datos y su procesamiento entre los años 1970-1980 impulsó la ciencia de la computación y por ende la programación en informática.

La digitalización de procesos [1] es un aspecto muy importante en la actualidad, y obtuvo más relevancia con el efecto pre-post pandemia COVID-19 que lo ha impulsado vertiginosamente, el teletrabajo abrió la posibilidad de operar desde cualquier lugar del mundo con las sedes de las empresas conectadas a internet.

Es así que la ciencia de Ingeniería de Sistemas ha encontrado un nicho explotable dentro de la Ingeniería Industrial, llegando hasta conceptos actuales de manufactura avanzada, fábrica inteligente o industria 4.0 [1], para referirse a un conjunto de transformaciones productivas e institucionales compuesta por extensión de una serie de tecnologías digitales a la manufactura.

La Ingeniería del Software ha incrementado su campo de acción como ciencia aplicada al área del mantenimiento industrial [3], brindando un conjunto de métodos, procesos, técnicas y herramientas para desarrollar sistemas fiables que cumplan con los estándares impuestos por la industria, sobre todo en la digitalización de procesos [1] y analítica de datos principalmente, impactando positivamente en los tiempos de producción y satisfacción de los clientes, en la competitividad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La investigación centra su trabajo en el diseño de una estrategia de productividad industrial, tema identificado con las herramientas de los Proyectos de Investigación enunciados en la sección contexto.

La productividad del mantenimiento industrial es muy sensible al acto comunicacional entre áreas operativas (realizado de forma personal y manual) y en particular, cuando existen desvíos al plan de producción, ocasionando retrasos e inconvenientes en la cadena de valor.

Optimizar los procesos involucrados, repercutirá en el cumplimiento de cronogramas, entregas a tiempo y por consecuencia, sobre la utilidad de las empresas.

El proyecto consta de varias etapas, la primera consiste en establecer el problema evidenciando el impacto de la variable tiempo de parada productiva a través de la recuperación de la información.

Luego se planteará una recomendación para la ejecución del proyecto con base en el análisis del estado del arte sobre el tema y una revisión teórica bajo el marco de las metodologías ágiles con la que se realizará el diseño de la estrategia de sistemas para la gestión del proceso tiempos de comunicación.

Se diseñará una planificación de actividades en las que destacan diseño de arquitectura TI en cuanto a las configuraciones de hardware y software, descripción y análisis de procesos TI en cuanto a: flexibilidad, conectividad, calidad y seguridad, más la adaptabilidad y capacitación del personal operativo.

Se comprobará la factibilidad de la recomendación a través de una prueba piloto programando una aplicación web para evaluar la efectividad sobre el indicador comunicación de fallas correctivas, que actualmente ocupa en

promedio un 19% del consumo del tiempo productivo en planta.

Para el desarrollo de la aplicación web se considerará DJANGO, PHYTON y MySQL, con el sistema de control de versiones GIT y el proyecto alojado en GitHub.

En la intranet del negocio piloto se publicará la interfaz gráfica que necesitarán los equipos de operación industrial para visualizar el estado de los activos en planta, esta prueba contemplará una identificación visual simple de los estados posibles con base en un código de colores similares a las utilizadas en el tránsito vehicular, donde el color rojo simbolizará máquina o línea parada, el color verde máquina o línea en producción, el color azul máquina o línea en lavado CIP, el color amarillo máquina o línea en mantenimiento y finalmente el color blanco máquina o línea sin programa.

La extensión de la aplicación de visualización con una red de monitores en el piso de planta, asistirá al personal técnico ante correcciones de fallas, como principal beneficio, además de otros cambios para todos los recursos humanos claves del complejo industrial.

La interfaz presentará comunicación en línea, simultánea 24/7 de los sectores principales de la planta productiva; servicios, preparación de fases, producción, envasado y paletizado.

Con la finalidad de no saturar el personal operativo se utilizó como recurso un lenguaje visual sencillo en código de colores.

El entrenamiento hacia el personal consistirá en la información general del proyecto, procesos del negocio, flujos de trabajos objetivos y su importancia en la productividad de la planta.

El análisis de esta prueba piloto ofrecerá los ajustes necesarios para el diseño de la estrategia y así proyectarla a corto, mediano y largo plazo.

3. RESULTADOS OBTENIDOS/ESPERADOS

La mejora en las comunicaciones ante fallas correctivas [3] generará como beneficio la reducción en tiempos de reparación de activos en planta y, en consecuencia, mejores indicadores de productividad y entrega oportuna de pedidos.

Actualmente el proyecto se encuentra en proceso de documentación del impacto de la estrategia aplicada para la asistencia continua a la productividad. Mientras que otros procesos y recursos limitantes en los tiempos de entrega de las herramientas informáticas se encuentran en fase de exploración.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo contemplado en el proyecto consistió, además de integrantes del proyecto de investigación de Ingeniería de Software de la UNSL, en un coordinador de proyectos, tres programadores y un especialista en automatización.

Para su ejecución se requieren los siguientes recursos: adiestramiento del recurso humano técnico, internet, computadoras, red de comunicaciones, autómatas programables, sistemas de automatización, software, programas, licencias y capacitaciones del personal operativo.

La industria requiere de muchas aplicaciones para la optimización de proceso más allá de la robótica y sensores inteligentes en las máquinas.

La interacción humana como la comunicación descrita en este proyecto es un tema poco tratado y genera impacto en los costos de las empresas.

Por lo que se espera que más industrias y profesionales se adhieran a esta línea de investigación con sus tesis de grado y postgrado.

El autor cuenta con una tesina de grado aprobada y en curso de la tesis de maestría al cual pertenece el proyecto aquí descrito.

5. BIBLIOGRAFIA

- [1] J. Motta, H. Moreno y R. Ascúa, “Industria 4.0 en MyPyMEs manufactureras de la Argentina”, Documentos de Proyectos (LC/TS.2019/93), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL), 2019. Recuperado 4 de septiembre de 2022, de https://repositorio.cepal.org/bitstream/handle/11362/45033/1/S1900952_es.pdf
- [2] López, Andrés y Ramos, (2018). Adrián “El sector de software y servicios informáticos en la Argentina. Evolución, competitividad y políticas públicas”. CECE. Recuperado 4 de septiembre de 2022, de <http://fcece.org.ar/wp-content/uploads/informes/software-servicios-informaticos-argentina.pdf>
- [3] Denia A., José C. (2018). “Procesos y gestión del mantenimiento.” Recuperado 30 de agosto 2022 de <https://fdocuments.ec/document/mantenimiento-y-seguridad-industrial-2018-9-12-procesos-y-gestin-del-mantenimiento.html>
- [4] Centro de e-learning (2022). “Gestión ágil de proyectos.” Universidad Tecnológica Nacional. Facultad Regional Buenos Aires. Secretaría de cultura y extensión universitaria. Centro de formación, investigación y desarrollo de soluciones de e-learning.
- [5] Landajuela, Iker (2018) “Taller gestión de proyectos con metodologías ágiles SCRUM” Recuperado 16 de septiembre 2022 de <https://ikerlandajuela.wordpress.com/2018/05/25/material-taller-gestion-de-proyectos-con-metodologias-agiles-scrum/?blogsub=confirming#subscribe-blog>
- [6] Van Bon, J., De Jong, A., Kolthof, A., Pieper, M., Tjassing, R., Van deer Veen, A., & Verheijen, T. (2008). Fundamentos de ITIL V3. Amersfoort, Holanda: Van Haren Publishing. doi: 9789087530600.
- [7] Scarano, Eduardo R. (2004) “Manual de redacción de escritos de investigación”. Ediciones Macchi, Argentina.
- [8] Moreno, Paula M. (2018). “Plan de negocio de una asesoría tecnológica de la industria 4.0 para PyMEs”. Rescatado el 15 de agosto de 2020 de <https://oa.upm.es>
- [9] Albrieu, Ramiro, Basco, Ana, López, Caterina, Azevedo, Belisario, Peirano, Fernando Rapetti, Martín, Vienni, Gabriel. (2019). “Travesía 4.0: hacia la transformación industrial argentina”. Instituto para la integración de América Latina y el Caribe (INTAL). Recuperado el 15 de julio de 2020 de <https://repositorio.cedes.org/handle/123456789/4621>
- [10] Kruchten, Philippe (1995). “Architectural Blueprints-The “4+1” View Model of Software Architecture”. Rational Software Corp. Paper published in IEEE Software 12 (6) November 1995, pp. 42-50. Recuperado 16 septiembre 2022 de <https://arxiv.org/ftp/arxiv/papers/2006/2006.04975.pdf>
- [11] Ayala G., Yarol J. (2020) “Creación e implementación de un sistema de información aplicado al mejoramiento y control del mantenimiento de la maquinaria de la compañía industrias Ayala HNOS Y CIA SAS. Recuperado 4 de septiembre 2022 de <https://repository.usta.edu.co/handle/11634/33617>
- [12] Calatayud, Agustina, Katz, Raúl. 2019. “Cadena de suministro 4.0: mejores prácticas internacionales y hoja de ruta para América Latina”. Recuperado el 20 de marzo de 2021 de https://www.researchgate.net/publication/336662115_Cadena_de_suministro_40_Mejor

es_practicas_internacionales_y_hoja_de_rut
a_para_America_Latina.

- [13] Castellanos, Luis F., Sebastián, Cely, Cortes R., Camilo y Castillo L., Erickson F. (2020) Modelo de negocios aplicado en la Universidad Antonio Nariño sede Bogotá con tres productos: Sistema de información para la gestión de equipos y mantenimientos (preventivo y correctivo), mantenimiento predictivo para motores de inducción y plan de actualización de equipos. Recuperado 4 de septiembre de 2022, de <http://186.28.225.13/bitstream/123456789/3090/8/2020FelipeCastellanos.pdf>
- [14] Comisión Económica para América Latina y el Caribe (CEPAL). (2018). Datos, algoritmos y políticas: la redefinición del mundo digital (LC/CMSI.6/4), Santiago, 2018. Recuperado el 20 de febrero de 2020 de <https://repositorio.cepal.org/handle/11362/43477>
- [15] Hernández-Sampieri, Roberto y Mendoza T., Christian P. (2018). “Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta.” McGRAW-HILL Interamericana Editores, S.A. de C. V. Ciudad de México.
- [16] Inga L., Alex M., Tucumbi P., Williams R. (2021). “Aplicación web para el proceso de control de mantenimiento de hardware del GAD municipal de la ciudad de Latacunga. Ecuador. Recuperado 5 de septiembre 2022 de <http://repositorio.utc.edu.ec/bitstream/27000/8790/1/PI-002002.pdf>
- [17] Manifiesto Ágil. Recuperado 16 de septiembre 2022 de <http://agilemanifesto.org/iso/es/manifesto.html>
- [18] Menzinsky, Alexander, López, Gertrudis, Palacio, Juan, Sobrino, Miguel A., Álvarez, Rubén y Riva, Verónica (2022) “Historias de Usuario. Ingeniería de Requisitos Ágil” Scrum Manager. Recuperado 16 septiembre 2022 de https://scrummanager.net/files/scrum_manager_historias_usuario.pdf
- [19] Molina Ríos, Jimmy Rolando, Ordóñez, Mariuxi Paola Zea, Castillo, Fausto Fabián Redrován, Mora, Nancy Magaly Loja, Pardo, Milton Rafael Valarezo, Tapia, Joofre Antonio Honores. (2018). “Snail, una metodología híbrida para el desarrollo de aplicaciones web”. Recuperado el 20 de mayo de 2020 de <https://www.3ciencias.com>
- [20] Prendeville, Philip (2017). “Tailoring the agile approach to machine building: Facilitating improved communication across departments.” Recuperado 16 de septiembre 2022 de https://www.researchgate.net/figure/A-summary-of-Scrum-roles-activities-and-practices-adapted-from-Rubin-2012_fig5_317014052/download
- [21] Serrano Junco, Claudia Liliana. (2022). “Metodologías ágiles en las PyMEs: un modelo integral de auditoría en la gestión interna”. Bogotá: Corporación Universitaria Minuto de Dios - UNIMINUTO, 2022. Recuperado el 25 de marzo de 2022 de https://biblioteca-cum.hosted.exlibrisgroup.com/F/?func=direct&doc_number=102471&local_base=UNM01
- [22] Vrancken, Lisandro E. (2018). “Utilización de metodologías en el lanzamiento de negocios del sector TIC Santa Fe. Una propuesta basada en aplicación de mejores prácticas de gestión de proyectos.” Recuperado 30 de agosto 2022 de <https://bibliotecavirtual.unl.edu.ar:8443/handle/11185/1106>

Implementación de una BCI mediante el diseño y desarrollo de Serious Games para la rehabilitación de pacientes neurológicos

Javier J. Rosenstein, Rodrigo Gonzalez, Juan Salvador Portugal, Julian Argañaráz, Nicolás Ignacio Zarate

Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat, Belgrano 721, 5501 Godoy Cruz, Mendoza, Mendoza, Argentina.

rosensteinjavier@uch.edu.ar, gonzalezrodrigo@uch.edu.ar, portugaljuan@uch.edu.ar, arganarazjulian@uch.edu.ar, zaratenicolas@uch.edu.ar

RESUMEN

En el desarrollo de los sistemas de realidad virtual totalmente interactivos, es necesario contemplar: a) el desarrollo de la interfaz gráfica que se le presenta al usuario, y b) la adquisición de datos provenientes de los movimientos del mismo como entrada de datos del sistema de realidad virtual. El usuario es quien voluntariamente debe comandar esta interfaz digital, de modo que todo esto se convierta en una interacción ágil, bidireccional y produzca el efecto de virtualidad deseado, conformando un sistema usuario-máquina tan parecido a la realidad como sea posible. Para lograr mediante este sistema cumplir con los objetivos que se deseen obtener, ya sea el caso del sólo efecto lúdico de entretener al participante o con fines mucho más específicos como un complejo sistema de rehabilitación cognitiva en pacientes con capacidades físicas disminuidas, conocidos como “Juegos Serios” o “Serious Games”.

Los dispositivos de adquisición de datos generalmente están asociados a funciones o características de los individuos que los utilizan para la participación en el sistema y se necesita integrar los movimientos que estos representan hacia una interfaz gráfica que permita la realimentación neuronal del paciente para de este modo lograr la mejora cognitiva (efecto de neurofeedback deseado). El presente tra-

bajo trata de la implementación de una BCI – Brain Computer Interfase mediante la captura de datos de EEG/EOG y de otros dispositivos desarrollados *ad hoc* de los cuales se pueda realizar la adquisición de movimientos del usuario y entre las partes de un entorno multimedia permitir la representación visual en un escenario virtual que permita la retroalimentación al usuario en tiempo real logrando una experiencia interactiva e inmersa. Finalmente de este modo posibilitar la producción de un sistema de capacitación / rehabilitación o mejor llamado Serious Game o Juego Serio.

Palabras clave: EOG, EEG, BCI, Serious Games, Juegos Serios, Mirror Neurons, Neurofeedback

CONTEXTO

El presente proyecto se desarrolla en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat (Godoy Cruz, Mendoza), en el marco de la Licenciatura en Sistemas de Información; en cooperación con el Laboratorio de I+D+i en Neurotecnologías de la empresa Neuromed Argentina S.A. (Godoy Cruz, Mendoza).

Este trabajo es parte del proyecto de investigación que dio inicio en Septiembre de 2022 denominado “Diseño y desarrollo de prototi-

pos de Juegos Serios destinados a la rehabilitación de problemas físicos / neurológicos”.

1. INTRODUCCIÓN

El objetivo principal de la presente línea de investigación consiste en el diseño y desarrollo de una BCI (Brain Computer Interface) [1] [2], que permita interactuar entre las señales generadas por un paciente neurológico [3] y una interfaz de realidad virtual. Este sistema debe permitir lograr el principio de neurofeedback [4], o retroalimentación hacia el paciente. De este modo se podrán mejorar sus capacidades cognitivas correspondientes. Esto tiene aplicación directa en los tratamientos de rehabilitación en pacientes de patologías neurológicas y cognitivas, principalmente en niños, este enfoque es importante ya que está demostrado que es la mayor causa de discapacidad en niños, según estadísticas como las publicadas por el Indec de su informe sobre los resultados del censo de discapacidades en el 2018 (ver figura1), en argentina dentro de las principales discapacidades en menores, la mental-cognitiva es la mayor de las discapacidades entre las edades de 6 a 14 años.

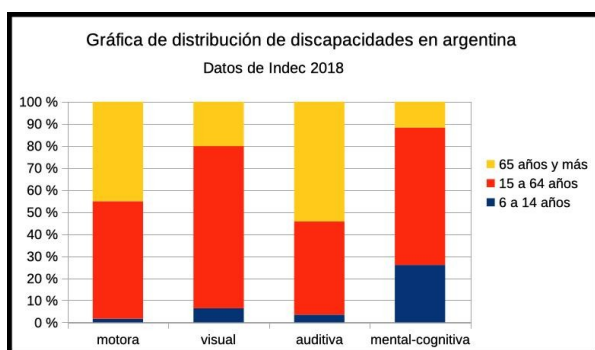


Figura 1: Discapacidades según Indec en argentina

Incluso, estudios indican la posibilidad de tratar patologías psiquiátricas como la depresión [5] [6].

La implementación del trabajo se organiza de acuerdo a las siguientes etapas:

- Adquisición de señales mediante técnicas de Electroencefalografía (EEG) [7] y adquisición de movimientos oculares mediante las técnicas de electrooculografía (EOG) [8] [9] [10] [11] [12], así como la adquisición de datos de otros dispositivos de adquisición desarrollados *ad hoc*.
- Análisis de estas señales en tiempo real para poder identificar la voluntad de movimiento del individuo, así como la dirección del movimiento.
- Transformar la voluntad de movimiento en comandos que permitan transmitir la información al componente software / hardware que la requiera.
- Como continuación al resultado obtenido en el punto anterior, Diseñar, programar y poder comandar una interfaz gráfica de aprendizaje o Juego Serio.

Finalmente se presenta la *Interfaz Cerebro / Computadora* (BCI – *Brain Computer Interface*) como un sistema de adquisición de datos, procesamiento de los mismos y luego la representación en nuestro modelo de prototipo de Serious Game [14] [15]. Una vez adquiridas las señales adquiridas estas se analizan y codifican para permitir su comunicación con las interfaces virtuales que interpreten este protocolo establecido.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo está compuesto por tres etapas o fases de trabajo. La primera de ellas corresponde a la implementación entre un sistema *simulado* de captura de datos provenientes de un paciente, y una interfaz virtual básica. La segunda etapa del proyecto pretende avanzar sobre la captura de datos *reales* para que, luego de procesados, se transfieran hacia la interfaz virtual generando el efecto neurofeedback deseado. Una breve descripción de cada una de las etapas se detalla a continuación:

1. La primer parte consiste en la implementación de un simulador de señales de adquisición que podrían ser del tipo EEG / EOG u otro dispositivo que se diseñe para la adquisición de muestras necesarias para el análisis e interpretación de la voluntad del usuario de la BCI. Estas señales una vez procesadas se deben codificar en comandos para poder ser transmitidas hacia una interfaz virtual. Esta debe poder interpretar las señales transmitidas y representar la voluntad inicial del usuario correspondiente. De este modo se cumplen los objetivos de captura, análisis, procesamiento, transmisión, recepción y representación, lo cual produce el efecto de neurofeedback deseado sobre una interfaz virtual de capacitación a nivel prototipo.
2. La segunda etapa consiste en el diseño y desarrollo del escenario virtual de rehabilitación cognitiva para pacientes neurológicos según las indicaciones concretas por parte del especialista en neurología. Se partirá de un relevamiento de las técnicas de aprendizaje que se requieren implementar y los resultados que se pretenden obtener, indicados por el neurólogo o experto afín.
3. La etapa final futura de este proyecto consistiría en el desarrollo y mejora del proyecto mediante la adquisición real de señales a) EEG/EOG por electrodos ubicados superficialmente sobre la cabeza del paciente o b) proveniente de dispositivos diseñados para la captura mediante otro

tipo de características que permitan adquirir la voluntad del paciente. Esta modificación al sistema requiere de un diseño e implementación electrónica así como del desarrollo del firmware que permita adquirir, analizar y preprocesar estas señales, que luego serán transmitidas hacia las interfaces virtuales de rehabilitación cognitiva de la etapa 2. Así, el sistema de adquisición real de datos se incorporará a futuro en forma transparente al proyecto implementado en las primeras 2 etapas.

3. RESULTADOS ESPERADOS

Los resultados esperados se pueden dividir en dos grandes grupos:

1. BCI e interfaz virtual gráfica que produzca en el paciente el efecto de neurofeedback, se espera obtener como producto final un sistema de retroalimentación a un usuario, en primera medida simulando la generación de señales que serán interpretadas en los escenarios de realidad virtual diseñados para este fin, estas señales se deberán enviar mediante comandos codificados que se transmitirán a través de una red Ethernet.
2. Adquisición de datos reales de EEG/EOG o dispositivos diseñados específicamente: análisis y preprocesamiento; se identifican los comandos necesarios para su transmisión hacia el prototipo funcional que implemente la técnica de la terapia y se logre el efecto real final de neurofeedback propuesto, cabe aclarar que para que esto puedan lograrse es necesario la interacción con equipos de trabajo de diseño y desarrollo de electrónica que colaboren en conjunto con el equipo de investigación formado.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de la tesis de maestría en teleinformática de la Universidad de Mendoza, de Javier J. Rosenstein, defendida a fines del 2020. Dicha tesis se centró en la implementación del protocolo VRPN demostrando su uso en un sistema BCI.

Esta línea de I+D Además cuenta con la dirección del Mgter. Javier J. Rosenstein (UCH) y como investigador principal el Dr. Rodrigo Gonzalez (UCH). En lo que respecta a la formación de estudiantes de la licenciatura, esta línea de investigación cuenta con tres tesis de grado en curso, pertenecientes a los estudiantes Juan Salvador Portugal, Julian Argañaraz y Nicolás Ignacio Zarate, cuyos planes de tesis se encuentran específicamente dentro del marco de este proyecto. Todos ellos han terminado de cursar la licenciatura en Sistemas de Información en la Universidad Champagnat.

Todos los avances logrados y las implementaciones realizadas relacionadas con el desarrollo del presente proyecto desde su primer etapa, se están utilizando como recurso para el dictado de talleres de comunicaciones, redes, programación de microcontroladores, programación en C/C++ y diseño y programación de interfaces virtuales de capacitación/rehabilitación en general, tanto para estudiantes de la universidad, como así también para alumnos externos a la UCH.

5. BIBLIOGRAFÍA

1. Y. Wang, X. Gao, B. Hong, and S. Gao, "Practical designs of brain-computer interfaces based on the modulation of EEG rhythms," in *Brain-Computer Interfaces*. Springer, 2009, pp. 137–154.
2. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
3. J. A. Pineda, "The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing"," *Brain Research Reviews*, vol. 50, no. 1, pp. 57–68, 2005.
4. S. Enriquez-Geppert, R. J. Huster, and C. S. Herrmann, "Boosting brain functions: Improving executive functions with behavioral training, neurostimulation, and neurofeedback," *International Journal of Psychophysiology*, vol. 88, no. 1, pp. 1–16, 2013.
5. R. Ramirez, M. Palencia-Lefler, S. Giraldo, and Z. Vamvakousis, "Musical neurofeedback for treating depression in elderly people." *Frontiers in neuroscience*, vol. 9, pp. 354–354, 2014.
6. W. Rief, "Getting started with neurofeedback," 2006.
7. J. D. Kropotov, *Quantitative EEG, event-related potentials and neurotherapy*. Academic Press, 2010.
8. A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 33, no. 4, pp. 741–753, 2011.
9. H. Singh and J. Singh, "A review on electrooculography," *International Journal of Advanced Engineering Technology*, vol. 3, no. 4, pp. 115–122, 2012.
10. D. P. Bautista, I. A. Badillo, D. De la Rosa Mejía, and A. H. H. Jiménez, "Interfaz humano-computadora basada en señales de electrooculografía para personas con discapacidad motriz," *ReCIBE*, vol. 3, no. 2, 2016.
11. S. Yathunathan, L. Chandrasena, A. Umakanthan, V. Vasuki, and S. Munasinghe, "Controlling a wheelchair by use of EOG signal," in *2008 4th International Conference on Information and Automation for Sustainability, IEEE*, 2008, pp. 283–288.
12. V. C. C. Roza, "Interface para tecnologia assistiva baseada em eletrooculografia," 2014.

13. A. C. Gaviria, I. C. Miller, S. O. Medina, and D. R. Gonzales, "Implementación de una interfaz hombre-computador basada en registros EOG mediante circuitos de señal mixta PSoC," in V Latin American Congress on Biomedical Engineering CLAIB 2011 May 16-21, 2011, Habana, Cuba. Springer, 2013, pp. 1194–1197.
14. P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: A survey and a classification towards a taxonomy," in 5th Iberian Conference on Information Systems and Technologies. IEEE, 2010, pp. 1–6.
15. J. S. Breuer and G. Bente, "Why so serious? On the relation of serious games and learning," *Eludamos. Journal for Computer Game Culture*, vol. 4, no. 1, pp. 7–24, 2010.
16. Instituto Nacional de Estadísticas y Censos, INDEC, Argentina, 2018, ISBN:978-950-896-519-6, https://www.indec.gob.ar/ftp/cuadros/poblacion/estudio_discapacidad_07_18.pdf
17. Rosenstein, Javier J. (2020). *Uso de VRPN para la implementación de una interfaz cerebro-computadora* [Tesis de Maestría en Teleinformática, Universidad de Mendoza].

La Inteligencia Artificial y su aplicación a los servicios de Internación Domiciliaria

Poch Miguel ¹, Rotella Carina¹, Ontiveros Patricia¹, Tagarelli Sandra^{1,2}, Caymes-Scutari Paola^{3,4}, Bianchini Germán³

¹Laboratorio de Gobierno Electrónico, Departamento de Ingeniería en Sistemas de Información. Facultad Regional Mendoza/Universidad Tecnológica Nacional

Rodríguez 273 (M5502AJE) Mendoza

²Laboratorio de Analítica de Datos, Departamento de Ingeniería en Sistemas de Información. Facultad Regional Mendoza/Universidad Tecnológica Nacional

Rodríguez 273 (M5502AJE) Mendoza

³Laboratorio de Investigación en Cómputo Paralelo/Distribuido

Departamento de Ingeniería en Sistemas de Información

Facultad Regional Mendoza/Universidad Tecnológica Nacional

Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

⁴Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

miguel.poch@docentes.frm.utn.edu.ar, crotella@frm.utn.edu.ar, pontiveros@frm.utn.edu.ar, stagarelli@frm.utn.edu.ar, pcaymesscutari@frm.utn.edu.ar, gbianchini@frm.utn.edu.ar

RESUMEN

En la presente línea de investigación se busca analizar los antecedentes teóricos y experiencias prácticas registradas en el campo de la Inteligencia Artificial (IA) aplicada a los servicios de Internación Domiciliaria (Home Care), para tratar de generar un enfoque integral que potencie la utilización de los nuevos dispositivos, la IA y los servicios médico asistenciales en domicilio.

La salud está fuertemente afectada por la Inteligencia Artificial y las Herramientas 4.0. La Internet de las Cosas, el Big Data, el Data Mining, la Robótica y la Impresión 3D son solamente algunas de las utilizadas, en especial para el diagnóstico y la prevención.

Sin embargo, hay pocas experiencias de Inteligencia Artificial enfocadas a la

Internación Domiciliaria como disciplina integral. Se avanzó sobre el equipamiento y el hogar inteligente, pero en mucha menor medida sobre el desarrollo de este equipamiento para los domicilios y en la prestación médico asistencial (médicos, kinesiología, enfermería, fonoaudiología, terapia ocupacional, etc.).

Palabras clave: Internación Domiciliaria, Inteligencia Artificial, Cuidados Domiciliarios, Home Care, Smart Home, Enfermería.

CONTEXTO

Este trabajo se contextualiza en el Proyecto PID TEUTIME0007658TC, homologado y acreditado por la Universidad Tecnológica

Nacional, en el cual se busca incentivar las actividades de investigación científica en docentes y estudiantes de grado, a través de equipos interdisciplinarios.

En particular, la presente línea de trabajo explora la necesidad de contar con equipos formados en la implantación de nuevas tecnologías, recursos que puedan aprovecharlas y la relación de los pacientes con estas innovaciones.

1. INTRODUCCIÓN

La Internación Domiciliaria consiste en una conjunción de RRHH, equipamiento y estudios diagnósticos dedicados a la atención de salud de personas que no pueden desplazarse a los centros de salud, o, por su condición, no justifica el estar internado en una clínica. Se aplica a todo tipo de pacientes, pero especialmente a la pediatría, la ancianidad, y a quienes requieren de cuidados paliativos [1] [2] [3].

Por sus grandes ventajas desde el tipo económico, social y de satisfacción del paciente y su grupo familiar, tuvo un crecimiento exponencial en el siglo XXI, siendo una solución para los financiadores de salud (Estado y privados). Solamente como medida podemos decir que, en 2011, aproximadamente dos millones de personas o el 5.6% de la población anciana de Medicare (USA) que vivía en la comunidad estaban confinadas a su casa [4].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se presentan estados de avance de la IA en distintos puntos de interés: Nuevos Dispositivos Médicos (Internet de las Cosas, Diagnóstico a Distancia, etc.) y Enfermería (base de los cuidados domiciliarios).

Dispositivos Médicos en Domicilios:

La revolución de Internet de las Cosas (IoT) es la disciplina que generó que objetos de cualquier índole (electrónicos o no) puedan

servir, a través de sensores incluidos en ellos y conectados a través de la nube, para medir valores que antes solamente se hacía a través de artefactos específicos para tal fin. Zapatillas que miden la velocidad, sensores para medir la humedad de un campo deportivo y regarlo cuando hace falta, o collares para mascotas que indican su ubicación en todo momento, pasaron de ser una quimera a una realidad.

En medicina, los cardiodesfibriladores implantables o los medidores continuos de glucosa son los primeros que se utilizaron para “independizar” a los pacientes de los centros asistenciales. Pero no fue sino hasta mediados de la segunda década del siglo XXI que se comenzó a utilizar el inmenso caudal de datos obtenidos de los sensores, equipos y dispositivos para trabajar en el aprendizaje automático [5].

Uno de los problemas que genera una internación domiciliaria es la dificultad que se tiene de trasladar al paciente para cada estudio que se le deba efectuar. Cada salida del domicilio implica una logística problemática, riesgosa y antieconómica. El desafío de la industria de los dispositivos médicos está en poder coordinar la toma de valores en un domicilio, con la transmisión de los datos de los sistemas de detección, análisis y predicción de enfermedades que utilicen la Inteligencia Artificial para tal fin.

Enfermería y Cuidado Domiciliario:

El rol de la enfermería en los cuidados domiciliarios es fundamental, ya que es el responsable del cuidado durante las 24 horas. Muchas de sus funciones básicas son las mismas que en un centro asistencial (control de signos vitales, higiene y confort, administración de medicamentos, realización de los procedimientos de su responsabilidad o de ayuda a los médicos, etc.). Sin embargo, en el domicilio cuenta con más tareas: la relación con el paciente y la familia, la contención, la higiene de su lugar, y (en especial) el monitoreo continuo del estado del paciente. El estudio de cualquier cambio de actitud, físico,

emocional o cognitivo puede significar un síntoma de alguna dificultad [6].

Como hemos señalado, un hogar inteligente es una ayuda inconmensurable para el paciente y la familia, y lo es también para la enfermería. El estudio y control permanente de un paciente en forma personal sin aparatología en guardias de 6 u 8 horas es una tarea quimérica e ineficaz. El buen manejo de sensores y alarmas (sensores de oximetría, de pulso, etc.) previene la ocurrencia de eventos con difícil resolución posterior[7].

Se empieza a aplicar la robótica para las tareas de asistencia o asistencia social. Entre los de asistencia (tareas más físicas) encontramos a los robots de servicio (movilidad, monitoreo), de ayuda a la movilidad, para servir y alimentar, portadores u otros (bañeras robotizadas, tutores de ejercicios, etc.). Entre los de asistencia social encontramos a los de cuidado emocional y los de atención cognitiva [8]. Estas innovaciones generan una mejora en la calidad de vida, con menores reingresos en instituciones de salud.

Dermody y otros investigadores resaltan la importancia de incorporar el conocimiento clínico de enfermería en hogares inteligentes, así como en características de inteligencia artificial. Según sus hallazgos, esta integración resulta fundamental para mejorar la calidad de atención y cuidado de los pacientes en el hogar. [9]. En los sistemas de automatización del hogar, la Inteligencia Artificial cumple múltiples funciones, tales como servir como una base de datos de conocimientos y reglas, tomar decisiones, implementar acciones y controlar dispositivos. En este sentido, su capacidad para procesar grandes cantidades de información y adaptarse a situaciones cambiantes resulta esencial para lograr un hogar más inteligente y conectado [10].

Para que una Inteligencia Artificial sea eficiente debe tener un alto volumen de información. Por un lado, información externa, de múltiples casos y, por el otro, del caso en particular. Para ello, el personal debe estar capacitado para cargar síntomas, alteraciones,

y demás eventos, para que el sistema pueda aprender en forma automática [11].

La buena utilización de chatbots provistos de IA son otra de las herramientas que están evolucionando. La función primordial es la mejora en la toma de decisiones, ya que un evento puede producirse en cualquier horario del día, cuando se tienen recursos de consulta o cuando no, y las primeras maniobras son las fundamentales para el buen desarrollo del caso.

3. RESULTADOS OBTENIDOS / ESPERADOS

En el presente trabajo se persigue que ambas líneas de estudio, la Inteligencia Artificial y la Internación Domiciliaria, converjan en una sola. En primer lugar, se pretende alcanzar una idea clara sobre la necesidad de dispositivos especiales para la aplicación de Inteligencia Artificial en las internaciones domiciliarias. La gran mayoría son dispositivos existentes diseñados para instituciones sanatorias. Su primera característica es que deben contar con facilidad de traslado con una alta capacidad para recoger valores por medio de sensores o ingreso de datos. Su otra condición importante es que tengan capacidad para conectarse en forma automática a unidades de procesamiento de Inteligencia Artificial para que comparen contra experiencias generales y del propio paciente. De esta manera, los recursos humanos afectados al cuidado del paciente en el domicilio podrían tener la información suficiente y a tiempo para la toma de decisiones. Por otro lado, se debe reformular la capacitación de los recursos humanos del personal dedicado a las tareas de cuidados domiciliarios en general, y de enfermería en particular. Los cambios tecnológicos que se van a ir produciendo en los próximos años (muy cercanos) van a poner de manifiesto que los planes de formación de enfermeros no priorizan los cambios tecnológicos ni la utilización de herramientas informáticas. Según lo publicado por el Ministerio de Educación de la República Argentina, en el

ámbito universitario, se pueden encontrar más de cien programas de grado y posgrado disponibles [7]. Analizando los planes de estudios de la carrera de algunas de las Universidades Estatales (UBA, UNLP, UNC, UNL, UNT, UNMDP, UNSa, UNCo, UNSL), podemos verificar que no se tratan temas tecnológicos y/o informáticos. La formación sobre estos temas debería ir delineándose, pudiéndose aprovechar las mismas estructuras de IA existentes para tal fin.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente proyecto se enfoca en contribuir a la formación y desarrollo de investigadores nóveles en temáticas vinculadas a la Inteligencia Artificial. Dos de los docentes-investigadores en la línea específica de IA aplicada a la salud, y otro de los docentes en la rama de IA aplicada a la Internación Domiciliaria. El proyecto se contará con la participación de 4 alumnos de la carrera de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

[1] Goswami, S. HOME BASED PALLIATIVE CARE, SUGGESTIONS FOR ADDRESSING CLINICAL AND NON-CLINICAL ISSUES IN PALLIATIVE CARE, Jul. 2021, doi: 10.5772/intechopen.98648.

[2] <https://www.hopkinsmedicine.org/health/care/giving/pediatric-home-health-care-what-you-need-to-know>. Accedido 3-3-2023

[3] Miskelly, FG. ASSISTIVE TECHNOLOGY IN ELDERLY CARE, *Age and Ageing*, Volume 30, Issue 6, November 2001, Pages 455–458, <https://doi.org/10.1093/ageing/30.6.455>

[4] Ornstein, KA; Leff, B; Covinsky, KE; Ritchie, CS; Federman, AD; Roberts, L; Kelley, AS; Siu, AL; Szanton, SL.

EPIDEMIOLOGY OF THE HOMEBOUND POPULATION IN THE UNITED STATES. *JAMA Intern Med.* 2015 Jul;175(7):1180-6. doi: 10.1001/jamainternmed.2015.1849. Erratum in: *JAMA Intern Med.* 2015 Aug;175(8):1426. PMID: 26010119; PMCID: PMC4749137.

[5] Badnjević, A., Avdihodžić, H. i Gurbeta Pokvić, L. (2021). ARTIFICIAL INTELLIGENCE IN MEDICAL DEVICES: PAST, PRESENT AND FUTURE. *Psychiatria Danubina*, 33 (suppl 3), 101-106. Preuzeto s <https://hrcak.srce.hr/257613>

[6] Maalouf, N; Sidaoui, A; Elhajj, IH; Asmar, D. ROBOTICS IN NURSING: A SCOPING REVIEW. *J Nurs Scholarsh.* 2018 Nov;50(6):590-600. doi: 10.1111/jnu.12424. Epub 2018 Sep 27. PMID: 30260093.

[7] Castro, D., Coral, W., Cabra, J., Colorado, J., Méndez, D., & Trujillo, L. (2017). SURVEY ON IOT SOLUTIONS APPLIED TO HEALTHCARE. *DYNA*, 84(203), 192-200.

[8] Dermody, G; Fritz, R. A CONCEPTUAL FRAMEWORK FOR CLINICIANS WORKING WITH ARTIFICIAL INTELLIGENCE AND HEALTH-ASSISTIVE SMART HOMES. *Nurs Inq.* 2019 Jan;26(1):e12267. doi: 10.1111/nin.12267. Epub 2018 Nov 12. PMID: 30417510; PMCID: PMC6342619.

[9] Guo, X.; Shen, Z.; Zhang, Y.; Wu, T. REVIEW ON THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN SMART HOMES. *Smart Cities* 2019, 2, 402-420. <https://doi.org/10.3390/smartcities2030025>

[10] Dermody, G; Fritz, R. A CONCEPTUAL FRAMEWORK FOR CLINICIANS WORKING WITH ARTIFICIAL INTELLIGENCE AND HEALTH-ASSISTIVE SMART HOMES. *Nurs Inq.* 2019 Jan;26(1):e12267. doi:

10.1111/nin.12267. Epub 2018 Nov 12.
PMID: 30417510; PMCID: PMC6342619.

[eria-la-carrera-que-eligen-43000-argentinos.](#)
Consultado el 02/03/2023

[11]
<https://www.argentina.gob.ar/noticias/enferm>

ABORDAJES CIENTÍFICO TECNOLÓGICOS BASADOS EN EL USO DE VANTS

Russo C., Serafino S., Cicerchia B., Merino M., Guiguet M.
Adorno S., Ramón H. D.

Instituto de Investigación y Transferencia en Tecnología (ITT)

Comisión de Investigaciones Científicas (CIC)

Escuela de Tecnología (ET)

Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento Nro. 1169 2° Piso, Junín (B) – TE: (0236) 4407750 INT 11610

{claudia.russo, sandra.serafino, lucas.cicerchia, marcelo.guiguet, sebastian.adorno,

hugo.ramon}@itt.unnoba.edu.ar

mariano.merino@nexo.unnoba.edu.ar

RESUMEN

Las especies exóticas invasoras son responsables por grandes pérdidas económicas en muchos sectores, a nivel mundial. Sin embargo, hay muy pocos estudios que evalúen el impacto de las mismas en Sudamérica. Entre los impactos identificados, el jabalí y los cerdos cimarrones provocan cambios en los procesos ecosistémicos debido a alteraciones en el ambiente, daños en suelos por efecto de las hozadas, pérdida de biodiversidad, predación de semillas y dispersión de plantas exóticas. El objetivo de este proyecto es el desarrollo multidisciplinario de un sistema de evaluación de daño y costo económico producido por jabalíes y cerdos cimarrones sobre un agroecosistema. La implementación del mismo incluye el uso de un dron así como la aplicación de técnicas de visión artificial para realizar una evaluación global y registro geoespacial de la zona de análisis como para la aplicación posterior de procesamiento digital de imágenes sobre los videos y fotografías captadas por cámaras de luz visible, infrarroja y térmica.

Palabras clave: VANT, PROCESAMIENTO DE IMÁGENES, AGROECOSISTEMAS.

Las líneas de investigación presentadas a continuación se enmarcan en el proyecto de

investigación "Abordajes científico tecnológicos basados en el uso de vants y visión artificial para la evaluación del impacto de jabalíes y cerdos cimarrones en agroecosistemas", con lugar de trabajo en el Instituto de Investigación y Transferencia en Tecnología (ITT), presentado en la convocatoria a Proyectos de investigación interdisciplinarios unnoba 2022, aprobado y financiado por la Secretaría de Investigación, Desarrollo y Transferencia (SIDT) de la UNNOBA.

Como hipótesis de trabajo se plantea que el manejo del relevamiento de especies exóticas invasoras son plausibles de ser abordadas desde el monitoreo mediante el uso de tecnologías de VANT y técnicas de procesamiento digital de imágenes. Mediante la producción de conocimiento científico-tecnológico, se espera estimar y elaborar medidas de mitigación de los impactos económicos que provocan los jabalíes y cerdos cimarrones sobre la producción agrícola. Elaborando soluciones agrotecnológicas que puedan ser replicadas en distintos establecimientos de la provincia de Buenos Aires y que ayuden en el área estratégica de la agroindustria.

1. INTRODUCCIÓN

Las especies exóticas son aquellas que no son nativas del país (en este caso Argentina) a la que llegaron de manera intencional o accidental, generalmente como resultado de actividades humanas. Las especies exóticas que se establecen en un nuevo sitio, y se reproducen y dispersan sin control, causando daños al ecosistema, a las especies nativas, a la salud o a la economía, son llamadas especies exóticas invasoras (EEI).

Las especies exóticas invasoras son responsables por grandes pérdidas económicas en muchos sectores, a nivel mundial. Sin embargo, hay muy pocos estudios que evalúen el impacto de las mismas en Sudamérica. La investigación del impacto de las especies invasoras es importante para motivar y orientar las respuestas políticas, aumentando la conciencia de las partes interesadas e identificando las prioridades de acción.

En Argentina existen una treintena de especies de mamíferos exóticos invasores entre ellos se encuentra el jabalí y los cerdos cimarrones (cerdos domésticos asilvestrados y/o formas híbridas con jabalíes), todos pertenecientes a la especie *Sus scrofa*.

Actualmente las poblaciones antes mencionadas ocupan una amplia área del país y están consideradas entre las 100 exóticas invasoras más dañinas a nivel mundial (Lowe et al., 2000) y la de mayor impacto en nuestro país (Novillo y Ojeda 2008). Su éxito invasor puede ser atribuido, entre otras causas, a sus hábitos alimenticios omnívoros, su adaptación a gran variedad de climas, la falta de depredadores naturales y su elevada capacidad reproductiva.

Entre los impactos identificados, el jabalí y los cerdos cimarrones provocan cambios en los procesos ecosistémicos debido a alteraciones en el ambiente, daños en suelos por efecto de las hozadas, pérdida de biodiversidad, predación de semillas y dispersión de plantas

exóticas. Además, compite por los recursos con las especies nativas.

En términos económicos, entre los impactos negativos se registran daños en la producción ganadera por invasión en las explotaciones, daños en cultivos y acopios, daños en infraestructura e importantes daños por accidentes viales en la mayoría de los lugares donde existen poblaciones naturalizadas (Cuevas et al., 2012; Barrios-García y Ballari, 2012).

En los últimos años el efecto de los jabalíes y cerdos cimarrones sobre las actividades agropecuarias se ha incrementado llegando a la tapa de los diarios tanto regionales como nacionales. Comenzando a preocupar a los productores que en algunos casos emprendieron algunos métodos de control para reducir la densidad de este mamífero en sus cultivos y así reducir sus pérdidas. También se han reportado daños en ganadería sobre todo en las explotaciones ovinas.

En el estudio del manejo de las especies exóticas invasoras se requiere el monitoreo continuo de su dinámica poblacional y cómo ésta impacta sobre las actividades humanas y sus costos asociados.

Tradicionalmente el monitoreo de las poblaciones de jabalíes y cerdos cimarrones, se realizaba con distintos métodos de muestreo en campo, entre los que destacan: el conteo directo de animales en líneas de marcha o transectas, la captura-recaptura en diferentes tipos de trampas, redes y cámaras-trampa; además, ha sido tradicional el empleo de métodos que involucran el rastreo para contar huellas, excrementos y/o cualquier otra evidencia de la presencia de las especies de interés (Sutherland, 2006).

El desarrollo de nuevas tecnologías permite abordar diferentes problemáticas de la biología aplicada al desarrollo de medidas de manejo de fauna; esto es especialmente vital en el caso de las EEI que afectan las producciones

agropecuarias. Un caso paradigmático reciente es el uso de los vehículos aéreos no tripulados (VANTs) que ha posibilitado la realización de múltiples tareas en el ámbito rural. Los comúnmente llamados drones han tenido un gran avance últimamente, son aeronaves pequeñas que pueden controlarse fácilmente desde un teléfono inteligente y son capaces de portar cámaras u otros dispositivos y sensores eléctricos, razón por la que se les usa en un sin número de proyectos científicos y aplicaciones comerciales.

El uso de los drones contribuye a la obtención de datos espaciales en un corto periodo de tiempo y con una alta resolución espacial a un costo reducido. Previo a su aparición, todo dependía de la disponibilidad de los satélites, de aviones tripulados o de la cartografía realizada recorriendo el campo. Como ejemplo de la adopción de los drones como nuevas herramientas, podemos mencionar a la agricultura de precisión la cual utiliza tecnologías de posicionamiento global satelital (GPS), sensores láser en tierra para medir el nivel del terreno e imágenes aéreas para evaluar las diferentes variables de cada determinada parcela.

Los drones como plataforma de toma de fotografías tienen múltiples aplicaciones, abarcando diversos campos del conocimiento; el medio ambiente, la agricultura, las actividades industriales, el urbanismo, etc.

Durante los últimos años ha habido un aumento en la aplicación de nuevas tecnologías que involucran la utilización de Vehículos Aéreos No Tripulados (VANT o drones) para el estudio, monitoreo y manejo de la fauna silvestre, especialmente en mamíferos terrestres medianos y grandes (e.g. Hodgson et al., 2013; Rutten et al., 2018; Pereira et al., en prensa). El uso de esta tecnología está creciendo rápidamente y provee datos de detección remota a escalas temporales y espaciales muy precisas, que nos permite

plantear y contestar preguntas que un tiempo atrás hubiera sido más difícil responder.

Además de los trabajos de identificación y demografía en especies animales, los estudios enfocados a la caracterización y mapeo de hábitats son cada vez más abundantes, con una aproximación de índole ecológica (Barasona et al., 2014).

El manejo de las EEI requiere de monitoreos continuos que nos permitan conocer los cambios espaciales y temporales en la abundancia de las poblaciones de estas especies, así como también los posibles factores que pudieran afectarles como las producciones agropecuarias, (Mandujano et al., 2017).

La aparición de los drones, como una herramienta innovadora para el monitoreo aéreo de fauna es prometedora y su uso en investigación de vida silvestre está creciendo de manera vertiginosa.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto se centra en la evaluación del impacto de una EEI, especie *Sus Scrofa*, representada por el jabalí y los cerdos cimarrones, sobre la actividad agropecuaria con pérdidas que actualmente desconocemos, pero que a vista de los productores podría alcanzar al 20%. Lo que implicaría pérdidas por millones de USD en granos, pérdida de calidad de suelo y daños en infraestructura. Por consiguiente buscamos evaluar las pérdidas en las actividades agropecuarias con la utilización de nuevas tecnologías de adquisición y procesamiento de imágenes, basadas en drones.

La utilización de drones con estos objetivos es un desarrollo totalmente inédito en el país, casi con nulo desarrollo en Sudamérica (Deferrari et al. 2020, Schroeder et al 2020). Se cuentan con algunas experiencias en Norteamérica y

Europa, (Mandujano et al., 2017; Rutten et al., 2018) en este último caso en superficies pequeñas en comparación con las producciones de la zona núcleo maicera y la depresión del Salado.

En cuanto a la evaluación económica de las pérdidas producidas por jabalíes y cerdos cimarrones en el país, actualmente son de carácter muy preliminar ya que están basadas en datos extrapolados de otras situaciones de manejo fuera del continente americano (Heringer et al., 2021). Esto se debe a la falta de información biológica y ecológica sobre muchos aspectos de esta especie, los cuales no llevan más de una década desde su inicio de análisis.

La estimación del impacto mediante las imágenes tomadas por los drones y la estimación de los costos económicos de este impacto, permitirá dentro de un manejo adaptativo proponer medidas de mitigación de los mismos, así como el desarrollo de protocolos específicos para su evaluación. Es importante destacar que la participación de los productores es fundamental para este proceso, ya que son ellos los que primariamente soportan el impacto negativo de esta EEI sobre sus producciones.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto está compuesto por cuatro nodos, el biológico, el agronómico, el tecnológico y el económico, todos están trabajando en aspectos productivos anclados al territorio de la universidad. Este proyecto por primera vez aglutina a estos nodos, que si bien tienen una historia de trabajo en común entre algunos de ellos, no se habían congregado para resolver una problemática puntual como la que contempla este proyecto, de alcance regional y

con proyección a las principales áreas agrícolas del país.

La problemática de las EEI y en particular la de los jabalíes y cerdos silvestres, está en franco crecimiento, siendo un tema recurrente en los medios de comunicación, con múltiples quejas de los productores por pérdidas que actualmente no están evaluadas; pero que aparecen como importantes en algunos cultivos.

A través de este proyecto nos proponemos evaluar este impacto, a fin de colaborar con la búsqueda de soluciones para esta problemática que en menos de una década puede hacer impacto en una de las principales actividades económicas del país, con serias pérdidas tanto económicas como ambientales.

La temática planteada por el presente proyecto es inédita, con antecedentes escasos y basados en datos provenientes de poblaciones de otras regiones. De llevarse a cabo el proyecto sería de vanguardia en un aspecto poco desarrollado en el mundo, y muy sensible para la producción agrícola del país, motor de las exportaciones nacionales.

4. FORMACIÓN DE RECURSOS HUMANOS

Se espera formar recursos humanos interdisciplinarios con una visión global de la problemática planteada y la solución propuesta. En particular en el nodo tecnológico se capacitaron investigadores en el uso de VANTs, y se forman otros específicamente en el uso de herramientas de visión artificial y procesamiento digital de imágenes. El equipo de trabajo está compuesto por investigadores formados y en proceso de formación, becarios del Sistema Nacional de Ciencia y Tecnología, becarios alumnos, graduados e investigadores externos.

Se han presentado planes de trabajo para becas, las cuales se encuentran actualmente en proceso de evaluación.

Se esperan finalizar, además, tres tesis doctorales, y una tesis de especialización.

5. BIBLIOGRAFÍA

1. Barasona J.A., Mulero-Pázmány M., Acevedo P., Negro J.J., Torres M.J., Gortázar C., Vicente, J. 2014. Unmanned Aircraft Systems for studying spatial abundance of ungulates: relevance to spatial epidemiology. *PLoS One* 9:e115608.
2. Barrios-García, M.N. y Ballari, S.A. (2012). Impact of wild boar (*Sus scrofa*) in its introduced and native range: a review. *Biol. Invasions* 14(11): 2283-2300. 27-36.
3. Cuevas, M.F., Mastrantonio, L., Ojeda, R.A. and Jaksic, F.M. (2012). Effects of wild boar disturbance on vegetation and soil properties in the Monte Desert, Argentina. *Mamm. Biol.* 77: 299-306.
4. Deferrari G.A., J.M. Escobar, J. Ponce y D.Quiroga.2020. Primeras pruebas del uso de vehículos aéreos no tripulados (VANTs) como herramienta de evaluación de impacto de castores en Tierra del Fuego (República Argentina). *Notas sobre Mamíferos Sudamericanos* Vol 2. <http://doi.org/10.31687/saremNMS.20.0.31>
5. Heringer G, Angulo E, Ballesteros-Mejia L, Capinha C, Courchamp F, Diagne C, Duboscq-Carra VG, Nuñez MA, Zenni RD (2021) The economic costs of biological invasions in Central and South America: a first regional assessment. In: Zenni RD, McDermott S, García-Berthou E, Essl F (Eds) *The economic costs of biological invasions around the world. NeoBiota* 67: 401–426. <https://doi.org/10.3897/neobiota.67.59193>
6. Hodgson JC et al. 2018 Drones count wildlife more accurately and precisely than humans. *Methods Ecol. Evol.* 9, 1160–1167. (doi:10.1111/2041-210X.12974)
7. Mandujano, s., M. C. Mulero Pazmany, y A. Rísquez-Valdepeña. 2017. Drones: una nueva tecnología para el estudio y monitoreo de fauna y hábitats. *Agroproductividad* 10:79–84.
8. Novillo, A. y Ojeda, R.A., 2008. The exotic mammals of Argentina. *Biological Invasions*, 10 (8): 1333-1344.
9. Rutten, A., J. Casaer, M. F. Vogels, E. A. Addink, J. Vanden Borre, and H. Leirs. 2018. Assessing agricultural damage by wild boar using drones. *Wildlife Society Bulletin* 42:568–576.
10. Schroeder NM, Panebianco A, Gonzalez Musso R, Carmanchahi P. 2020 An experimental approach to evaluate the potential of drones in terrestrial mammal research: a gregarious ungulate as a study model. *R. Soc. open sci.* 7: 191482. <http://dx.doi.org/10.1098/rsos.191482>
11. Sutherland, W.J. 2006. Predicting the ecological consequences of environmental change: a review of the methods. *Journal of Applied Ecology*, 43, 599–616

INNOVACIÓN TECNOLÓGICA A TRAVÉS DE LA HIPER AUTOMATIZACIÓN

Russo C. , Serafino S. , Cicerchia B. , Alvarez E., Luengo P.3, Useglio G. , Di Cicco A., Charne J.,
Guiguet M., Pérez G., Cintora F., Argento F.,
Instituto de Investigación y Transferencia en Tecnología (ITT)
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)
Sarmiento Nro. 1169 2° Piso, Junín (B) – TE: (0236) 4407750 INT 11610

{ claudia.russo, sandra.serafino, lucas.cicerchia, eduardo.alvarez, pablo.luengo, gustavo.useglio,
carlos.dicicco, javier.charne, marcelo.guiguet, gabriel.perez}@itt.unnoba.edu.ar
fedecintora@gmail.com, facundoargento@comunidad.unnoba.edu.ar

RESUMEN

Las nuevas estrategias tecnológicas deben ayudar a las organizaciones a ser más ágiles en la toma de decisiones, y más sostenibles. En el mismo sentido ayudar a que logren una mejor adaptación a las circunstancias, gracias a las capacidades de automatización y digitalización de procesos operativos. La innovación tecnológica en equipos que permiten digitalizar todo tipo de señales y que se viene generando en las últimas décadas ha desatado una nueva ola de técnicas de procesamiento digital asociadas. Al mismo tiempo, los problemas también han evolucionado debido a la globalización de la información existente en relación a los mismos, su entorno y la transversalidad de datos de otras áreas de interés que los impactan. Los problemas que no plantean soluciones desde la tecnología, hoy son potenciales candidatos a utilizarla. Incluidos problemas que tienen incorporadas soluciones tecnológicas en sus procesos requieren de nuevas alternativas basadas en un mayor nivel de automatización también aplicada ahora a sus estrategias de negocios y toma de decisiones. Ésta propuesta propone identificar, evaluar, y abordar problemas interdisciplinarios con soluciones basadas en técnicas de hiper automatización.

Palabras clave: HIPER AUTOMATIZACIÓN. PROCESAMIENTO DE SEÑALES. IMÁGENES.

CONTEXTO

Las líneas de investigación presentadas a continuación se enmarcan en el proyecto de investigación "Innovación tecnológica a través de la hiper automatización", con lugar de trabajo en el Instituto de Investigación y Transferencia en Tecnología (ITT), presentado en la convocatoria a Subsidios de Investigación Bianuales (SIB) 2022, aprobado y financiado por la Secretaría de Investigación, Desarrollo y Transferencia (SIDT) de la UNNOBA.

El estudio, análisis, diseño y desarrollo de nuevas herramientas tecnológicas es parte activa de nuestro trabajo como docentes e investigadores en el marco del Instituto de Investigación y Transferencia en Tecnología de la UNNOBA, teniendo como objetivos aportar y transferir innovación tecnológica a diferentes tipos de problemas fundamentalmente de impacto regional. Particularmente se viene trabajando sobre áreas como agricultura de precisión, salud, sistemas GIS, vehículos aéreos no tripulados (UAV), seguridad y educación en diferentes tipos de proyectos con financiamiento interno y externo, becas de formación, dirección y codirección de tesis de grado, entre otras actividades.

1. INTRODUCCIÓN

Partiendo de este nuevo paradigma de problemas a los que nos enfrentamos como investigadores y docentes universitarios, consideramos que es necesario ajustar el alcance y potencialidad de las soluciones existentes desde el punto de vista tecnológico. La automatización en general y la de procesos en particular es una innovación tecnológica diseñada para que una máquina o algoritmo se encargue de una tarea en específico. Este tipo de innovación tecnológica, muy utilizada en el sector manufacturero en principio (tareas repetitivas por ejemplo en líneas de producción o tareas peligrosas), han ido aumentando su popularidad en los últimos años para otro tipo de actividades como procesos de servicio al cliente, inventariado, control de calidad y análisis de datos, entre otros. Según un reporte de la consultora Deloitte, en 2020 el 78% de las compañías ya había empezado a implementar la automatización de procesos, el 16% estaba preparándose para aplicarlo y tan solo el 6% no tenía aún planes de hacerlo [1]. De acuerdo con previsiones hechas por las grandes consultoras como Gartner [2] o Forrester [3] sobre las tecnologías más relevantes para 2021, la idea es apuntar hacia una estrategia tecnológica apoyada en la combinación de distintas tecnologías capaces de facilitar la digitalización de procesos de negocio de extremo a extremo buscando su eficiencia y el incremento de las habilidades de las personas en cada área en una empresa o institución.

En consecuencia, automatizar es necesario en la actualidad como parte de un proceso innovador desde el punto de vista tecnológico, pero claramente ya no será suficiente. Por lo tanto, el camino parece indicar que será necesaria la combinación de diferentes tecnologías de optimización, predicción y automatización cuyo objetivo es hacer que las operaciones sean más eficientes, sostenibles, y

dirigidas a una mejor solución capitalizando el conocimiento y la información para una toma de decisiones más eficiente. Se hacen entonces presentes términos como Hiper Automatización (Gartner), o Digital Process Automation (Forrester), o Intelligent Process Automation (IDC), para englobar este nuevo tipo de estrategias.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se presentan las principales líneas de investigación del proyecto:

Línea 1: Agropecuario, forestal y agroindustrial

Dentro de los sectores productivos en los que se puede aplicar la hiper automatización es el sector agropecuario, más específicamente en el área de Agricultura de Precisión (AP). De este modo, se ha convertido en una herramienta fundamental para lograr un manejo adecuado del suelo y sus cultivos, teniendo en cuenta su variabilidad dentro de un lote [4], permitiendo adaptarse a las exigencias de la agricultura moderna [5]. Es a partir de esto que se considera que dicha disciplina tiene un gran potencial a la hora de aplicar este tipo de tecnologías ya que allí el proceso de automatización permitiría agilizar y facilitar la tarea del productor agropecuario, brindándole además una herramienta soporte a la toma de decisiones.

Dentro de ésta área además se destaca el Sensado Remoto (SR) que incluye tanto a los Satélites como los Vehículos Aéreos No Tripulados (VANT - Drones) [6][7]. En este caso, estas tecnologías permiten sensor objetos o coberturas terrestres de manera remota y a partir del procesamiento de estas señales brindar información relevante. En el caso de la agricultura se pueden evaluar diferentes aspectos con el fin de dar soporte a los productores para la toma de decisiones, que

inevitablemente impactan en una mejor utilización del suelo y un mejor tratamiento del medio ambiente.

Línea 2: Salud

Otros de los sectores en la cual se puede realizar un aporte es en el ámbito de la salud. El avance de la tecnología en éste ámbito ha hecho que el sector comience a utilizarla cada vez más. Desde un sistema de software, con la utilización de aparatología para realizar diferentes mediciones médicas, el diagnóstico por imágenes o la utilización de modelos de reconstrucción de órganos en 3D hasta robots que permiten realizar cirugías de manera remota. Es por este motivo que se considera que la hiper automatización puede permitir, por un lado automatizar algunas tareas que los profesionales realizan actualmente de manera manual agilizando la tarea del profesional y por otro puede brindarle un soporte a la toma de decisiones o definiciones de diagnóstico.

En el ámbito de la cardiología, una rama particular de la medicina, la resonancia magnética presenta ventajas sobre las demás técnicas de diagnóstico por imagen, ya que permite visualizar las estructuras cardíacas en cualquier plano del espacio y obtiene imágenes con gran discriminación entre distintos tejidos [8]. Se estudian distintos aspectos de la patología cardiovascular, y su mayor desafío consiste en lograr imágenes de alta calidad de una estructura en constante movimiento. En la práctica habitual, el post procesamiento de las imágenes de resonancia cardíaca se realiza de manera manual [8][10]. La segmentación manual es una tarea que consume mucho tiempo y requiere personal con experiencia [10][11]. Se vienen desarrollando herramientas de segmentación semiautomática con distintos grados de intervención del operador y precisión [9][11], muchas basadas en inteligencia artificial.

Otro caso es el de la kinesiología, dónde un aspecto a tener en cuenta es la estimación de las posturas de los pacientes para el desarrollo de una terapia adecuada a la problemática en cuestión. Se analiza la medición de la biomecánica de las personas. Existen trabajos que apuntan a realizar estas mediciones mediante la utilización de visión artificial y técnicas de procesamiento de imágenes [12][13]. Es de interés poder reconstruir el movimiento y analizarlo en el laboratorio con un enfoque interdisciplinario, como también realizar evaluaciones específicas como por ejemplo evaluar la fuerza, la actividad muscular, la movilidad articular, entre otras cualidades del movimiento. Y en el caso de la investigación aportando en la simplificación de tareas o automatización de tests.

En el caso de la salud aplicada a diagnósticos tempranos, es posible trabajar con sistemas de toma de información en tiempo real para individuos que realizan una actividad deportiva. Información tal como esfuerzo realizado, desplazamiento, ritmo cardíaco, entre otros, pueden ser relevados y capturados para su análisis. La mejor aplicación a nuestros fines de diagnóstico temprano se da en su uso en colegios de nivel medio, donde una cardiopatía detectada en edad temprana puede llegar a salvar una vida. Se plantea además analizar dicha información no solo a la vista humana, sino que además, aplicar análisis de IA para dar mayor sensibilidad al sistema.

Línea 3: Seguridad

En el área de seguridad la utilización de diversos sensores es cada vez más común, y cada vez es más habitual encontrarse con la utilización de cámaras tanto por entidades gubernamentales, privadas como por parte de civiles. En este aspecto, se considera que la hiper automatización podría aportar un valor agregado, donde por ejemplo, mediante la utilización de cámaras conjuntamente con

actuadores [14]. Así se pueden desarrollar automatismos que permitan desde la prevención hasta el desarrollo de herramientas de hardware y/o software que sirvan de apoyo al personal de seguridad o a la justicia.

Línea 4: Educación

En los últimos tiempos las tecnologías de la información y la comunicación (TIC) han tenido un papel destacado y sostenido en la innovación educativa, especialmente en el último año. 2020, el año de la pandemia global generada por el Covid-19, ha traído consigo una nueva realidad educativa sostenida por la tecnología y basada en el uso de nuevas herramientas y plataformas educativas.

Para ello ha sido y será fundamental en el futuro la utilización de los conceptos de la hiper automatización de procesos que, en el terreno educativo, permite que los ambientes de aprendizaje sean adaptativos, basándose en las tecnologías emergentes. Así, analizan las actividades de los usuarios y su comportamiento, aprendiendo de ellos y adaptándolo. No se trata de simples tareas repetitivas, sino de procesos más complejos. En ellos se involucra el análisis de datos, monitoreo y toma de decisiones en tiempo real.

3. RESULTADOS OBTENIDOS/ESPERADOS

El equipo ya se encuentra trabajando sobre las cuatro líneas expuestas con resultados comprobados y otros en desarrollo. Respecto de la línea 1, se espera desarrollar herramientas de análisis que integren el uso de Vants, imágenes satelitales, y plataformas robóticas de recorrido a campo. En relación a la línea 2, se trabaja con la Sociedad Argentina de Cardiología para el desarrollo tanto de un dataset de imágenes de RM de corazón a nivel nacional para análisis de variables cualitativas y cuantitativas de forma automática. En cuanto al dispositivo de captura de datos de actividad

física, se espera obtener un sistema multicanal de hasta 32 canales simultáneos. Sobre la línea 3, se espera generar herramientas de control automatizado en ambientes externos y avanzar sobre la integración de actuadores con sistema de seguridad basados en video. Sobre la línea 4 se espera estudiar y desarrollar técnicas de educación inmersiva basadas en el uso de realidad virtual y realidad aumentada.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por investigadores formados y en proceso de formación, becarios del Sistema Nacional de Ciencia y Tecnología, becarios alumnos, graduados e investigadores externos.

Se han presentado planes de trabajo para becas, las cuales se encuentran actualmente en proceso de evaluación.

Se esperan finalizar, además, dos tesis doctorales, una tesis de maestría y presentarse a una convocatoria a beca postdoctoral del CONICET.

5. BIBLIOGRAFÍA

1. “Deloitte named a Leader by Gartner in Public Cloud Infrastructure Professional and Managed Services Worldwide | Deloitte | Press release.”
<https://www2.deloitte.com/us/en/pages/about-deloitte/articles/press-releases/deloitte-named-leader-gartner-public-cloud-infrastructure-professional-and-managed-services-worldwide.html>. (accessed Apr. 16, 2022).
2. D. Groombridge, “Las principales tendencias tecnológicas estratégicas para 2021,” Gartner, p.2021, 2021, [Online]. Available:
<https://www.gartner.es/es/tecnologia-de-la-informacion/tendencias/tendencias-tecnologicas-estrategicas-2021> .

3. "Forrester Helps Organizations Grow Through Customer Obsession." <https://www.forrester.com/bold> (accessed Apr. 16, 2022).
4. Y. Wang, K. Lee, S. Cui, E. Risch, and J. Lian, "Agriculture robot and applications," in *Future information engineering and manufacturing science : proceedings of the 2014 International Conference on Future Information Engineering and Manufacturing Science (FIEMS 2014)*, 2015, pp.43–46, doi: <https://doi.org/10.1201/b18167>.
5. S. Blackmore, "The role of yield maps in Precision Farming," Cranfield University, 2003.
6. S. Mandujano, "Drones: Una nueva tecnología para el estudio y monitoreo de fauna y hábitats", *AP*, vol. 10, n.o 10, mar. 2018.
7. A. Rutten, J. Casaer, M. F. A. Vogels, E. A. Addink, J. Vanden Borre, and H. Leirs, "Assessing agricultural damage by wild boar using drones," *Wildl. Soc. Bull.*, vol. 42, no. 4, pp. 568–576, Dec. 2018, doi: 10.1002/WSB.916.
8. O. Bernard et al., "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?," *IEEE Trans. Med. Imaging*, vol. 37, no. 11, pp. 2514–2525, 2018.
9. P. Peng, K. Lekadir, A. Gooya, L. Shao, S. E. Petersen, and A. F. Frangi, "A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging," *Magn. Reson. Mater. Physics, Biol. Med.*, vol. 29, no. 2, pp. 155–195, 2016.
10. J. Schulz-Menger et al., "Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update : Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing," *J. Cardiovasc. Magn. Reson.*, vol. 22, no. 1, p. 19, 2020.
11. T. Leiner et al., "Machine learning in cardiovascular magnetic resonance: Basic concepts and applications," *J. Cardiovasc. Magn. Reson.*, vol. 21, no. 1, pp. 1–14, 2019.
12. A. : Lenin, G. Aguilar, S. Dirigido, V. Espartaco, and R. Bykbaev, "Diseño y desarrollo de un módulo para determinar la postura humana empleando técnicas de visión artificial y reconocimiento de patrones como herramienta de," 2020, Accessed: Apr. 16, 2022. [Online]. Available: <https://dspace.ups.edu.ec/handle/123456789/18689>.
13. J. Reig Doménech, "Estudio del estado del arte de los métodos de estimación de la pose humana en 3D," Jun. 2018, Accessed: Apr. 16, 2022. [Online]. Available: <http://rua.ua.es/dspace/handle/10045/76995>.
14. Giacchetta, R. C., Brizuela, J., Cruza, J. F., & Romero, F. A. "HYFLIERS: drones inteligentes para automatizar la inspección por ultrasonido de infraestructuras de petróleo y gas". 2019, 14 Congreso AEND Vitoria-Gasteiz. H2020.

Tecnologías de Datos Espaciales, Visualización y Realidad Virtual

Luis Reynoso, Silvia Amaro, Lidia Lopez, Viviana Sanchez, M. José Rotter, Sergio Cotal

Departamento de Programación - Facultad de Informática

Universidad Nacional del Comahue

{luis.reynoso, silvia.amaro, lidia.lopez, viviana.sanchez, mariajoserotter, sergio.cotal}
@fi.uncoma.edu.ar

RESUMEN

Los avances tecnológicos y la situación de que buena parte de las actividades humanas tengan un componente locacional, han provocado que en la actualidad se disponga de un importante volumen de datos georreferenciados. En los últimos años, y alrededor del mundo se han multiplicado las iniciativas con las Infraestructuras de Datos Espaciales (IDE) para lograr la interoperabilidad de datos. Asimismo, la realidad virtual es cada vez más accesible. Junto con la realidad aumentada y la aplicación de los procesos de bigdata en este contexto, se busca dar un valor agregado fundamental a todo el proceso. Otra importante aplicación en la que se enfocará el proyecto, es en la creación y evaluación de Modelos de Valoración Automatizados (MVA) definido como un modelo matemático y económico para hallar de forma transparente, rápida y segura el valor de un conjunto de inmuebles. Los resultados pretendidos corresponden a utilizar estas tecnologías en el desarrollo de aplicaciones y tableros de control sobre datos abiertos y brindar una forma de interactuar con los datos y resultados.

Palabras clave: *Datos Espaciales, Visualización de Datos, Realidad Virtual, Bigdata*

CONTEXTO

Esta investigación se desarrolla en el contexto del Proyecto de investigación Tecnologías de Datos Espaciales,

Visualización y Realidad Virtual de la Facultad de Informática de la Universidad Nacional del Comahue. El desarrollo de estas tecnologías requiere del diseño de iniciativas, la construcción de aplicaciones y tableros de información entre la academia, la sociedad civil y el sector público y privado, para su explotación de información empleando tecnologías de infraestructura de datos espaciales, visualización, big data, realidad virtual, y datos abiertos.

El proyecto es financiado por la propia Universidad Nacional del Comahue

1. INTRODUCCIÓN

Visualización de datos [13], Infraestructura de Datos Espaciales [4], Big Data [11], Open Data (Datos Abiertos) [1], Open Gov (Gobierno Abierto) y Realidad Aumentada, son modelos tecnológicos disruptivos que de manera independiente o combinada comienzan a ser utilizados con mayor frecuencia en nuestra vida cotidiana. Estas tecnologías están despertando actualmente un gran interés en muchos organismos públicos y empresas por su enorme potencial estratégico.

Una de las principales variables para medir el impacto y alcance de estas tecnologías, al igual que en el caso de datos abiertos y geoservicios [6], es su agregación y reutilización. En la cadena de valor de datos abiertos y geoservicios, contribuyen tanto los creadores de datos abiertos como desarrolladores y enriquecedores que los explotan. La agregación de diferentes

conjuntos de datos, permite crear nuevos datos que pueden conducir a nuevos servicios o productos de datos.

Varios organismos estatales ponen a disposición cantidades cada vez mayores de datos abiertos en Internet [3]. Esto se debe fundamentalmente a garantizar el efectivo ejercicio del derecho de acceso a la información pública, promover la participación ciudadana y la transparencia de la gestión pública, en el marco de La Ley 27275 de Acceso a la Información Pública. Estos conjuntos de datos abiertos (datasets) están creciendo día a día. A pesar de que para acceder a Datos Abiertos se utilizan cada vez con mayor frecuencia los portales nacionales de Datos Abiertos, seguidos de los que acceden directamente a las propias administraciones públicas, los conjuntos de datos abiertos, pueden ser tan grandes, que mirar la información en modo tabular no nos dice mucho. En cambio geolocalizar la información en mapas ó explotar los datos a partir de técnicas de big data y visualización [7, 12, 14], nos ayuda también a comprender los datos, sacar conclusiones, y es muy ilustrativa una vez reflejada en un mapa. Del mismo modo los tableros de control, son una herramienta de visualización de los principales indicadores y datos que nos ayudan a facilitar y simplificar la comprensión del estado general de una temática. Por otro lado, no podemos dejar de tener en cuenta, que con el cambio de época a partir del año 2020, y el comienzo de la pandemia [5], todos los organismos públicos y privados, han potenciado el uso de la web, el trabajo a distancia, y la posibilidad de acceder a recursos y productos a partir del uso de tecnologías web. Todo esto configura nuevas formas de producción y trabajo colaborativo en formato digital que requiere inexorablemente el empleo de las mejores

técnicas y estrategias para visibilizar y difundir la información sustantiva primaria de organismos y empresas (las cuales son fuente auténtica de información).

En este contexto los objetivos de esta investigación y desarrollo abarcan:

- Análisis de modelos, técnicas, estándares y software de visualización existentes, Infraestructura de Datos Espaciales (IDE), realidad virtual, big data y data science aplicadas a conjuntos diversos (incluidos datos abiertos) .
- Difundir iniciativas en el uso, explotación e integración de datos abiertos, geoservicios, e interoperabilidad de sistemas (iniciativas propias del proyecto ó de otros proyectos en el marco de IDE).
- Desarrollar aplicaciones web y de realidad virtual, geoservicios, y tableros de control que utilicen, difundan, integren y exploten información geolocalizada a partir de las tecnologías mencionadas, empleando técnicas y metodologías científicas.
- Difundir prácticas de datos abiertos en la Universidad Nacional del Comahue e integrar información territorial generada por distintas Unidades Académicas en el marco de tecnologías de Infraestructura de Datos Espaciales de la institución.
- Fortalecer y consolidar el nodo IDE de la Universidad Nacional del Comahue (IDEUNCo).
- Colaborar a través del Nodo IDEUNCo con la localización de Objetivos de Desarrollo Sostenible (ODS).
- Especificar un modelo de evaluación de IDEs.
- Articular trabajo conjunto con las unidades académicas, organismos públicos y privados para aplicar explotación de información con el uso de las tecnologías investigadas en el proyecto, como así también difusión de

técnicas y tecnologías a partir de capacitaciones.

- Documentar las propuestas realizadas (considerado éste un proceso continuo).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En este proyecto podemos identificar las siguientes líneas de investigación y desarrollo:

a. Infraestructuras de Datos Espaciales (IDE)

Una IDE de esta forma mejora el acceso y la publicación de información geográfica en diversos formatos. Una IDE implica también un conjunto de políticas, estándares, organizaciones y recursos tecnológicos que faciliten la producción, obtención, uso y acceso a la información georreferenciada (Linares, 2015). Una IDE almacena y administra datos y atributos geográficos lo suficientemente bien documentados para lograr su aplicabilidad y confiabilidad, posibilita un medio sencillo de búsqueda, visualización y evaluación a través de catálogos y servidores de mapas, entre otros servicios potenciales. Por otro lado, la literatura sobre IDE [2, 4] menciona que un elemento saliente de las IDE es su organización, su marco institucional, sus políticas, que dependerán del objetivo con el que sea creada y las lógicas de funcionamiento establecidas.

b. Realidad Virtual

La realidad virtual (RV) se basa en computación gráfica, tecnología de simulación y tecnología de multimedia; para simular funciones visuales, auditivas, táctiles y otras funciones sensoriales humanas [9]. A través de esta tecnología las personas pueden

sumergirse en un entorno virtual con la información generada por una computadora y experimentar algún tipo de interacción. El recorrido virtual es una rama de aplicaciones de RV y consiste en la creación de información del entorno virtual en un nuevo espacio multidimensional, basándose en datos reales mediante el uso de la tecnología de RV [10]. Con el entorno creado se puede realizar el recorrido a larga distancia del mundo real, con las características de las 3I: interacción, inmersión e imaginación.

c. BigData

Se refiere a grandes volúmenes de datos que se originan desde diferentes fuentes; en donde la variedad y la velocidad juegan un rol preponderante, y es muy importante utilizar herramientas y procesos específicos para la recolección, almacenamiento, procesamiento y análisis de los datos. El ciclo de vida de BigData es diferente al ciclo de vida de un desarrollo de software tradicional, en él se pueden observar las siguientes 6 actividades: evaluación del caso de negocio, identificación y recolección de datos, preparación de datos, análisis de datos, modos de análisis, visualización y utilización de los resultados. En este escenario es fundamental utilizar herramientas de visualización de datos [14], que permitan explotar los resultados obtenidos. La posibilidad de utilizar RV [9] en este contexto puede dar un valor agregado fundamental a todo el proceso.

d. Modelos de Valoración Automatizados

La aplicación de las técnicas mencionadas en la creación y evaluación de Modelos de Valoración Automatizados (MVA). Un MVA es un modelo matemático y económico para hallar de forma transparente, rápida y segura el valor de un conjunto de inmuebles.

Este modelo fundamentalmente cuenta con un conjunto de variables independientes o explicativas, y una variable dependiente que es el valor de mercado. Asociado al modelo hay un conjunto muy rico y amplio de técnicas que se utilizan para optimizarlo. El resultado que se obtiene al utilizar el modelo es obtener un valor estimativo en una fecha concreta para un conjunto de parcelas en las etapas de especificación y calibración del modelo.

Alternativamente estas técnicas utilizadas para MVA también pueden ser utilizadas para la especificación y evaluación de Modelos Tributarios en la visualización de datos.

3. RESULTADOS

OBTENIDOS/ESPERADOS

A continuación se enumeran y describen algunos de los resultados esperados a partir de la investigación y aplicación de las tecnologías de interés. Asimismo se indican algunos resultados preliminares obtenidos.

a. Observatorio de Hidrocarburos;

Se desarrolló un tablero sobre Hidrocarburos haciendo uso de datasets del Ministerio de Energía, de gran utilidad para el área del Comahue donde se encuentra localizada la UNCo. El propósito final es construir un Observatorio de Hidrocarburos con el propósito de lograr vinculación tecnológica, y visualización del proceso de explotación. Una versión preliminar del tablero se puede encontrar *en:* [\[http://opendata.fi.uncoma.edu.ar/TableroHidrocarburos.html\]](http://opendata.fi.uncoma.edu.ar/TableroHidrocarburos.html)

b. UNCovid

Se encuentra en desarrollo un tablero Covid de la Universidad Nacional del Comahue (UNCo) y en una etapa de validación que requiere mayor desarrollo y explotación de la

información. Este tablero utiliza como insumos datos abiertos del ministerio de salud, y combina a su vez el geoservicio WMTS del IGN ArgenMap: <https://opendata.fi.uncoma.edu.ar/TableroUNCovid.html>

c. En conjunto con la Secretaría de Planificación y Desarrollo Institucional,

Se desarrollará un tablero que visualice la oferta académica de la Universidad Nacional del Comahue, con datos de contacto de las distintas Unidades Académicas (UUAA), e información de recursos humanos de las UUAA. Dicho tablero requiere ser ampliado con información de grupos de extensión e investigación en las distintas UUAA. Por otro lado, es posible construir un mapa de graduados de la UNCo, para conocer donde están trabajando los egresados de nuestra casa de estudios. El tablero de la UNCo, se encuentra en desarrollo preliminar, dos de sus últimas versiones son:

[\[https://opendata.fi.uncoma.edu.ar/TableroUNComa.html\]](https://opendata.fi.uncoma.edu.ar/TableroUNComa.html)

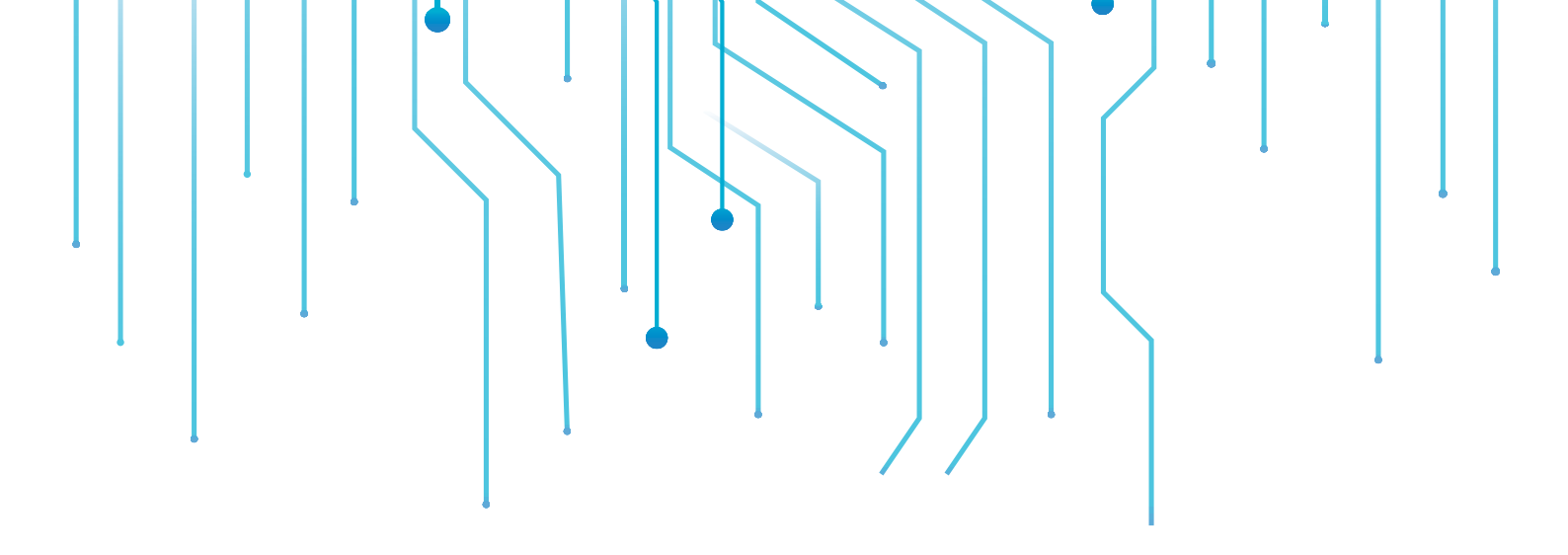
[\[https://opendata.fi.uncoma.edu.ar/TableroUNComaCuarto.html\]](https://opendata.fi.uncoma.edu.ar/TableroUNComaCuarto.html)

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta investigación está conformado por un Doctor en Computación, una Magíster en Ciencias de la Computación, y cuatro Licenciadas y Licenciados en Ciencias de la Computación, que están realizando la Maestría en Ciencias de la Computación de la Facultad de Informática de la Universidad Nacional del Comahue y actualmente están en proceso de armado de la propuesta de tesis. En temas relacionados con las líneas de investigación del proyecto.

5. BIBLIOGRAFÍA

- [1] Berends, J., Carrara, W., Engbers, W., Vollers, H. (2020) Re-using Open Data. A Study of Companies transforming Open Data into Economic & Social Value. European Data Portal, European Commission. ISBN: 978-92-78-41872-4
https://data.europa.eu/sites/default/files/re-using_open_data.pdf
- [2] Guisande González, C. , Vaamonde Liste, A. Barreiro Felpeto, A.. Tratamiento de datos con R, Statistica y SPSS. Madrid: Ediciones Díaz de Santos, 2013
- [3] Instituto Geográfico Nacional (2021a). Argenmap mapa base.
<https://www.ign.gob.ar/AreaServicios/Argenmap/Introduccion> [accedido 1 de Junio de 2021]. Instituto Geográfico Nacional (2021b). Datos abiertos del IGN
- [4] Linares, S. (2015) Una IDE para consultar, usar y compartir geodatos en docencia e investigación universitaria. Párrafos geográficos. IGEPAT. Año 2016 Volumen 15 N° 1 ISSN 1853-9424 en línea: http://igeopat.org/parrafosgeograficos/index.php?option=com_content&view=article&id=51&catid=25
- [5] Ministerio de Salud de la Nación, Argentina (2021b). COVID-19. Casos registrados en la República Argentina. Datos Abiertos del Ministerio de Salud. Dataset: <http://datos.salud.gob.ar/dataset/covid-19-casos-registrados-en-la-republica-argentina> [accedido 20 de Junio de 2021]
- [6] OGC Open Geospatial Consortium (2021), Organización de Estándares internacionales Disponible en: www.opengeospatial.org
- [7] Philipp K. Janert. Data Analysis with Open Source Tools. O'Reilly 2010.
- [8] Yin, R. K. Case Study Research. Design and Methods Fifth Edition. SAGE Publications, Inc., 2014
- [9] Burdea, G.C. and COIFFET, P. Virtual Reality Technology, New Yersey: John Wiley&Sons, 2003
- [10] Meraz, José Manuel Falcón, and Carlos Domenzain Domínguez. "Del museo sin muros, al museo como simulación fotográfica: experiencias contemporáneas en los museos en línea." Kepes 14, no. 16 (2017): 185-217
- [11] T. Erl, W. Khatkhat, and P. Buhler, Big data fundamentals: concepts, drivers & techniques. PrenticeHall Boston, 2016, vol. 1
- [12] A. Bahga and V. Madiseti, Big data science & analytics: A hands-on approach. Arshdeep Bahga & Vijay Madiseti, 2016
- [13] Card, S., Mackinlay, J., Shneiderman, B., Readings in Information Visualization – Using Vision to Think, Morgan Kaufmann, 1999
- [14] Alexandru C. Telea, Data Visualization: Principles and Practice, Second Edition, AK Peters / CRC Press, 2014



PDP

Procesamiento Distribuido y en Paralelo

PARALELIZACION DE ALGORITMOS Y EVALUACION DE RENDIMIENTO EN PLATAFORMAS DE CÓMPUTO DE ALTAS PRESTACIONES

Marcelo Naiouf⁽¹⁾ , Armando De Giusti⁽¹⁾⁽²⁾ , Laura De Giusti⁽¹⁾⁽³⁾ , Franco Chichizola⁽¹⁾ , Victoria Sanz⁽¹⁾⁽³⁾ ,
Adrián Pousa⁽¹⁾ , Enzo Rucci⁽¹⁾⁽³⁾ , María José Basgall⁽¹⁾⁽²⁾ , Mariano Sánchez⁽¹⁾ , Manuel Costanzo⁽¹⁾ ,
Silvana Gallo⁽¹⁾⁽²⁾ , Emmanuel Frati⁽¹⁾⁽⁴⁾ , Adriana Gaudiani⁽⁵⁾ , Sergio Leandro Calderón⁽¹⁾

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires
Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CIC – Comisión de Investigación Científica de la Provincia de Buenos Aires

⁴Universidad Nacional de Chilecito

⁵Universidad Nacional de General Sarmiento

{mnaiouf, degiusti, ldgiusti, francoch, vsanz, apousa, erucci, sgallo, fefrati,
msanchez, mjbassgall, scalderon}@lidi.info.unlp.edu.ar, agaudi@ungs.edu.ar

RESUMEN

El eje central de la línea de I/D es investigar en temas de cómputo paralelo y distribuido de alto desempeño, tanto en lo referido a los fundamentos como a la construcción, evaluación y optimización de las aplicaciones en arquitecturas multiprocesador. Se aplican los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos con el fin de obtener soluciones de alto rendimiento.

También incluye la construcción de ambientes para la enseñanza de la programación concurrente y paralela.

En la dirección de tesis de postgrado existe colaboración con el grupo HPC4EAS (High Performance Computing for Efficient Applications and Simulation) del Dpto. de Arquitectura de Computadores y Sistemas Operativos de la Universidad Autónoma de Barcelona; con el Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid; y con el grupo Soft Computing and Intelligent Information Systems (SCI2S) de la Universidad de Granada, entre otros.

Palabras clave: Cómputo paralelo y distribuido de altas prestaciones. Algoritmos paralelos y distribuidos. Arquitecturas multiprocesador. Ambientes de enseñanza.

CONTEXTO

La línea de I/D que se presenta es parte del Proyecto “Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, Fog-Edge-Cloud, Big Data, Robótica, y Tiempo Real” del III-LIDI acreditado por el Ministerio de Educación, y de proyectos acreditados y subsidiados por la Facultad de Informática de la UNLP. Además, existe cooperación con Universidades de Argentina, Latinoamérica y Europa a través de proyectos acreditados. Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD).

1. INTRODUCCIÓN

El área de cómputo de altas prestaciones (HPC, High-Performance Computing) es clave dentro de las Ciencias de la Computación, debido al creciente interés por el desarrollo de soluciones a problemas con alta demanda computacional y de almacenamiento, produciendo transformaciones profundas en las líneas de I/D [1].

El rendimiento en este caso está relacionado con dos aspectos: las arquitecturas de soporte y los algoritmos que hacen uso de las mismas, y el desafío se centra en cómo aprovechar las prestaciones obtenidas a partir de la evolución de las arquitecturas físicas. En esta línea la mayor importancia está en los algoritmos paralelos y en los métodos utilizados para su construcción y análisis a fin de optimizarlos.

Uno de los cambios de mayor impacto ha sido el uso de manera masiva de procesadores con más de un núcleo (*multicore*), produciendo plataformas distribuidas híbridas (memoria compartida y distribuida) y generando la necesidad de desarrollar sistemas operativos, lenguajes y algoritmos que las usen adecuadamente. También creció la incorporación de placas aceleradoras a los sistemas multicore constituyendo plataformas paralelas de memoria compartida con paradigma de programación propio asociado como pueden ser las unidades de procesamiento gráfico (GPU, Graphic Processing Unit) de NVIDIA y AMD, los coprocesadores Xeon Phi de Intel [2] o los aceleradores basados en circuitos integrados reconfigurables (FPGAs, Field Programmable Gate Array) [3]. En la actualidad, se comercializan placas de bajo costo como Raspberry PI [4], Odroid [5], NVIDIA Jetson [6], entre otras. Estas placas, con arquitectura similar a la que podemos encontrar en dispositivos móviles, poseen múltiples núcleos de baja complejidad y en algunos casos son procesadores multicore asimétricos (AMPs) con el mismo repertorio de instrucciones. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones [7]. Asimismo, los entornos

de computación cloud introducen un nuevo foco desde el punto de vista del HPC, brindando un soporte “a medida” sin la necesidad de adquirir el hardware.

La creación de algoritmos paralelos en arquitecturas multiprocesador no es un proceso directo [8]. El costo puede ser alto en términos del esfuerzo de programación y el manejo de la concurrencia adquiere un rol central en el desarrollo. Si bien en las primeras etapas el diseñador de una aplicación paralela puede abstraerse de la máquina sobre la que ejecutará el algoritmo, para obtener buen rendimiento debe tenerse en cuenta la plataforma de destino. En las máquinas multiprocesador, se deben identificar las capacidades de procesamiento, interconexión, sincronización y escalabilidad. La caracterización y estudio de rendimiento del sistema de comunicaciones es de interés para la predicción y optimización de performance, así como la homogeneidad o heterogeneidad de los procesadores [9].

Muchos problemas algorítmicos se vieron impactados por los multicore y clusters de multicore. El desarrollo de algoritmos que aprovechen adecuadamente las arquitecturas, motiva el estudio de performance en sistemas híbridos. Es necesario estudiar la utilización de lenguajes y bibliotecas ya que aún no se cuenta con un standard, aunque puede mencionarse el uso de los tradicionales MPI, OpenMP y Pthreads [10][11] o los más recientemente explorados UPC, Chapel y Titanium del modelo PGAS [12].

La combinación de arquitecturas de múltiples núcleos con aceleradores dio lugar a plataformas híbridas con diferentes características. Más allá del acelerador utilizado, la programación de estas plataformas representa un desafío. Para lograr aplicaciones de alto rendimiento, los programadores enfrentan dificultades como: estudiar características específicas de cada arquitectura y aplicar técnicas de programación y optimización particulares de cada una, lograr un balance de carga adecuado entre los dispositivos de procesamiento y afrontar la ausencia de estándares para este tipo de sistemas.

Por otra parte, los avances en las tecnologías de virtualización han llevado a que Cloud Computing sea una alternativa a los tradicionales sistemas de cluster [13]. El uso de cloud para HPC presenta desafíos atractivos, brindando un entorno reconfigurable dinámicamente sin la necesidad de adquirir hardware, y es una excelente plataforma para testear escalabilidad de algoritmos.

Métricas de evaluación del rendimiento y balance de carga

La diversidad de opciones vuelve complejo el análisis de performance de los Sistemas Paralelos, ya que los ejes sobre los cuales pueden compararse dos sistemas son varios. Existe un gran número de métricas para evaluar el rendimiento, siendo las tradicionales: tiempo de ejecución, speedup, eficiencia. Por su parte, la *escalabilidad* permite capturar características de un algoritmo paralelo y la arquitectura en que se lo implementa. Posibilita testear la performance de un

programa sobre pocos procesadores y predecirla en un número mayor, así como caracterizar la cantidad de paralelismo inherente en un algoritmo.

Un aspecto de interés que se ha sumado como métrica, a partir de las plataformas con gran cantidad de procesadores, es el del *consumo* y la *eficiencia energética* [14]. Muchos esfuerzos están orientados a tratar el consumo como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirlo.

El objetivo principal del cómputo paralelo es reducir el tiempo de ejecución haciendo uso eficiente de los recursos. El *balance de carga* es un aspecto central y consiste en, dado un conjunto de tareas que comprenden un algoritmo y un conjunto de procesadores, encontrar el mapeo (asignación) de tareas a procesadores tal que cada una tenga una cantidad de trabajo que demande aproximadamente el mismo tiempo, y esto es más complejo si hay heterogeneidad. Dado que el problema general de mapping es *NP-completo*, pueden usarse enfoques que dan soluciones subóptimas aceptables. Las técnicas de planificación a nivel micro (dentro de cada procesador) y macro (en un cluster) deben ser capaces de obtener buen balance de carga. Existen técnicas estáticas y dinámicas cuyo uso depende del conocimiento que se tenga sobre las tareas de la aplicación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Investigar en temas de cómputo paralelo y distribuido de alto desempeño, en lo referido a los fundamentos y a la construcción y evaluación de las aplicaciones. Esto incluye los problemas de software asociados con el uso de arquitecturas multiprocesador:
 - Lenguajes, modelos y paradigmas de programación paralela (puros e híbridos a distintos niveles).
 - Asignación de procesos a procesadores optimizando el balance de la carga de procesamiento.
 - Métricas de evaluación de complejidad y rendimiento: speedup, eficiencia, escalabilidad, consumo energético, costo de programación.
- Construir, evaluar y optimizar soluciones utilizando algoritmos concurrentes, paralelos y distribuidos sobre diferentes plataformas de software y arquitecturas con múltiples procesadores:
 - Arquitecturas de trabajo homogéneas, heterogéneas e híbridas: multicores, clusters, GPU, Xeon Phi, FPGA, placas de bajo costo y entornos cloud.
 - Aplicar los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos (aplicaciones científicas, búsquedas, simulaciones, imágenes, realidad virtual y aumentada, bioinformática, big data, n-body).

- Analizar y desarrollar ambientes para la enseñanza de programación concurrente y paralela.
 - Caracterizar diferentes modelos de arquitecturas paralelas.
 - Representar distintos modelos de comunicación/sincronización.
 - Definir métricas de evaluación de rendimiento y eficiencia energética.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollar y optimizar algoritmos paralelos sobre diferentes modelos de arquitectura. En particular, en aplicaciones numéricas y no numéricas de cómputo intensivo y tratamiento de grandes volúmenes de datos.
- Estudiar y comparar los lenguajes sobre las plataformas multiprocesador para diferentes modelos de interacción entre procesos.
- Investigar la paralelización en plataformas que combinan clusters, multicore y aceleradores. Comparar estrategias de distribución de trabajo teniendo en cuenta las diferencias en potencias de cómputo y comunicación, dependencia de datos y memoria requerida.
- Evaluar la performance (speedup, eficiencia, escalabilidad, consumo energético) de las soluciones propuestas. Analizar el rendimiento de soluciones paralelas a problemas con diferentes características (dependencia de datos, relación cómputo / comunicación, memoria requerida).
- Mejorar y adecuar las técnicas disponibles para el balance de carga (estático y dinámico) entre procesos a las arquitecturas consideradas.

En este marco, pueden mencionarse los siguientes resultados:

- Para la experimentación se han utilizado y analizado diferentes arquitecturas homogéneas o heterogéneas, incluyendo multicores, cluster de multicores (con 128 núcleos), GPU y cluster de GPU, Xeon Phi y FPGA.
- Se experimentó la paralelización en arquitecturas híbridas, con el objetivo de estudiar el impacto del mapeo de datos y procesos, así como de los lenguajes y librerías.
- Respecto de las aplicaciones y temas estudiados, se trabajó fundamentalmente con los siguientes problemas:

➤ **Aceleración de aplicaciones con cómputo colaborativo CPU-GPU.** Las computadoras comerciales actuales incluyen decenas de cores y al menos una GPU. El uso de ambas unidades de procesamiento de forma colaborativa puede mejorar significativamente el rendimiento de una aplicación. Sin embargo, esto supone un desafío para los programadores ya que dichas unidades difieren en arquitectura, modelo de programación y rendimiento. En [15] se propuso un modelo híbrido para estructurar código a ser ejecutado sobre un sistema heterogéneo con múltiples cores y 1 GPU (utilizando todos los recursos disponibles). Utilizando este modelo se desarrolló un algoritmo paralelo de pattern matching para sistemas

heterogéneos CPU-GPU [16]. Los resultados revelaron que este algoritmo supera en rendimiento a trabajos previos, desarrollados para sistemas multicore y GPUs, para datos de tamaño considerable. En [17] se presentó una solución al problema de pattern matching que aprovecha toda la potencia computacional de los procesadores Intel Xeon Phi KNL 7230 mediante el uso de SIMD y paralelismo de hilos. Se mostró que el algoritmo propuesto alcanza aceleraciones significativas. En trabajos futuros interesa investigar sobre el cómputo colaborativo incluyendo este tipo de arquitecturas.

➤ **Bioinformática y biología computacional.** Estas áreas incluyen aplicaciones que se nutren de las capacidades de HPC para brindar soluciones en un tiempo de respuesta aceptable, siendo cada día más debido al crecimiento exponencial de datos genómicos en los últimos años [18]. Es posible encontrar aplicaciones en alineamiento de secuencias, acoplamiento y dinámica molecular, predicción y búsqueda de estructuras moleculares, entre otras.

➤ **Alineamiento de secuencias.** Si bien el algoritmo de Smith-Waterman es considerado el método de alineamiento más preciso, este algoritmo resulta costoso debido a su complejidad computacional cuadrática. Dado el surgimiento del estándar SYCL y de implementaciones como Intel oneAPI, se migró una versión CUDA de este algoritmo a su equivalente en SYCL, analizando el esfuerzo de programación requerido y la relación costo-beneficio entre overhead y portabilidad para diferentes arquitecturas CPU y GPUs [19]. A futuro, se propone extender el estudio realizado contemplando otras generaciones de GPUs y arquitecturas como FPGAs.

➤ **Identificación de biomarcadores.** entre todas las plataformas, servicios y paquetes de software disponibles para esta aplicación, se encuentra Multiomix [20], quien se destaca del resto por facilitar el proceso de integración y minería de datos oncogenómicos públicos y cargados por los usuarios, proporcionando una interfaz gráfica de usuario amigable para los usuarios no expertos. A pesar de ello, Multiomix aun funciona de forma secuencial por ser un desarrollo reciente (una ejecución habitual en Multiomix puede tomar varias horas para datasets de pequeño porte). Se propone acelerar el procesamiento de Multiomix de manera de disminuir los tiempos de respuesta y así mejorar la productividad de sus usuarios.

➤ **Cálculo de los caminos mínimos.** Es uno de los problemas básicos y de mayor antigüedad de la teoría de grafos teniendo aplicación en el dominio de las comunicaciones, del ruteo de tráfico, de la bioinformática, entre otros. El algoritmo de Floyd-Warshall (FW) permite computar la distancia mínima

entre todos los pares de un grafo. Además de poseer una alta demanda de ancho de banda, FW resulta costoso computacionalmente al ser $O(n^3)$. Habiendo desarrollado una versión optimizada para la arquitectura Intel Xeon Phi KNL [21], se propone tomarla como base para implementar una equivalente para procesadores multicore convencionales. Al mismo tiempo, se espera encontrar posibles oportunidades de optimización paralela.

➤ **Simulación de N cuerpos computacionales con atracción gravitacional.** Su propósito es aproximar en forma numérica la evolución de un sistema de cuerpos en el que cada uno interactúa con todos los restantes. El uso más conocido de esta simulación quizás sea en la astrofísica, donde cada cuerpo representa una galaxia o una estrella particular que se atraen entre sí debido a la fuerza gravitacional. Si bien existen diferentes métodos para procesar la simulación de los N cuerpos, en todos los casos se requiere alta demanda computacional. Se estudió la paralelización de la versión directa de esta simulación sobre diferentes lenguajes no convencionales en el ámbito de HPC, en particular Python y Rust. Se focalizó en cómo diferentes técnicas de optimización logran mejorar el rendimiento y en su comparación con un lenguaje tradicional como es C, considerando no sólo el rendimiento sino también el esfuerzo de programación. A futuro, interesa extender el estudio incorporando otras aplicaciones y arquitecturas multicore [22][23][24].

➤ **Problemas de optimización de simulación de sistemas dinámicos complejos mediante heurísticas.** La búsqueda de un conjunto de parámetros de entrada que optimicen el funcionamiento de un simulador de un sistema físico es un proceso de alto costo computacional que puede considerarse intratable y requiere de heurísticas que permitan disminuir el tiempo de ejecución. En los nuevos avances realizados en esta línea se aprovechó la continuidad en los valores de los parámetros físicos distribuidos sobre el dominio del sistema, de manera de realizar búsquedas de parámetros ajustados sobre espacios de búsquedas de tamaño mucho más reducidos que los utilizados en una primera etapa del trabajo. Esta metodología mucho más eficiente se puede extender a otros simuladores de fenómenos físicos y en particular con simuladores de inundaciones de ríos con las que se llevaron adelante las experiencias. Las experiencias se pueden correr de manera colaborativa beneficiándose ampliamente del uso de plataformas de clusters de procesadores [25] [26].

➤ **Ambientes para la enseñanza de concurrencia.** Se desarrolló el entorno CMRE para la enseñanza de programación concurrente y paralela a partir de cursos iniciales en carreras de Informática. Incluye un entorno visual que representa una ciudad en la que pueden definirse varios robots que interactúan. Combina

aspectos de memoria compartida y distribuida mediante instrucciones para bloquear y liberar esquinas de la ciudad y el concepto de pasaje de mensajes a través de primitivas de envío y recepción. Además, se incluyen los conceptos de heterogeneidad (diferentes velocidades de los robots) y consumo energético [27]. Se ha integrado con el uso de robots físicos (Lego Mindstorm 3.0) que ejecutan en tiempo real las mismas instrucciones que los robots virtuales y se comunican con el entorno mediante bluetooth [28]. Se ha ampliado para incorporar conceptos básicos de computación en la nube (Cloud Computing) [28]. Actualmente, se está desarrollando una nueva herramienta para la enseñanza de programación concurrente en cursos avanzados. Su objetivo principal es visualizar los conceptos de sincronización y comunicación entre procesos.

➤ **Aplicaciones en Big Data.** En los últimos años, los escenarios de grandes volúmenes de datos (Big Data) son cada vez más comunes debido a la constante generación de datos a partir de fuentes como sensores o Internet, por ejemplo. Para poder trabajar con conjuntos de datos Big Data, se requiere de cómputos de altas prestaciones y de herramientas software específicas capaces de procesar significativos tamaños de datos en tiempos razonables, y que puedan ser ejecutados de manera paralela y distribuida.

Una de las herramientas más populares para el procesamiento de Big Data es Apache Spark, la cual presenta varias características deseables que la hacen ser ampliamente elegida. El uso intensivo de memoria RAM, los mecanismos eficientes de tolerancia a fallos, las distintas estructuras de datos que provee altamente optimizadas, la amplia variedad de librerías que habilitan utilizar algoritmos de Machine Learning, procesamiento en Streaming, entre otros, son algunas de sus características más relevantes.

En esta línea de trabajo, se están desarrollando nuevas técnicas para el preprocesamiento de datos para problemas de clasificación en Big Data, utilizando el framework Apache Spark sobre Cómputo de Altas Prestaciones. Como parte del trabajo realizado se presenta una propuesta metodológica para la reducción dual (reducción de instancias y de características) de un conjunto de datos tabulares que representa un problema de clasificación [29].

Se trata de un diseño sencillo y escalable, capaz de analizar la reducción de datos de forma vertical y horizontal, recibiendo la visión del experto en datos mediante el establecimiento de dos simples condiciones: un umbral de calidad predictiva asociado al modelo obtenido a partir del conjunto reducido, y un rango de porcentajes de reducción a comprobar. Además, su implementación utiliza operaciones paralelas y utilidades totalmente optimizadas proporcionadas por Apache Spark.

Esta línea de trabajo se lleva a cabo en colaboración con el Dr. Alberto Fernández Hilario del grupo de

investigación "Soft Computing and Intelligent Information Systems" (SCI2S) de la Universidad de Granada, España.

➤ **Cifrado de grandes volúmenes de datos.** Hoy en día, la cantidad de datos sensibles que se generan para ser almacenados y/o transmitidos a través de la red aumenta constantemente. Para proteger los datos confidenciales de amenazas potenciales, se utilizan estrategias de encriptación. Además, el tiempo involucrado en el cifrado de datos está directamente relacionado con la cantidad de datos que se cifrarán y puede ser significativo. Para reducir el tiempo de cifrado es natural recurrir a soluciones de encriptación paralelas, que explotan todo el poder computacional proporcionado por las arquitecturas emergentes. AES (Advanced Encryption Standard) es uno de los algoritmos de cifrado más utilizados y el gobierno de los Estados Unidos lo considera lo suficientemente seguro como para proteger la información nacional. Hay varias implementaciones de AES, tanto en hardware como en software. En [30] se presenta una comparación de rendimiento de una solución AES basada en hardware para CPU multinúcleo con el de otras dos soluciones AES basadas en software para CPU multinúcleo y GPU, respectivamente. El primero se implementa con las nuevas instrucciones de Intel AES y el segundo con la biblioteca OpenSSL. Los resultados revelan que utilizar cómputo paralelo para el proceso de cifrado reduce significativamente el tiempo de ejecución respecto a una solución secuencial. Los resultados también muestran que el mayor rendimiento se alcanza utilizando una solución multicore con AES basado en hardware. Sin embargo, la solución por software utilizando 2 GPUs puede ser una alternativa competitiva a la solución multicore por hardware cuando se tienen pocos núcleos o una CPU que no soporta AES. Asimismo, en [31] se muestra que una solución colaborativa, que utiliza al mismo tiempo múltiples cores y 2 GPUs para el cifrado de grandes volúmenes de datos, permite alcanzar mejoras en el rendimiento.

Organización de Eventos

En 2022 se organizaron las X Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2022), con participación de especialistas académicos del país y el exterior, además de empresas con experiencia en Cloud Computing [32][33]. En junio de 2023 se tiene prevista la organización de las XI JCC-BD&ET [34].

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D se concluyó 1 tesis doctoral, 1 Trabajo Final de Especialización y 2 Tesinas de Grado de Licenciatura. Se encuentran en curso en el marco del proyecto 3 tesis doctorales, 2 de maestría, 2 trabajos de Especialización y 4 Tesinas de grado.

Se participa en el dictado de las carreras de Doctorado en Cs. Informáticas y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática UNLP, por lo que potencialmente pueden generarse más Tesis y Trabajos Finales.

Hay cooperación con grupos de otras Universidades del país y del exterior, y tesis de diferentes Universidades realizan su trabajo con el equipo del proyecto.

5. BIBLIOGRAFÍA

- [1]. Giles MB, Reguly I. "Trends in high-performance computing for engineering calculations". *Phil.Trans.R.Soc.A* 372: 20130319. 2014. <http://dx.doi.org/10.1098/rsta.2013.0319>
- [2]. Jeffers J., Reinders J. "Intel Xeon Phi Coprocessor High Performance Programming". Morgan Kaufmann.
- [3]. Sean Settle. "High-performance Dynamic Programming on FPGAs with OpenCL". *IEEE High Performance Extreme Computing Conf (HPEC '13)*. 2013. <https://doi.org/10.1016/j.physa.2017.04.159>.
- [4]. Raspberry PI. <https://www.raspberrypi.org/> Accedido 21 de Marzo de 2016.
- [5]. Odroid <http://www.hardkernel.com> Accedido 2016.
- [6]. Nvidia Jetson <https://www.nvidia.com/es-la/autonomous-machines/embedded-systems/> Accedido en Febrero de 2023
- [7]. Annamalai A., Rodrigues R., Koren I., Kundu S. "Dynamic Thread Scheduling in Asymmetric Multicores to Maximize Performance-per-Watt". *2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*, pp. 964-971, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012.
- [8]. McCool M. "Structured Parallel Programming: Patterns for Efficient Computation". Morgan Kaufmann.
- [9]. De Giusti L, Naiouf M., Chichizola F., Luque E., De Giusti A. "Dynamic Scheduling in Heterogeneous Multiprocessor Architectures. Efficiency Analysis". *Computer Science and Technology Series – XV Argentine Congress of Computer Science Selected Papers*. La Plata (Buenos Aires): Editorial de la Universidad de La Plata (edulp). 2010. p85 - 95. isbn 978-950-34-0684-7.
- [10]. Chapman B., Jost G., Van der Pas. "Using OpenMP – Portable Shared Memory Parallel Programming". 2008. UK: MIT Press.
- [11]. Hager G., Wellein G. "Introduction to HPC for Scientists and Engineers". 2011. EEUU: CRC Press.
- [12]. De Wael M., Marr S., De Fraine B., Van Cutsem T., De Meuter W. "Partitioned Global Address Space Languages". *ACM Computing Surveys* 47 (4), 2015.
- [13]. Amazon Elastic Compute Cloud (Amazon EC2). <http://aws.amazon.com/es/ec2/>. Febrero 2013.
- [14]. Balladini J., Rucci E., De Giusti A., Naiouf M., Suppi R., Rexachs D., Luque E. "Power Characterisation of Shared-Memory HPC Systems". *Computer Science & Technology Series – XVIII Argentine Congress of*

- Computer Science Selected Papers. ISBN 978-987-1985-20-3. Pp. 53-65. EDULP, La Plata (Argentina), 2013
- [15]. Sanz V., Pousa A., Naiouf M., De Giusti A. "Accelerating Pattern Matching with CPU-GPU Collaborative Computing". Proceedings of the International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2018), ISBN: 978-3-030-05051-1, págs. 310-322, doi: https://doi.org/10.1007/978-3-030-05051-1_22, 2018.
- [16]. Sanz V., Pousa A., Naiouf M., De Giusti A. "Efficient Pattern Matching on CPU-GPU Heterogeneous Systems". In: Wen S., Zomaya A., Yang L. (eds) Algorithms and Architectures for Parallel Processing. ICA3PP 2019. Lecture Notes in Computer Science, vol 11944. Springer, Cham. 2020.
- [17]. Sanz V., Pousa A., Naiouf M., De Giusti A. "Accelerating Pattern Matching on Intel Xeon Phi Processors". In: Qiu M. (eds) Algorithms and Architectures for Parallel Processing. ICA3PP 2020. Lecture Notes in Computer Science, vol 12452. Springer, Cham. https://doi.org/10.1007/978-3-030-60245-1_18.
- [18]. Rucci E., De Giusti A., Naiouf M., García Sánchez C., Botella Juan G., Prieto-Matías M.. "State-of-the-art in Smith-Waterman Protein Database Search". Big Data Analytics in Genomics. Ka-Chun Wong (Editor). ISBN: 978-3-319-41278-8 (print) 978-3-319-41279-5 (online), Springer, págs. 197-223, 2016.
- [19]. Costanzo M., Rucci E., García-Sánchez C., Naiouf M., Prieto-Matías M.. "Migrating CUDA to oneAPI: A Smith-Waterman Case Study." International Work-Conference on Bioinformatics and Biomedical Engineering 2022, ISBN 978-3-031-07802-6, págs. 103-116, Cham, doi: 10.1007/978-3-031-07802-6_9. 2022.
- [20]. Camele G., Menazzi S., Chanfreau H., Marraco A., Hasperué W., Butti M. D., Abba M.C. "Multimix: a cloud-based platform to infer cancer genomic and epigenomic events associated with gene expression modulation". Bioinformatics, Volume 38, Issue 3, 1 February 2022, Pages 866–868, <https://doi.org/10.1093/bioinformatics/btab678>
- [21]. Costanzo M., Rucci E., Costi U., Chichizola F., Naiouf M. "Comparison of HPC Architectures for Computing All-Pairs Shortest Paths. Intel Xeon Phi KNL vs NVIDIA Pascal". En: Computer Science – CACIC 2020. Revised Selected Papers., Springer International Publishing, doi. 10.1007/978-3-030-75836-3_3, 2021.
- [22]. Costanzo M., Rucci E., Naiouf M., De Giusti A. "Performance vs Programming Effort between Rust and C on Multicore Architectures: Case Study in N-Body", Proceedings of 2021 XLVII Latin American Computing Conference (CLEI), ISBN: 978-1-66549-503-5, págs. 1-10, doi. 10.1109/CLEI53233.2021.9640225, 2021.
- [23]. Milla A., Rucci E. "Performance Comparison of Python Translators for a Multi-threaded CPU-Bound Application". In: Pesado, P., Gil, G. (eds) Computer Science – CACIC 2021. Communications in Computer and Information Science, vol 1584. Springer, Cham. 2022. https://doi.org/10.1007/978-3-031-05903-2_2.
- [24]. Milla A., Rucci E. "Comparación de Rendimiento y Esfuerzo de Programación entre Numba y Cython para una Aplicación Multi-hilada de Alto Rendimiento". Congreso Argentino de Ciencias de la Computación (CACIC 2022). 2022.
- [25]. Trigila M., Gaudiani A., Luque E. "Agile Tuning Method in Successive Steps for a River Flow Simulator". In: Shi Y. et al. (eds) Computational Science – ICCS 2018. ICCS 2018. Lecture Notes in Computer Science, vol 10862. Springer, Cham.
- [26]. Trigila M., Gaudiani A., Luque E. "Adjustment of a simulator of a complex dynamic system with emphasis on the reduction of computational resources". Actas del Workshop. p. 142-147. 2018.
- [27]. Castro J., De Giusti L., Gorga G., Sánchez M., Naiouf M. "ECMRE: Extended Concurrent Multi Robot Environment". Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3, Springer, Cham, págs. 285-294. 2018.
- [28]. De Giusti L., Chichizola F., Rodríguez Eguren S., Sánchez M., Paniego J. M., De Giusti A. "Introduciendo conceptos de Cloud Computing utilizando el entorno CMRE". Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) – Workshop de Innovación en Educación en Informática. 2016.
- [29]. Basgall, M. J., Naiouf, M., & Fernández, A. "Análisis y diseño de técnicas de preprocesamiento de instancias escalables para problemas no balanceados en Big Data. Aplicaciones en situaciones de emergencias humanitarias". Tesis Doctoral. 2022.
- [30]. Sanz V., Pousa A., Naiouf M., De Giusti A. "Comparison of Hardware and Software Implementations of AES on Shared-Memory Architectures". In: Naiouf M., Rucci E., Chichizola F., De Giusti L. (eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2021. Communications in Computer and Information Science, vol 1444. Springer, Cham. 2021. https://doi.org/10.1007/978-3-030-84825-5_5.
- [31]. Sanz, V., Pousa, A., Naiouf, M., De Giusti, A. "Performance Analysis of AES on CPU-GPU Heterogeneous Systems". In: Rucci, E., Naiouf, M., Chichizola, F., De Giusti, L., De Giusti, A. (eds) Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2022. Communications in Computer and Information Science, vol 1634. Springer, Cham. 2022. https://doi.org/10.1007/978-3-031-14599-5_3
- [32]. Rucci E., Naiouf M., Chichizola F., De Giusti L., De Giusti A.. "Cloud Computing, Big Data & Emerging Topics". Communications in Computer and Information Science. Springer Cham. 2022. ISBN 978-3-031-14599-5. <https://doi.org/10.1007/978-3-031-14599-5>.
- [33]. Naiouf M., De Giusti A., Chichizola F., De Giusti L., Rucci E. (eds). "Short papers of the 10th Conference on Cloud Computing, Big Data & Emerging Topics". 2022. <https://doi.org/10.35537/10915/139373>.
- [34]. XI Jornadas de Cloud Computing, Big Data & Emerging Topics. <https://jcc.info.unlp.edu.ar/>

Análisis y modelado de sistemas utilizando Cloud Computing, IaaS privados y públicos.

Diego Encinas^{1,2}, Brian Galarza¹, Román Bond¹, Gonzalo Zaccardi¹, Nicolás Benquerença Mendes¹, Jorge Osio¹, David Duarte¹, Martín Morales^{1,3}

¹Proyecto de Investigación SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - UNAJ

²Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

³Centro CodApli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, bgalarza@unaj.edu.ar, rbond@unaj.edu.ar, gzaccardi@unaj.edu.ar, nbenquerenca@unaj.edu.ar, josio@unaj.edu.ar, davito.duarte.22@gmail.com, martin.morales@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio del rendimiento de las arquitecturas tipo cloud a través del despliegue de IaaS y utilización de IaaS públicos, en particular en el área de cómputo paralelo de altas prestaciones (HPC). Enfocando en la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente, especialmente la entrada/salida y las comunicaciones.

Palabras clave: *Cloud Computing. OpenStack. Sistemas de Archivos en clústers. Redes definidas por Software.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación “Simulación, Computación de Altas Prestaciones (HPC) y optimización de aplicaciones sociales – SimHPC” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 183/21. Además, el proyecto colabora con el Programa “Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social” – TICAPPS de la UNAJ.

Introducción

Cloud Computing es un paradigma que ha estado en constante crecimiento. Cada vez más compañías y grupos de investigación trabajan en conjunto con el fin de explotar las oportunidades ofrecidas por el mismo [1]. Dicho paradigma ofrece muchas ventajas, tales como el bajo costo de implementación, ya que no se necesitan

computadoras de última tecnología debido a que éstas trabajan conjuntamente (Clustering) con la posibilidad de escalar horizontalmente de manera sencilla. Además, hay software Open Source disponible para los nodos en el clúster como las infraestructuras Eucalyptus, OpenNebula, CloudStack u OpenStack integradas con GNU/Linux y compatibles, por ejemplo, con Amazon WebServices.

Despliegue de IaaS

Las comunicaciones en Cloud Computing son una parte fundamental del paradigma que consisten en utilizar distintos nodos y lograr hacerlos funcionar conjuntamente.

Para lograr una comunicación sincronizada entre estos nodos se propone utilizar OpenStack [2] como también OpenNebula [3].

OpenStack es una plataforma de tecnología open source que utiliza recursos virtuales agrupados para diseñar y gestionar nubes privadas y públicas a través de múltiples servicios que, de manera coordinada, cumplen diferentes propósitos para lograr el correcto funcionamiento de, por ejemplo, una "Infraestructure as a service" (IaaS). Algunos de los servicios ofrecidos por OpenStack son: hypervisor (Nova), autenticación (Keystone), Imágenes (Glance), Dashboard (Horizon), Networking (Neutron) y block storage (Cinder). Según las necesidades se pueden requerir de ciertos servicios u otros. La Arquitectura básicamente consiste en dos tipos de nodos: "Compute Node" y "Controller Node". Se llaman Compute Node a todos aquellos que se encargan del procesamiento de servicios específicos mientras que Controller Node es aquel que comunica a cada uno de los anteriores [4] [5] [6].

Fuel es una herramienta open source desarrollada por Mirantis en la cual se ejecuta un script que permite configurar, de manera más amigable respecto a OpenStack, los recursos que se desean otorgar a la infraestructura, como la cantidad de nodos, los núcleos de procesador, la memoria RAM, entre otros [7].

Fuel trabaja con un nodo master el cual es el encargado de controlar a los nodos slaves que contendrán la infraestructura OpenStack. Es decir, desde el nodo Fuel Master se indican qué paquetes se van a instalar en cada nodo slave (Glance, Nova-Compute, Keystone, etc.) para luego en los slaves tener armados los nodos compute y controller, sin necesidad de realizar configuraciones manuales en cada uno de los mismos.

OpenNebula es un software de código abierto que permite el despliegue de IaaS. Busca reducir la complejidad generada por OpenStack y ofrece soporte con hipervisores tales como KVM y VMware vCenter.

OpenNebula clasifica a los nodos en dos tipos, Front – end los cuales entran en contacto con los usuarios y a su vez se comunican con los nodos de la infraestructura en los cuales se lanzarán las instancias y los nodos virtualizados los que a su vez deben contar con los paquetes correspondientes de storage, autenticación y networking para poder funcionar correctamente.

La implementación de estas infraestructuras ofrece ventajas en las cuales los clústeres virtualizados trabajan en conjunto ofreciendo un buen rendimiento a bajos costos y con posibilidad de escalabilidad al poder agregar mayor cantidad de nodos para procesamiento de manera sencilla.

Sistemas de Archivos Paralelos en clústeres

Amazon Webservices [8], mediante el servicio EC2 (Amazon Elastic Compute Cloud) permite desplegar clústeres virtuales mediante instancias de VMs y almacenamiento para las mismas. Mediante este servicio y el correspondiente clúster conformado, se utiliza un sistema de archivos paralelo (PVFS2) que permite la gestión de datos particionados y distribuidos en los distintos nodos, mediante múltiples tareas de una aplicación ejecutada sobre el clúster.

PVFS2 utiliza una estructura cliente-servidor. Dependiendo de su rol, existen tres tipos de nodos dentro de un clúster con este sistema de archivos: servidores de datos, servidores de metadatos y clientes, en donde cada uno de los nodos puede cumplir los tres roles.

PVFS2 [9] contiene una herramienta interna que permite volcar información en logs durante la ejecución de tareas dentro del sistema de archivos. Por cada uno de los nodos, el administrador del clúster puede obtener información de depuración (GOSSIP) conformado por registros de debug, de acceso, contadores de rendimiento y errores producidos durante la ejecución. Referido a los contadores de rendimiento, se obtienen mediante el software de monitoreo Atop. Esta herramienta permite obtener reportes de la actividad de los procesos y la utilización de los diferentes recursos del sistema (memoria, disco, CPU, red, etc) [10].

Asimismo, es posible especificar los datos a obtener en función de las distintas capas de PVFS2, del rol específico de cada nodo (cliente-servidor), así como también de otros factores relacionados con el funcionamiento interno y operaciones asociadas a la gestión de archivos.

Teniendo conocimiento del funcionamiento del código fuente correspondiente a PVFS2 y del modo en el

que realiza la escritura de registros en los logs, resulta factible realizar modificaciones que permitan obtener otro tipo de parámetros de interés a partir de esta herramienta.

Finalmente se propone el análisis, uso y configuración de distintas herramientas no invasivas para determinar la performance del sistema de archivos en clústeres virtuales.

Redes definidas por Software

Las redes definidas por software (SDN) son un paradigma de gestión y administración de redes por medio de software que permiten tener un control más flexible con respecto al control del tráfico de datos por medio de hardware ya que permite cambiar en tiempo real las normas y políticas establecidas en la red [11].

Mininet[12] es un emulador de redes SDN open source que permite generar tráfico artificial entre nodos virtuales de la red.

Se ha llevado a cabo el montaje de un laboratorio SDN utilizando Mininet junto con Amazon WebServices con el fin de analizar el comportamiento de este nuevo paradigma de redes de computadoras en un entorno de Cloud Computing.[13]

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.

- Plataformas de software para implementar y administrar Clouds públicos, privados e híbridos.
- Sistemas de Archivos Paralelos.

Resultados y Objetivos

Investigación experimental

- Implementación de un IaaS encargado de realizar operaciones en procesamiento paralelo aumentando la eficiencia y reduciendo los costes generados.
- Implementación de OpenStack Dashboard y de un sistema desarrollado para poder controlar/administrar de manera visual (web) y más básica cada uno de los servicios.
- Implementación de OpenNebula en un sistema con las mismas características que el implementado por OpenStack con el fin de poder realizar pruebas en entornos similares.
- Utilización de Fuel para instalar/configurar OpenStack como sistema de administración de nube (Cloud Computing) a partir de la infraestructura de 2 nodos compute y el controller [14].
- Análisis del rendimiento de un Cloud privado en la ejecución de instancias personalizadas.
- Ejecutar diferentes benchmarks en la infraestructura desplegada sobre OpenStack y OpenNebula para así comparar los resultados obtenidos de ambas infraestructuras y poder realizar un análisis del rendimiento en cada caso.
- Introducción al estudio e integración de sensores físicos y los servicios en la nube [15].
- Medición de servicios en la nube enfocados a IaaS y PaaS [16].
- Análisis y configuración de clústeres virtuales. Análisis y configuración de herramientas no invasivas para la obtención de métricas en las distintas capas de software de los sistemas de archivos paralelos.
- Utilización de otros sistemas de archivos paralelos como Lustre [17] y Beegfs [18] para obtener métricas en Metadataservidores [19] [20] [21].

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería Informática de la UNAJ. También aportan trabajos de alumnos de las materias Sistemas Operativos 1, Redes de Computadoras 2, Programación en Tiempo Real y Organización y Arquitecturas de Computadoras. Por otro lado, algunos integrantes participan en el dictado de la Diplomatura en Ciencia de Datos de la UNAJ.

Durante 2021 se han realizado publicaciones nacionales. Además, se encuentran en desarrollo y concluidas varias Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Informática.

En esta línea de I/D existe cooperación a nivel nacional. Hay 4 investigadores realizando carreras de postgrado y alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Kondo, D., Javadi, B., Malecot, P., Cappello, F., Anderson, D. P.: "Cost-benefit analysis of Cloud Computing versus desktop grids". In: IPDPS '09 Proceedings.

- IEEE International Symposium on Parallel and Distributed Processing. Washington, USA (2009).
2. OpenStack Cloud Software: Open source software for building private and public clouds. <http://www.openstack.org>. Febrero 2015.
3. OpenNebula. <https://opennebula.org/>. Febrero 2019
4. Galarza, B.; Tuamá, C.; Zaccardi, G.; Encinas, D.; Morales, M. "Implementaciones de Cloud Computing y aplicaciones en el ámbito universitario". 1° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2013). Ciudad de Córdoba, Argentina.
5. Zaccardi, G.; Galarza, B.; Encinas, D.; Morales, M. "Implementación de Cloud Computing utilizando OpenStack". 2° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2014). Ciudad de San Luis, Argentina.
6. Galarza, B.; Zaccardi, G.; Encinas, D.; Morales, M. "Análisis de despliegue de una IaaS utilizando Openstack". XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015). Ciudad de Junín, Argentina.
7. OpenStack Deployment Fuel. <https://www.mirantis.com/products/mirantis-openstack-software/openstack-deployment-fuel/>. Febrero 2016
8. Amazon Web Services (AWS)-Cloud Computing Services. <https://aws.amazon.com> Marzo 2019
9. T. PVFS2, "PVFS 2 File System Semantics Document," tech. rep., PVFS Development Team, 2015
10. Atop Tool. <https://www.atoptool.nl/index.php>
11. Kreutz, D., Ramos, F. M. V., Esteves Verissimo, P., Esteve Rothenberg, C., Azodolmolky, S., & Uhlig, S. (2015). Software-Defined Networking: A Comprehensive Survey. *Proceedings of the IEEE*, 103(1), 14–76.
12. Mininet. <http://mininet.org>
13. Montes de Oca, F.; Galarza, B.; Morales, M.; Encinas, D. "Redes Definidas por Software en Entorno de Cloud Computing". 6° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2018). Mar del Plata, Argentina.
14. Zaccardi, G.; Galarza, B.; Morales, M.; Encinas, D. "Despliegue y ejecución de un cloud privado". 4° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2016). Ciudad de Salta, Argentina.
15. Armano, M.; Navarro, G.; Morales, M.; Encinas, D. "Utilización de servicios de Cloud Computing y sensores". 8° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.
16. González, N.; Lescano, N.; Pinto, C.; Morales, M.; Encinas, D. "Análisis de rendimiento de IaaS y PaaS". 8° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.
17. Lustre Manual. https://doc.lustre.org/lustre_manual.xhtml#idm140436306123424. Febrero 2021.
18. Heichler, Jan. "An introduction to BeeGFS." (2014).
19. Benquerença Mendes, N.; Bond, R.; Morales, M.; Encinas, D. "Rendimiento de sistema de archivos en arquitecturas distribuidas y paralelas". 8° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba.
20. E. Párraga, B. León, R. Bond, D. Encinas, A. Bezerra, S. Mendez, D. Rexachs, E. Luque. (2021) Analyzing the I/O Patterns of Deep Learning Applications. *Cloud Computing, Big Data & Emerging Topics. JCC-BD&ET 2021. Communications in Computer and Information Science*, vol 1444. Springer, Cham. https://doi.org/10.1007/978-3-030-84825-5_1
21. Benquerença Mendes, N.; Bond, R.; Morales, M.; Encinas, D. "Sistema de Archivos Paralelos con Aplicaciones de Machine Learning.". XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022). La Rioja, Argentina

Estrategias de modelado y simulación en sistemas de HPC y salud.

Diego Encinas^{1,2}, Jimena Jara¹, Román Bond¹, Daniel Rosatto¹, Diego Alvarez¹, Gustavo García Krahn¹, Braian Dorrego¹, Adriana Gaudiani³, Martín Morales^{1,4}

¹Proyecto de Investigación SimHPC - Programa TICAPPS - Instituto de Ingeniería y Agronomía - UNAJ

²Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

³Área de Computación - Instituto de Ciencias - UNGS

⁴Centro CodAppli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, elchejime@gmail.com, rbond@unaj.edu.ar, drosatto@unaj.edu.ar, diegoalvarez3219@gmail.com, krahn_gustavo@hotmail.com, rayan_22_dorrego@hotmail.com, agaudiani@ungs.edu.ar, martin.morales@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de diferentes técnicas de modelado y simulación para entornos de Cómputo en Altas Prestaciones (HPC, High Performance Computing). El enfoque del estudio es la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios y reconfigurar el sistema físico. Además, se analizan los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o reconfigurarse.

Por otra parte, se ha llevado a cabo el desarrollo de un simulador para entornos en el área de salud, en el sector de emergencias hospitalarias y de propagación de enfermedades.

Palabras clave: *Arquitecturas Multiprocesador. Simulación. Sistema E/S paralela. Modelado y Simulación basado en agentes (Agent-Based Modeling and Simulation, ABMS). Cloud Computing. CloudSim. Simulación y Salud.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación “Simulación, Computación de Altas Prestaciones (HPC) y optimización de aplicaciones sociales –

SimHPC” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 183/21. Además, el proyecto aporta al Programa “Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social” – TICAPPS de la UNAJ.

En el tema existe un convenio de colaboración en actividades de Investigación y Postgrado con el Instituto de Investigación en Informática – LIDI de la Universidad Nacional de La Plata y el Área de Computación del Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

Introducción

El crecimiento sostenido en la demanda del poder de cómputo remarca la necesidad de sistemas con enfoques de paralelización masiva y cómputo de alta performance (HPC, High Performance Computing) [1]. Los clusters se han convertido en uno de los enfoques principales para lograr paralelismo a bajo costo. Una noción extendida lo constituye la utilización de grid computing y más recientemente cloud computing. Independientemente de la solución, estos sistemas constan de un gran número de componentes incluyendo nodos de procesamiento, bancos de memoria, discos, entre otros.

En cuanto a las herramientas de simulación para Cloud Computing, CloudSim es un framework desarrollado en Java que provee las APIs necesarias para que el usuario genere una simulación de un Data Center funcionando como servidor de nube capaz de simular la ejecución de CloudLets.

Por otra parte, los servicios de urgencias hospitalarias son considerados como una de las unidades del sistema sanitario de mayor complejidad y fluidez, lo que unido a la variabilidad de su actividad da lugar a que su gestión operativa sea una tarea muy complicada [2]. Es por ello que resultaría de mucha utilidad para sus responsables disponer de un sistema de ayuda a la toma de decisiones (Decision Support System-DSS) tan flexible como un simulador, que permitiría tomar medidas disponiendo de información suficiente sobre las alternativas posibles.

Sistemas de E/S Paralela

Las exigencias en los sistemas de E/S paralelos se han incrementado debido al aumento en número, velocidad y potencia de las unidades de procesamiento en los clusters. También las aplicaciones científicas que utilizan cómputo de altas prestaciones acrecientan estos requerimientos.

En muchos casos, el cuello de botella de los sistemas paralelos es la E/S a causa de las exigencias que debe afrontar [3]. La E/S Paralela es esencial para emparejar el avance de las arquitecturas de los procesadores y el rápido crecimiento de la capacidad computacional. Aunque la arquitectura jerárquica de memoria multinivel puede evitar grandes pérdidas de prestaciones debido a los retardos de acceso a disco, la capacidad de memoria es limitada. Además, como la capacidad computacional aumentará, la disponibilidad de memoria por core decrecerá, especialmente si la escala de los sistemas de HPC se proyecta a millones de cores o más. Varias simulaciones científicas y de ingeniería de áreas críticas de investigación, tales como la nanotecnología, astrofísica, clima y energía física están convirtiéndose en aplicaciones intensivas de datos. Para poder disminuir la brecha entre CPUs-E/S se deben identificar los factores que influyen en las

prestaciones y proponer nuevas soluciones [4] [5].

En el área de tolerancia a fallas en sistemas de cómputo de alta prestaciones se puede notar la importancia de la unidad de E/S en las arquitecturas paralelas como un punto a mejorar para lograr cubrir las exigencias de las aplicaciones que utilizan HPC. Una manera de llevar a cabo este trabajo es utilizar técnicas de simulación para evaluar el efecto de los cambios de los factores con mayores influencias en las prestaciones del sistema de E/S paralelo.

Se puede disminuir la complejidad y la probabilidad de errores en la generación de sistemas híbridos desarrollando una simulación específica de éstos utilizando diferentes frameworks [6] [7] [8].

Las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos. Contar con una herramienta que permita predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores de aplicaciones paralelas como para administradores de centros de cómputo. Por otro lado, ABMS ha sido utilizado para modelar problemas y sistemas complejos en diversas áreas de la ciencia.

Evaluar las prestaciones del subsistema de E/S con diferentes configuraciones y la misma aplicación, permite adaptar la configuración de E/S teniendo en cuenta el patrón de acceso de la aplicación. Pero también puede ser una gran ventaja analizar las necesidades de las aplicaciones antes de configurar el sistema físico. Una manera de predecir el comportamiento de las aplicaciones en el sistema de cómputo, ante distintas configuraciones, es utilizando técnicas de modelado y simulación.

Se está desarrollando modelos e implementando una simulación de la arquitectura de E/S paralela, por medio de técnicas de simulación basadas en agentes o Sistemas Multi-Agente, (MAS-MultiAgentSystems), para evaluar el efecto de dimensionar el sistema de E/S o cambiar componentes como la red de almacenamiento, dispositivos de E/S, entre otros [9].

Simulación de arquitecturas de Cloud Computing

CloudSim [10] [11] es un Framework de simulación generalizado y extensible que permite el modelado y la simulación de diferentes infraestructuras y servicios de aplicaciones de Cloud Computing. Un ejemplo de utilización es la simulación de muchos centros de datos.

Su arquitectura consiste en entidades específicas que se representan como clases Java que pueden ser heredadas o instanciadas. Estas clases representan centros de datos, hosts físicos, máquinas virtuales, servicios a ejecutar en los centros de datos y servicios en la nube de usuarios [12] [13]. Además, CloudSim soporta la inserción dinámica de los elementos de simulación y proporciona aplicaciones de paso de mensajes y la topología de la red del centro de datos.

La versatilidad de CloudSim es la principal ventaja del sistema. La integración de nuevos parámetros y conceptos de la simulación es implementada desde abstracciones preestablecidas convenientemente por los autores. Las abstracciones principales son SimEvent [13], SimEntity [14], DataCenterCharacteristics y Vm.

El aporte de un desarrollo de nuevos actores al componente Vm (Virtual Machine) posibilita que por medio de simulación se obtengan métricas de entrada/salida. Las estadísticas que aporta la nueva implementación de Vm dan soporte a un espacio de memoria ram en tiempos de simulación. Dicho espacio de memoria principal está controlado por otro componente que es capaz de procesar instrucciones guardadas en el espacio de memoria sintético denominado RamEntity. Si las instrucciones se guardan lógicamente, la nueva versión de Vm es capaz de administrar procesos en la nueva capa de ejecución. La idea general de la implementación consiste en la creación de una nueva SimEntity y sus agregaciones necesarias para manipular el espacio de memoria proporcionada por RamEntity, en tiempos de simulación. De tal forma que cada celda de RamEntity no sólo guarda datos, sino que también es posible que mantenga objetos interpretables como eventos para la nueva SimEntity que se denomina InitEntity.

Lo destacable es que el entorno de la cola de procesos es en el espacio de memoria de

RamEntity y que cada instrucción que compone a un código objeto del proceso es en realidad un evento con todos sus parámetros.

Simulación y Salud

Al analizar las necesidades presentes en los servicios de salud, se obtiene que dentro de una sala de urgencias se encuentre una gran diversidad de escenarios posibles donde cada uno de estos puede afectar a resultados sensibles como, por ejemplo, la tasa de mortalidad de personas. Para solucionar este problema se desarrolló un simulador con el objetivo de ser una herramienta capaz de recrear una gran cantidad de escenarios y así poder tomar decisiones rápidas.

El simulador desarrollado se centra en la prevalencia puntual de infecciones intrahospitalarias [15] en una sala de urgencias y cómo la afectan distintos factores relacionados con la gestión hospitalaria.

Para llevar a cabo el modelado del simulador se utilizó el paradigma de Modelado y Simulación basado en Agentes (ABMS). El framework utilizado es Repast Symphony [16], una herramienta especializada en ABMS, la cual provee una serie de ventajas respecto al modelado e implementación de agentes, así como también la coordinación concurrente de los agentes.

De esta manera, se clasificaron distintos agentes intervinientes en el ambiente de una sala de urgencias, como los pacientes y médicos, entre otros. Dichos agentes definen su comportamiento mediante máquinas de estado, las cuales determinan las acciones correspondientes tanto a la atención hospitalaria como también al estado de salud respecto de una enfermedad intrahospitalaria específica.

La propagación de la infección intrahospitalaria estará modelada mediante interacciones entre estos agentes, por ejemplo, uno de los focos de infección más importante de las salas de urgencias son las salas de espera. La forma más efectiva para calibrar el simulador se da mediante el grado de interacción de agentes, ajustando así la tasa de transmisión de la enfermedad en cuestión.

Una vez calibrado el simulador, se pueden obtener resultados. El trabajador de la salud dispone de diversos parámetros para configurar

la simulación, por ejemplo, la cantidad de pacientes, la cantidad de camas disponibles, la cantidad de insumos hospitalarios, etc. Al realizar distintas ejecuciones, se pueden obtener resultados analizando distintos posibles cuellos de botella, configurando la cantidad de médicos clínicos disponibles, la cantidad de recepcionistas encargados de la admisión, el triage, entre otros. De esta manera, es posible la toma de decisiones respecto a la asignación de recursos y personal para agilizar la estancia hospitalaria de los pacientes y evitar los contagios producidos por sus interacciones.

Además, se analiza y trabaja la optimización y paralelización del simulador con los ambientes Repast HPC para clusters y Flame para ejecución sobre GPU, obteniendo mejor performance en términos de tiempos.

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.
- Arquitectura de E/S paralela considerando el software, hardware, comunicaciones entre módulos y dispositivos de almacenamiento.
- Nuevos aportes de desarrollos que mejoren los modelos de simulaciones con CloudSim para el análisis de la performance en sistemas de arquitecturas de software de Cloud Computing.
- Modelado y simulación para la administración de sistemas de salud.

Resultados y Objetivos

Investigación experimental

- Diseño y desarrollo de modelos mediante técnicas de Modelado y simulación basada en agentes (ABMS) para analizar el comportamiento de las distintas capas de la pila de software de E/S.
- Análisis, modelado e implementación de las operaciones típicas de E/S: read, write,

open, close, flush. Se ha utilizado Amazon Web Services para creación de cluster virtuales y obtener métricas de la pila de software de E/S.

- Incorporación de tiempos de entrenamiento, obtenidos en AWS, en el simulador. Con esto se logra una salida más detallada y un método para validar tiempos y métricas del simulador con AWS.
- Implementación de comandos para ejecutar desde command center en NetLogo. Con esto se logró sintetizar el benchmark IOR correspondiente a la capa de aplicación de la pila de E/S, logrando introducir nuevos parámetros como tamaño de archivo y cantidad de nodos de E/S (metadata server y data server). De ejecutar este comando, se obtiene una nueva salida similar a la del benchmark IOR [17].
- Obtención de un método de desarrollo de nuevos actores genéricos CloudSim que mejoran el modelado y la producción de estadísticas virtuales. Implementación de la entidad InitEntity que procesa instrucciones en el espacio de memoria de las máquinas virtuales. Se vinculan exitosamente los tiempos de procesamiento de un cloudlet y las instrucciones en las máquinas virtuales [18].
- Diseño de prototipos de pruebas para estudiar entrada/salida como en el área de salud, utilizando microcontroladores [19] [20].
- Contraste de una simulación de un cluster en la nube y uno idéntico desplegado en un sistema de cloud computing público [21].
- Modelado de infraestructuras de sistemas de salud. Implementación de un simulador para analizar el contagio de enfermedades intrahospitalarias [22].
- Paralelización en la ejecución de simuladores.
- Implementación de un simulador para analizar la propagación de enfermedades [23].

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería en Informática de la UNAJ. También aportan

trabajos de alumnos de las materias Redes de Computadoras 2 y Programación en Tiempo Real. Por otro lado, algunos integrantes participan en el dictado de la Diplomatura en Ciencia de Datos de la UNAJ.

Durante 2022 se han realizado publicaciones nacionales e internacionales. Además, se encuentran en desarrollo y concluidas varias Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Informática.

En esta línea de I/D existe cooperación a nivel nacional e internacional. Hay dos investigadores realizando estudios de postgrado, 1 becario EVC CIN, un becario de Iniciación a la Investigación UNAJ, 1 becario Manuel Belgrano y 2 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Second Edition. Pearson Addison Wesley, 2003.
2. R. Galeano, C. Villalba, D. Rexachs, E. Luque. Agent-Based Model to Simulate Outpatient's Consultations at the "Hospital de Clínicas". The Eighth International Conference on Advances in System Simulation (SIMUL 2016). 1:46-51
3. H Hennessy, J. L., Patterson, and D. A., Computer Architecture, Fourth Edition: A Quantitative Approach. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2006.
4. J. M. May, Parallel I/O for high performance computing. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001.
5. V. Balaji, Earth system modelling – Volume 4. IO and Postprocessing. Springer, 2013.
6. D. Encinas, Utilización de un reloj global para el modelado de un ambiente simulado distribuido. XVIII Congreso Argentino de Ciencias de la Computación. 2012
7. D. Encinas, Simulación de una red CAN para dimensionar las comunicaciones de una IMU. VII Congreso Argentino de Tecnología Espacial. 2013.
8. D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
9. D. Encinas et al., Modeling I/O System in HPC: An ABMS Approach. The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015.
10. R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose and R. Buyya "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms" Published online 24 August 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/spe.995.
11. <http://www.cloudbus.org/cloudsim> 2018.
12. Hamza Ouarnoughi, Jalil Boukhobza, Frank Singhoff, Stephane Rubini, Erwann Kassis. "Considering I/O Processing in CloudSim for Performance and Energy Evaluation". OpenStack Cloud Software: Open source software for building private and public clouds. © Springer International Publishing AG 2016 M. Tauber et al. (Eds.): ISC High Performance Workshops 2016, LNCS 9945, pp. 591–603, 2016. DOI: 10.1007/978-3-319-46079-6_40.
13. www.icsa.inf.ed.ac.uk/research/groups/hase/simj ava 2018.
14. F. Howell, R Mc Nab. A discrete event simulation library for java. International Conference on Web-Based Modeling and Simulation. 1998.
15. Prevalencia puntual de infección nosocomial, disponible en internet: https://www.researchgate.net/publication/242363760_Prevalencia_puntual_de_infeccion_nosocomial. Fecha: 29/03/2020
16. Repast Symphony Frequently Asked Questions, disponible en internet: <https://repast.github.io/docs/RepastFAQ/RepastFAQ.html>. Fecha: 29/03/2020
17. D. Encinas, S. Mendez, M. Naiouf, A. De Giusti, D. Rexachs del Rosario, and E. Luque, An Agent-Based Model for Analyzing the HPC Input/Output System International journal on advances in systems and measurements vol. 13, num. 3 & 4, págs. 192-202, 2020.
18. D. Rosatto, R. Bond, M. Belizán, M. Morales, D. Encinas. Modelado y simulación de arquitecturas de Cloud Computing con CloudSim: comunicación entre entidades. XXIII Congreso Argentino de Ciencias de la Computación. ISBN 978-950-34-1539-9. 2017
19. L. Barrera, M. Rodriguez, R. Bond, M. Morales, D. Encinas Diseño de un oxímetro de pulso. Prototipo de pruebas. XXVIII Congreso Argentino de Ciencias de la Computación. ISBN 978-987-1364-31-2. 2022
20. M. López, R. Bond, M. Morales, D. Encinas. Diseño de un prototipo de pruebas de Entrada/Salida con Arduino. 10º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2022). Entre Ríos. 2022
21. T. Rosales, J. Spinelli, M. Di Nardo, R. Bond, D. Rosatto, D. Encinas, F. Romero. Análisis de una plataforma de simulación para Cloud Computing. Un caso de estudio. XXVI Congreso Argentino de Ciencias de la Computación. 2020
22. L. Maccallini, D. O. Encinas, and F. Romero. "An Approach to the Modeling and Simulation of Intra-Hospital Diseases". Journal of computer science and technology (ISSN 1666-6038), vol. 21, num. 2, págs. 157-169, doi. 10.24215/16666038.21.e14, 2021.
23. J. Baez, A. Barreto, B. Galarza, M. Morales, D. Encinas. Simulación para estimar propagación de enfermedades. 8º Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2020). San Francisco, Córdoba. 2020

Arquitecturas Multiprocesador en HPC: Software de Base, Modelos y Aplicaciones

Armando De Giusti ⁽¹⁾⁽²⁾ , Marcelo Naiouf⁽¹⁾ , Fernando G. Tinetti ⁽¹⁾⁽³⁾ , Horacio Villagarcía ⁽¹⁾⁽³⁾ , Franco Chichizola⁽¹⁾ , Laura De Giusti⁽¹⁾⁽³⁾ , Enzo Rucci⁽¹⁾⁽³⁾ , Adrián Pousa⁽¹⁾ , Victoria Sanz ⁽¹⁾⁽³⁾ , Diego Montezanti ⁽¹⁾ , Diego Encinas ⁽¹⁾ , Ismael Rodríguez⁽¹⁾ , Sebastián Rodríguez Eguren⁽¹⁾ , Leandro Libutti⁽¹⁾ , Manuel Costanzo ⁽¹⁾ , Juan Manuel Paniego⁽¹⁾ , Martín Pi Puig⁽¹⁾ , César Estrebow⁽¹⁾ , Javier Balladini⁽⁴⁾ 

¹Instituto de Investigación en Informática LIDI (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata – Comisión de Investigaciones Científicas de la
Provincia de Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴Universidad Nacional del Comahue

{degiusti,mnaiouf,fernando,hvw,francoch,ldgiusti,erucci,apousa,vsanz,dmontezanti,dencinas,ismael,seguren,
llibutti,mcostanzo,jmpaniego,mpipuig,cesarest}@lidi.info.unlp.edu.ar; javier.balladini@gmail.com

Resumen

El eje de esta línea de I/D lo constituye el estudio de las arquitecturas multiprocesador que integran sistemas distribuidos y paralelos. Incluye como temas centrales:

- Arquitecturas many-core (GPU, procesadores MIC, TPUs), FPGAs, híbridas (diferentes combinaciones de multicores y aceleradores), y asimétricas.
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas y su evaluación de rendimiento computacional y energético.
- Estudio y optimización de código heredado.
- Desarrollo y evaluación de estrategias de resiliencia.
- Modelado y simulación de E/S en HPC.

Palabras clave: *Sistemas Paralelos. Clusters. Arquitecturas asimétricas. GPU, FPGA, TPU. Eficiencia energética. Resiliencia. Portabilidad. Código heredado. E/S paralela.*

Contexto

La línea de investigación que se presenta es parte del proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de Rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real.” del III-LIDI y de proyectos específicos apoyados por organismos nacionales e

internacionales. También es parte del proyecto “Procesamiento Eficiente de Grandes Datos mediante Cómputo de Altas Prestaciones, Fog y Edge” financiado por la Facultad de Informática de la UNLP y del proyecto “Unidad Inteligente para Control de Consumo Energético” financiado por la Secretaría de Políticas Universitarias y la UNLP.

En los temas relacionados existen cooperaciones con varias Universidades de Argentina, y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por ERASMUS, CyTED y la OEI (Organización de Estados Iberoamericanos). En particular, se colabora con el proyecto “Computación de Altas Prestaciones Eficiente y Segura para Aplicaciones de Servicios de Salud Inteligentes”, de la Universidad Autónoma de Barcelona.

Por otra parte, se cuenta con financiamiento de diferentes empresas de Argentina, en particular para la formación de recursos humanos en las temáticas de Cloud, Fog y Edge Computing.

Se participa en iniciativas como el Consorcio de I+D+I en Cloud Computing, Big Data y Emergent Topics, conformado por Universidades de Argentina y España.

Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del Ministerio de Educación, Cultura, Ciencia y Tecnología de la Nación.

Introducción

Una de las áreas de creciente interés lo constituye el cómputo de altas prestaciones, en el cual el rendimiento está relacionado con dos aspectos: por un lado, las arquitecturas de soporte, y por otro, los algoritmos que hacen uso de éstas.

A la aparición de arquitecturas *many-core* (como las GPU o los procesadores MIC), se ha sumado el uso de FPGAs, debido a su potencia de cómputo y rendimiento energético. Su combinación en sistemas HPC da lugar a plataformas híbridas con diferentes características [22]. Lógicamente, esto trae aparejada una revisión de los conceptos del diseño de algoritmos paralelos (incluyendo los mismos lenguajes de programación y el software de base), así como la evaluación de las soluciones que éstos implementan.

Si bien la heterogeneidad del hardware permite maximizar las ganancias de rendimiento y eficiencia energética, implica el uso de lenguajes, librerías y herramientas específicas para cada una de ellas, lo que aumenta el costo de desarrollo y mantenimiento del software [21].

GPUs y Cluster de GPUs

Las GPUs son el tipo de acelerador dominante en la comunidad de HPC hoy en día debido a su alto rendimiento y bajo costo de adquisición. En la actualidad, tanto NVIDIA como AMD trabajan especialmente en mejorar la eficiencia energética de sus placas y disminuir el alto costo de programación.

La combinación de GPUs con otras plataformas paralelas, como clusters y multicores, brindan un vasto conjunto de posibilidades de investigación en arquitecturas híbridas, a partir de diferentes combinaciones como son:

- Máquinas multicore con más de una GPU, en las que se combinan herramientas de programación paralela como OpenMP/CUDA o Pthread/CUDA.
- Cluster de máquinas multicore, cada una con una o más placas de GPU, lo que permite

combinar OpenMP/MPI/CUDA o Pthread/MPI/CUDA.

Los desafíos que se plantean son múltiples, sobre todo en lo referido a distribución de datos y procesos en tales arquitecturas híbridas, a fin de optimizar el rendimiento de las soluciones.

FPGAs

Una FPGA (*Field Programmable Gate Array*) es una clase de acelerador basado en circuitos integrados reconfigurables. La capacidad de adaptar sus instrucciones de acuerdo con la aplicación objetivo le permite incrementar la productividad de un sistema y mejorar el rendimiento energético para ciertos tipos de aplicaciones. Tradicionalmente han sido utilizadas para el procesamiento digital de señales. Sin embargo, en los últimos años, existen tres tendencias claras para extender su uso a otros dominios. En primer lugar, el establecimiento de alianzas estratégicas entre fabricantes de procesadores y de FPGAs para integrar estos dispositivos en arquitecturas híbridas (Intel con Altera; IBM con Xilinx) [10][11]. En segundo lugar, el desarrollo de nuevas herramientas de programación para FPGAs empleando estándares familiares para HPC, con las cuales se espera reducir los tradicionales tiempos y costos de programación [27][31]. Por último, la incorporación de FPGAs a los servicios de Cloud abre nuevas oportunidades para la explotación de esta clase de aceleradores.

TPUs

Las unidades de procesamiento tensorial (TPU) son una clase de Circuitos Integrados de Aplicación Específica (ASIC) desarrolladas por Google con el propósito de acelerar las cargas de trabajo de aprendizaje automático que requieren las aplicaciones desarrolladas en su framework TensorFlow [29]. Su uso provee una alternativa a otras arquitecturas ya conocidas como CPUs, GPUs y MICs. En ese sentido, interesa analizar las tasas de aceleración y eficiencia energética provistas por esta nueva arquitectura, en comparación con el resto.

Eficiencia energética

La mejora de la eficiencia energética es una de las principales preocupaciones en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores. Muchos esfuerzos están orientados a tratar la eficiencia energética y a las metodologías para medirla como ejes de I/D, como una métrica de evaluación relevante.

Entre los puntos de interés pueden mencionarse:

- Análisis de metodologías y herramientas para medir y optimizar el consumo energético.
- Estudio de técnicas para reducir el consumo energético en aplicaciones de HPC de acuerdo con las arquitecturas utilizadas.
- Evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas.
- Optimización de la eficiencia energética. A partir de los valores de energía que brindan los contadores hardware es posible definir estrategias de programación que lleven a reducir el consumo, manteniendo a su vez el rendimiento en valores aceptables [25].

Código heredado

La mayoría de los programas de simulación numérica que se emplean hoy en día fueron desarrolladas cuando las arquitecturas paralelas no existían. Por ello, este conjunto de aplicaciones abre oportunidades para desarrollar técnicas y herramientas que permitan optimizar el código, tanto desde el punto de vista computacional como desde la ingeniería de software [28].

Resiliencia

En la actualidad, lograr sistemas resilientes resulta un verdadero desafío considerando el creciente número de componentes, la cercanía a los límites físicos en las tecnologías de fabricación y la complejidad incremental del software. La corrección de las aplicaciones y la eficiencia en su ejecución se torna más importante en HPC debido a los extensos tiempos de ejecución. En ese sentido, resulta relevante desarrollar estrategias de detección y

recuperación de fallos, especialmente a través de librerías de software [15]. Por otro lado, resulta de interés la aplicación de mecanismos de resiliencia en aplicaciones distribuidas [34].

Simulación y Entrada/Salida paralela

A pesar de los avances tecnológicos, las operaciones de E/S en los centros de supercómputo siguen siendo un cuello de botella para determinadas aplicaciones HPC. El rendimiento de un sistema depende de la carga de trabajo (patrones de E/S de las aplicaciones) y de su configuración (hardware y software) [19]. Contar con herramientas que permitan modelar y predecir el comportamiento de este tipo de aplicaciones en HPC resulta fundamental para mejorar su rendimiento [5].

El análisis y diseño de modelos de simulación basados en la arquitectura de E/S paralela, permite disminuir la complejidad y cubrir las exigencias de las aplicaciones en HPC, al poder identificar y evaluar los factores que influyen en las prestaciones [6]. Además, considerando la relevancia que ha cobrado la simulación de transmisión de enfermedades intra-hospitalarias, resulta de interés utilizar la experiencia desarrollada para el modelado de estos sistemas [13] [32].

Dispositivos de bajo costo con capacidades para cómputo paralelo

En la actualidad se comercializan placas de bajo costo como Raspberry PI [20] u Odroid [33], que poseen múltiples núcleos simples. Asimismo, existen diversos dispositivos móviles con capacidades similares. Es de interés estudiar la manera de explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones [1].

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas many-core (CPU, GPU y TPU) y FPGA. Análisis de este tipo de máquinas y de técnicas para desarrollar código optimizado.

- Arquitecturas híbridas (diferentes combinaciones de clusters, multicores, manycores y FPGAs). Diseño de algoritmos paralelos sobre las mismas. Técnicas de resiliencia.
- Exploración de nuevos lenguajes, modelos y estándares de programación para HPC.
- Consumo energético en las diferentes arquitecturas de alto desempeño, en particular en relación con los algoritmos paralelos y la configuración de la arquitectura. Análisis de metodologías y herramientas de medición. Modelado y estimación del consumo de potencia de arquitecturas HPC.
- Análisis y desarrollo de modelos e implementación de simuladores de la pila de software de E/S en HPC.

Resultados y Objetivos

Investigación experimental por realizar

- Desarrollar y evaluar algoritmos paralelos sobre nuevas arquitecturas paralelas. Analizar rendimiento, eficiencia energética y costo de programación.
- Analizar las capacidades de lenguajes no convencionales para procesamiento paralelo, considerando rendimiento y costo de programación.
- Realizar el desarrollo de nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [24][25].
- Analizar proceso de migración de algoritmos entre arquitecturas diferentes, considerando portabilidad, rendimiento y esfuerzo de programación.
- Calibrar y sintonizar el modelado y simulación de E/S en HPC conseguido para distintos escenarios e infraestructuras
- Desarrollar técnicas de tolerancia a fallas que permitan aumentar la resiliencia de sistemas paralelos y distribuidos.

Resultados obtenidos

- Se compararon soluciones paralelas para el método de cifrado AES en diversas arquitecturas de memoria compartida [26].

- Se evaluaron capacidades de paralelismo de lenguajes no convencionales en arquitecturas multicore, como Python [14][23] y Rust [3].
- Se realizó un estudio de migración de códigos CUDA a DPC++ usando el ecosistema oneAPI para el ámbito genómico y se evaluó el esfuerzo de programación requerido, la portabilidad y el overhead en la ejecución [2].
- Se realizaron y analizaron modificaciones al framework TensorFlow para permitir la maleabilidad de hilos [12].
- Se analizó la paralelización de un prototipo de simulador de transmisión de enfermedades intra-hospitalarias [13].
- Se exploró el uso de diferentes microcontroladores para aplicaciones de aprendizaje automático [7] [8].
- Se evaluó el impacto de las comunicaciones en un cluster heterogéneo de placas RPi [30]
- Se desarrolló y validó un modelo estadístico para consumo de potencia en placas RPi de diferentes generaciones [18].
- Se diseñó la herramienta SEDAR y evaluó su desempeño para detección y recuperación de fallos transitorios [15].
- Se desarrollaron técnicas de modelado y simulación de E/S en HPC que permiten predecir cómo los cambios realizados en los diferentes componentes afectan a la funcionalidad y al rendimiento del sistema [5][6].
- Se desarrolló una plataforma distribuida, escalable, híbrida y redundante que permite aprovechar recursos de cómputo basados en arquitecturas ARM a través de un servicio de cloud [9].
- Se analizó el impacto de la E/S en aplicaciones de aprendizaje automático profundo [19].
- Se desarrolló un modelo que permite predecir el consumo energético de un sistema al que, cuando ocurre una falla, se le aplican diferentes estrategias de ahorro de energía [16].

Organización de Eventos

En el año 2022 se han organizado las X Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2022) en

Argentina, con participación de especialistas académicos del país y el exterior, además de empresas con experiencia en Cloud Computing [17] [4]. En junio de 2023 se tiene prevista la organización de las XI JCC-BD&ET.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyeron 2 Tesis Doctorales. Al mismo tiempo se encuentran en curso 3 tesis de Doctorado en Ciencias Informáticas y 3 tesis de Maestría.

Además, se participa en el dictado de las carreras de Doctorado en Ciencias Informáticas, y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática de la UNLP (acreditadas por la CONEAU con categoría A, B y A, respectivamente), por lo que potencialmente pueden generarse nuevas Tesis de Doctorado y Maestría, además de Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior, y hay tesis de diferentes Universidades realizando su Tesis con el equipo del proyecto.

Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación tres materias directamente relacionadas con los temas de ésta: “Taller de Programación sobre GPUs”, “Cloud Computing y Cloud Robotics” y “Conceptos y Aplicaciones en Big Data”.

Referencias

[1] P. S. Rodríguez Eguren, F. Chichizola, and E. Rucci, “Análisis del uso de un cluster de Raspberry Pi para cómputo de alto rendimiento”. Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018), ISBN: 978-950-658-472-6, págs. 134-144, 2018.

[2] Manuel Costanzo, Enzo Rucci, Carlos García-Sánchez, Marcelo Naiouf, Manuel Prieto-Matías. “Migrating CUDA to oneAPI: A Smith-Waterman Case Study.” International Work-Conference on Bioinformatics and Biomedical Engineering 2022, ISBN 978-3-031-07802-6, págs. 103-116, Cham, doi: 10.1007/978-3-031-07802-6_9, junio de 2022.

[3] M. Costanzo, E. Rucci, M. Naiouf, and A. D. Giusti, “Performance vs Programming Effort between Rust and

C on Multicore Architectures: Case Study in N-Body”, Proceedings of 2021 XLVII Latin American Computing Conference (CLEI), ISBN: 978-1-66549-503-5, págs. 1-10, doi: 10.1109/CLEI53233.2021.9640225, 2021.

[4] Marcelo Naiouf, Armando Eduardo De Giusti, Franco Chichizola, Laura Cristina De Giusti, Enzo Rucci (editores). “Short papers of the 10th Conference on Cloud Computing, Big Data & Emerging Topics”. ISBN 978-950-34-2126-0, Facultad de Informática (UNLP), 106 páginas, doi: 10.35537/10915/139373, julio de 2022.

[5] D. Encinas, M. Naiouf, A. De Giusti, S. Méndez, D. Rexachs del Rosario, and E. Luque, “On the Calibration, Verification and Validation of an Agent-Based Model of the HPC Input/Output System” Proceedings of the The Eleventh International Conference on Advances in System Simulation (SIMUL 2019), ISBN: 978-1-61208-756-6, págs. 14-21, 2019.

[6] D. Encinas, S. Mendez, M. Naiouf, A. De Giusti, D. Rexachs del Rosario, and E. Luque, “An Agent-Based Model for Analyzing the HPC Input/Output System” International journal on advances in systems and measurements vol. 13, num. 3 & 4, págs. 192-202, 2020.

[7] César A. Estrebou, Martín Fleming, Marcos D. Saavedra, Federico Adra, Armando E. De Giusti. “Lightweight Convolutional Neural Networks Framework for Really Small TinyML Devices.”. Third International Conference on Smart technologies, Systems and Applications (SmartTech-IC), ISBN 978-3-030-99170-8, págs. 3-16, Cham, doi: 10.1007/978-3-030-99170-8_1, marzo de 2022.

[8] César Armando Estrebou, Marcos David Saavedra, Federico Adra, Martín Fleming. “TinyML for Small Microcontrollers.” X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022), ISBN 978-950-34-2126-0, págs. 42-46, julio de 2022.

[9] David Petrocelli, Armando De Giusti, Marcelo Naiouf. “Collaborative, Distributed, Scalable and Low-Cost Platform Based on Microservices, Containers, Mobile Devices and Cloud Services to Solve Compute-Intensive Tasks.” European Conference on Parallel Processing 2021, ISBN 978-3-031-06156-1, págs. 545-548, Cham, doi: 10.1007/978-3-031-06156-1_47, junio de 2022.

[10] IBM. “IBM and Xilinx Announce Strategic Collaboration to Accelerate Data Center Applications”. Disponible en <https://www-03.ibm.com/press/us/en/pressrelease/48074.wss>

[11] Intel. “Intel Acquisition of Altera”. Disponible en <http://intelacquiesaltera.transactionannouncement.com>

[12] L. Libutti, L. De Giusti, and M. Naiouf. “PLANIFICACIÓN Y ACELERACIÓN DE ALGORITMOS DE MACHINE LEARNING”. Investigación joven (ISSN 2314-3991), vol. 7, num. 2, págs. 552-553, abril de 2021.

[13] Martin Paradiso, Lucas Maccallini, Agustina Vericat, Fernando Romero, Diego Encinas. An

Approach to Parallelization of Respiratory Disease Spread Simulations in Emergency Rooms. 2022 International Conference on Computational Science and Computational Intelligence (CSCI'22). En prensa.

[14] Milla, A., Rucci, E. (2022). Performance Comparison of Python Translators for a Multi-threaded CPU-Bound Application. In: Pesado, P., Gil, G. (eds) Computer Science – CACIC 2021. CACIC 2021. Communications in Computer and Information Science, vol 1584. Springer, Cham. https://doi.org/10.1007/978-3-031-05903-2_2

[15] D. Montezanti, E. Rucci, A. D. De Giusti, M. Naiouf, D. Rexachs, and E. Luque, “Soft errors detection and automatic recovery based on replication combined with different levels of checkpointing”. *Future generation computer systems* (ISSN 0167-739X), vol. 113, págs. 240-254, doi. <https://doi.org/10.1016/j.future.2020.07.003>, 2020.

[16] M. Morán, J. Balladini, D. Rexachs, and E. Rucci. “Some Issues to Consider in the Management of Energy Consumption in HPC Systems with Fault Tolerance”. Pertenece al libro: *Short papers of the 10th Conference on Cloud Computing, Big Data & Emerging Topics*, ISBN: 978-950-34-2126-0. Páginas: 17-22, 2022.

[17] Enzo Rucci, Marcelo Naiouf, Franco Chichizola, Laura De Giusti, Armando De Giusti. “Cloud Computing, Big Data & Emerging Topics. 10th Conference, JCC-BD&ET 2022, La Plata, Argentina, June 28–30, 2022, Proceedings”. ISBN 978-3-031-14598-8. <https://doi.org/10.1007/978-3-031-14599-5>

[18] J. M. Paniego, L. Libutti, M. P. Puig, F. Chichizola, L. De Giusti, M. Naiouf, and A. De Giusti, “Unified Power Modeling Design for Various Raspberry Pi Generations Analyzing Different Statistical Methods”. En: *Computer Science – CACIC 2019. communications in Computer and Information Science.*, ISBN: 978-3-030-48325-8, Springer International Publishing, págs. 53-65, 2020.

[19] E. Párraga, B. León, R. Bond, D. Encinas, A. Bezerra, S. Mendez, D. Rexachs, and E. Luque. “Analyzing the I/O Patterns of Deep Learning Applications”. *Proceedings from the 9th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2021)*, vol. 1444, págs. 3-16, doi. 10.1007/978-3-030-84825-5_1, 2021.

[20] Raspberry PI. <https://www.raspberrypi.org/>

[21] Manuel Costanzo, Enzo Rucci, Carlos García-Sánchez, Marcelo Naiouf. “Early Experiences Migrating CUDA codes to oneAPI.” *IX Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2021)*, ISBN 978-950-34-2016-4, págs. 14-18, julio de 2021.

[22] Rucci, Enzo: “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.

[23] Andrés Milla, Enzo Rucci. “Comparación de Rendimiento y Esfuerzo de Programación entre Numba

y Cython para una Aplicación Multi-hilada de Alto Rendimiento”. *Congreso Argentino de Ciencias de la Computación (CACIC 2022)*. En prensa.

[24] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias: “ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems”. In: *ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing)*. 2015.

[25] Saez, J.C., Pousa, A., Rodríguez-Rodríguez, R., Castro, F., Prieto-Matias, M. “PMCTrack: Delivering performance monitoring counter support to the OS scheduler”. *The computer journal* Volume 60, Issue 1 January 2017.

[26] Victoria Sanz, Adrián Pousa, Marcelo Naiouf, Armando De Giusti. “Performance Analysis of AES on CPU-GPU Heterogeneous Systems”. En: *Cloud Computing, Big Data & Emerging Topics, 10th Conference, JCC-BD&ET 2022*. ISBN 978-3-031-14599-5, Springer International Publishing, pág. 31-42, agosto de 2022.

[27] Sean Settle: “High-performance Dynamic Programming on FPGAs with OpenCL”. In: *IEEE High Performance Extreme Computing Conference*. 2013.

[28] F. G. Tinetti, M. J. Perez, A. Fraidenraich, and A. E. Altenberg, “Legacy code and parallel computing: updating and parallelizing a numerical model”. *The journal of supercomputing* (ISSN 1573-0484), doi. 10.1007/s11227-020-03172-7, 2020.

[29] Google Inc. “Cloud Tensor Processing Unit (TPU)” Disponible en <https://cloud.google.com/tpu/docs/tpus?hl=es-419>

[30] O. C. Valderrama Riveros and F. G. Tinetti, “MPI Communication Performance in a Heterogeneous Environment with Raspberry Pi”, *Advances in Parallel & Distributed Processing, and Applications. Transactions on Computational Science and Computational Intelligence.*, ISBN: 978-3-030-69984-0, págs. 451-460, doi. 10.1007/978-3-030-69984-0_33, 2021.

[31] Xilinx Inc. “SDAccel Development Environment”. [Online]. Disponible en <http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html>

[32] Maccallini, L., Encinas, D. O., & Romero, F. (2021). An Approach to the Modeling and Simulation of Intra-Hospital Diseases. *Journal of Computer Science and Technology*, 21(2), e14. <https://doi.org/10.24215/16666038.21.e14>

[33] Odroid <http://www.hardkernel.com> Accedido 21 de marzo de 2016.

[34] Suarez, Sergio Leonel. “Análisis de patrones de resiliencia en una arquitectura basada en microservicios”. Tesis de Licenciatura en Sistemas. Universidad Nacional de La Plata. 2022. <http://sedici.unlp.edu.ar/handle/10915/149187>.

Arquitecturas Edge-Fog-Cloud en Procesamiento Distribuido. Aspectos de Eficiencia y Resiliencia

Armando De Giusti ⁽¹⁾⁽²⁾ , Marcelo Naiouf ⁽¹⁾ , Santiago Medina, Diego Montezanti ⁽¹⁾ , Laura De Giusti ⁽¹⁾⁽³⁾ ,
Fernando G. Tinetti ⁽¹⁾⁽³⁾ , Franco Chichizola ⁽¹⁾ , Enzo Rucci ⁽¹⁾⁽³⁾ , Adrián Pousa ⁽¹⁾ , Victoria Sanz ⁽¹⁾⁽³⁾ ,
Diego Encinas ⁽¹⁾ , Ismael Rodríguez ⁽¹⁾ , Sebastián Rodríguez Eguren ⁽¹⁾ , Leandro Libutti ⁽¹⁾ , Manuel
Costanzo ⁽¹⁾, Lucas Gómez D'Orazio ⁽¹⁾, Francisco Garay ⁽¹⁾

¹Instituto de Investigación en Informática LIDI (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata – Comisión de Investigaciones Científicas de la
Provincia de Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

{degiusti, mnaiouf, fernando, smedina, dmontezanti, ldgiusti, fernando, francoch, erucci, apousa, vsanz, dencinas,
ismael, seguren, llibutti, mcostanzo, lgomez, fgaray}@lidi.info.unlp.edu.ar

Resumen

El eje de esta línea de I/D lo constituye el estudio de la integración de arquitecturas distribuidas, que van desde el nivel de los nodos sensores basados en microcontroladores (Edge Computing), pasando por una capa intermedia de preprocesamiento (Fog Computing), hasta la capa final de procesamiento en la nube (Cloud Computing). Los temas centrales de investigación son:

- La distribución equilibrada del procesamiento y del almacenamiento de datos en cada nivel.
- El análisis de la integridad y la performance de las comunicaciones, según el grado de distribución del procesamiento.
- La migración de un cierto nivel de inteligencia a los nodos finales, para reducir el consumo energético y los tiempos de comunicación.
- La integración y evaluación de servicios en las capas de Edge y Fog Computing.
- La incorporación y evaluación de estrategias de resiliencia y tolerancia a fallos a la arquitectura distribuida.
- La integración y evaluación de diversas plataformas IoT y servicios específicos.
- El desarrollo y evaluación del desempeño de aplicaciones que utilizan distintos niveles de procesamiento.
- La caracterización de los tiempos de cómputo y comunicaciones, así como de la

eficiencia energética, para aplicaciones en arquitecturas distribuidas

Palabras clave: *Sistemas Distribuidos. Cloud Computing. Fog Computing. Edge Computing. IoT, Procesamiento Distribuido. Eficiencia. Resiliencia*

Contexto

La línea de investigación que se presenta es parte del proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de Rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real.” del III-LIDI y de proyectos específicos apoyados por organismos nacionales e internacionales. También es parte del proyecto “Procesamiento Eficiente de Grandes Datos mediante Cómputo de Altas Prestaciones, Fog y Edge” financiado por la Facultad de Informática de la UNLP y del proyecto “Unidad Inteligente para Control de Consumo Energético” financiado por la Secretaría de Políticas Universitarias y la UNLP.

En los temas relacionados existen cooperaciones con varias Universidades de Argentina, y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por ERASMUS, CyTED y la OEI (Organización de Estados Iberoamericanos). En particular, se colabora con el proyecto “Computación de Altas Prestaciones Eficiente y Segura para

Aplicaciones de Servicios de Salud Inteligentes”, de la Universidad Autónoma de Barcelona.

Por otra parte, se cuenta con financiamiento de diferentes empresas de Argentina, en particular para la formación de recursos humanos en las temáticas de Cloud, Fog y Edge Computing.

Se participa en iniciativas como el Consorcio de I+D+I en Cloud Computing, Big Data y Emergent Topics, conformado por Universidades de Argentina y España.

Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del Ministerio de Educación, Cultura, Ciencia y Tecnología de la Nación.

Introducción

El constante aumento en la cantidad de datos a procesar que generan los múltiples sensores “inteligentes” en la actualidad, y la necesidad de acotar los tiempos de respuesta trae como consecuencia la necesidad de plantear nuevas arquitecturas que sean capaces de gestionar esa información.

La integración de capas de procesamiento y servicios, en lo que se denomina Edge y Fog Computing, ha dado lugar a un nuevo modelo de arquitectura denominado “Edge-Fog-Cloud Computing” [1] que busca generar una serie de ventajas:

- Reducir el tráfico de comunicaciones y los tiempos de respuesta, mediante la realización de cierto grado de procesamiento en los mismos nodos sensores o en una capa intermedia, y evitando así algunas comunicaciones con el Cloud.
- Obtener un cierto grado de resiliencia frente a fallos de comunicación o caídas de conexión.
- Lograr una cierta capacidad de adaptación de las instancias de procesamiento en función del contexto, incluyendo mayor procesamiento en los nodos sensores.

Este modelo de arquitectura trae nuevos desafíos [2], [3], [4], [5], como son:

- La administración de varios niveles de procesamiento y almacenamiento de datos heterogéneos, manteniendo sin embargo la integridad y seguridad del sistema.
- La definición de protocolos y mecanismos de interoperabilidad entre las diferentes capas de la arquitectura y los servicios en el Cloud.
- El desarrollo de estrategias para la tolerancia a fallos.
- El análisis de la distribución óptima de tareas en cada nivel, en función de mejorar los tiempos de respuesta.
- El análisis de estrategias para disminuir el consumo energético.
- La integración y la evaluación de diferentes plataformas, aplicaciones y servicios.
- El análisis de la escalabilidad con el aumento del número de sensores en la capa “Edge” (no lineal).

Cloud Computing

El cómputo en la nube (Cloud Computing), proporciona facilidad en el acceso y la utilización de grandes conjuntos de recursos físicos y lógicos (como pueden ser infraestructura, plataformas de desarrollo, almacenamiento y/o aplicaciones), por medio de una interfaz de administración web, con un modelo de arquitectura “virtualizada” [6] [7]. Estos recursos son proporcionados como servicios (“*as a service*”) y pueden ser dinámicamente reconfigurados para adaptarse a una carga de trabajo variable (escalabilidad), logrando una mejor utilización y evitando el sobre o sub dimensionamiento (elasticidad) [8].

En un modelo de procesamiento distribuido que integra desde los sensores hasta el cómputo en la nube, este último nivel se reserva para el procesamiento centralizado de algoritmos complejos, con gran volumen de datos. Parte de estos datos pueden haber sido pre-procesados en las capas inferiores de Edge o Fog [9] [10].

Fog Computing

El modelo de Fog Computing surge como respuesta al crecimiento de los desarrollos relacionados con Internet de las Cosas (IoT). Estos desarrollos requieren procesamiento en la nube, pero presentan características que hacen compleja su implementación mediante el uso exclusivo de tecnologías de Cloud Computing [11] [12]. Por lo tanto, Fog Computing consiste en una plataforma intermedia que provee capacidades de procesamiento, almacenamiento y servicios de comunicación en red entre los dispositivos Edge (que adquieren los datos en el modelo de IoT) y el Cloud. Las características de Fog Computing están orientadas al manejo de [13] [14] [15]:

- Aplicaciones distribuidas de tiempo real (y por lo tanto, bajas latencias).
- Gran número de nodos, que puede escalar dinámicamente.
- Heterogeneidad de los nodos.
- Movilidad de los nodos y conocimiento de su ubicación.
- Comunicaciones predominantemente inalámbricas.

Edge Computing

El modelo de Edge Computing está impuesto por el crecimiento exponencial del número de dispositivos sensores con inteligencia local disponibles. La Internet de las Cosas (IoT) está en continuo crecimiento: en la actualidad existen del orden de 50.000 millones de dispositivos conectados a Internet, con un tráfico del orden de 800 Zbytes. Por ello, se requiere capacidad de procesamiento cerca de los sensores, integración de datos locales y geográficamente distribuidos y posibilidad de preprocesar los mismos para enviarlos con una menor sobrecarga de comunicaciones a los niveles superiores (Fog o Cloud) [16] [17] [18] [19]. Las ventajas del modelo Edge son relativamente evidentes [20] [21]:

- Posibilidad de dar respuestas automáticas al usuario, en tiempo real, utilizando la capacidad local.
- Disminución del tráfico de datos.
- Mayor seguridad obtenida a través del procesamiento en servidores locales y privados.
- Posibilidad de disminuir el consumo energético, debido a las características de los componentes.
- Mejora de la eficiencia global para un sistema distribuido, débilmente acoplado.

Aplicaciones de tiempo real.

Este nuevo modelo Edge-Fog-Cloud Computing es especialmente aplicable a problemas de tiempo real (Cloud Robotics, vehículos autónomos, monitoreos de salud personalizados, etc.), al permitir que las capas cercanas a los sensores y al usuario resuelvan en menor tiempo (y con menor sobrecarga de comunicaciones), pudiendo dar una respuesta casi “inmediata” y dejando para el Cloud el procesamiento de los datos masivos *offline* (por ejemplo para ajustar/perfeccionar un modelo de comportamiento que luego se transforma en mejoras a los algoritmos en las capas Edge y Fog [22] [23]).

Eficiencia energética

La mejora de la eficiencia energética es un tema central en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores. Muchos esfuerzos están orientados a tratar la eficiencia energética y a las metodologías para medirla como ejes de I/D, como una métrica de evaluación relevante. En el caso del modelo Edge-Fog-Cloud Computing, hay una mejora del consumo por dos efectos:

- Normalmente los procesadores que se emplean en la capa Edge son de bajo consumo.
- La reducción del tráfico de comunicaciones produce la disminución del consumo energético asociado al mismo.

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas aplicables en Edge Computing y en Fog Computing. Modelos de referencia.
- Administración e integración de recursos y datos en Edge y Fog Computing.
- Seguridad e integridad en los datos en Edge y Fog Computing.
- Vinculación de las capas Edge y Fog con el Cloud. Servicios requeridos.
- Vinculación de las capas Edge y Fog con plataformas dedicadas a IoT.
- Estrategias de distribución óptima de procesamiento y datos entre capas.
- Escalabilidad de las aplicaciones sobre Edge-Fog-Cloud.
- Migración de inteligencia al nivel Edge para reducir consumo y tiempos de comunicación.
- Métricas de eficiencia considerando tiempos de respuesta / costo de comunicaciones / consumo energético.
- Estrategias para incorporar resiliencia a las plataformas basadas en Edge, Fog y Cloud [24].
- Desarrollo de diferentes nodos finales: nodos sensores, robots, drones, etc.
- Algoritmos colaborativos en tiempo real que integren las capas de Edge, Fog y Cloud.
- Sistemas distribuidos inteligentes para controlar y reducir el consumo energético.
- Aplicaciones de las tecnologías de Cloud, Fog y Edge Computing a los procesos de Transformación Digital en organizaciones e industria.

Objetivos y Resultados Esperados

Investigación experimental por realizar

- Análisis comparativo de plataformas IoT y servicios específicos para su utilización a nivel de Edge y Fog.
- Planteo de arquitecturas que integren servicios y plataformas en las capas de Edge y Fog.

- Análisis de performance de protocolos de comunicación para la gestión de nodos utilizando plataformas IoT.
- Estudios de consumo energético en aplicaciones que utilizan los niveles de Edge, Fog y Cloud.
- Estudio comparativo para el balanceo de la carga de procesamiento en aplicaciones distribuidas que utilizan la arquitectura de 3 niveles.
- Integración de aplicaciones móviles.
- Análisis de estrategias para el control de fallos en las diferentes capas de la arquitectura.
- Aplicaciones: red de sensores para el monitoreo de CO₂.
- Aplicaciones: sistema de nodos inteligentes para control del consumo energético en edificios.

Resultados obtenidos

- Se han analizado, desplegado y configurado diferentes servicios orientados a la gestión de nodos IoT.
- Se han realizado análisis comparativos de las características de Plataformas IoT desplegadas en diferentes arquitecturas.
- Se han realizado pruebas para el despliegue de sensores de CO₂ comunicados a una Plataforma IoT [25].
- Se han estudiado protocolos de comunicación y volúmenes de tráfico en aplicaciones distribuidas en tiempo real.
- Se ha desplegado un servidor a nivel de Edge (basado en una Raspberry Pi) para la gestión de nodos sensores y prueba de servicios específicos.
- Se han estudiado distintas estrategias de resiliencia aplicadas a una arquitectura distribuida basada en microservicios [26].

Organización de Eventos

En 2022 se han organizado las X Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET 2022) en Argentina, con participación de especialistas académicos del país y el exterior, además de empresas con experiencia en Cloud Computing [27] [28]. En junio de 2023 se tiene prevista la organización de las XI JCC-BD&ET.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D están en curso 3 Tesis de Maestría y 1 Trabajo Final Integrador de Especialización.

Asimismo, se presentó 1 trabajo final de grado y se están desarrollando otros 2 trabajos finales en temáticas relacionadas.

Además, se participa en el dictado de las carreras de Doctorado en Ciencias Informáticas, y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática de la UNLP (acreditadas por CONEAU), por lo que potencialmente pueden generarse Tesis de Doctorado y Maestría, además de Trabajos Finales de Especialización.

Se está trabajando en la definición de una carrera de Maestría en Transformación Digital, en la que se dictarían los cursos relacionados con el empleo de Cloud, Fog y Edge Computing.

Existe cooperación con grupos de otras Universidades del país y del exterior con posibilidad de realizar Tesis en colaboración.

Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación dos asignaturas directamente relacionadas con los temas de ésta: “Cloud Computing y Cloud Robotics” y “Conceptos y Aplicaciones en Big Data”. Asimismo, todos los años se desarrollan proyectos con alumnos, relacionados básicamente con aplicaciones de tiempo real con robots y drones.

Referencias

[1] Mohan N., Kangasharju J. “Edge-Fog Cloud: A Distributed Cloud for Internet of Things Computations”, 2016 Cloudification of the Internet of Things (CIoT), Paris, France, 2016, pp. 1-6.
[2] P. Garcia Lopez et al. “Edge-centric computing: Vision and challenges,” *SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 37–42, Sep. 2015.
[3] M. Yannuzzi et al. “Key ingredients in an iot recipe: Fog computing, cloud computing, and more fog computing,” in *IEEE CAMAD*, 2014.
[4] K. Hong et al. “Mobile fog: A programming model for large-scale applications on the internet of things,” in *ACM SIGCOMM Workshop on Mobile Cloud Computing*, 2013.
[5] A. Chandra, J. Weissman, and B. Heintz, “Decentralized edge clouds,” *Internet Computing, IEEE*, vol. 17, no. 5, pp. 70–73, Sept 2013.

[6] E. Rucci, M. Naiouf, F. Chichizola, and L. De Giusti “Cloud Computing, Big Data & Emerging Topics. 8th Conference, JCC-BD&ET 2020, La Plata, Argentina, September 8-10, 2020, Proceedings”. Springer CCIS, ISBN: 978-3-030-61218-4, 2020.
[7] Xing, Y., Zhan, Y.: “Virtualization and Cloud Computing”. In: Proceedings pp.305-312, Springer Link. ISBN 978-3-642-27323-0. (2012). Morgan Kaufmann. 2013.
[8] Velte, A.T., Velte, T.J., Elsenpeter, R.: “Cloud Computing: A Practical Approach”. McGraw Hill Professional. 2009.
[9] Ashkan Yousefpour, Caleb Fung, Tam Nguyen, Krishna Kadiyala, Fatemeh Jalali, Amirreza Niakanlahiji, Jian Kong, Jason P. Jue. “All one needs to know about fog computing and related edge computing paradigms: A complete survey”, *Journal of Systems Architecture*, Volume 98, 2019, Pages 289-330, ISSN 1383-7621,
[10] Pi Puig M. et al. (2019) Intelligent Distributed System for Energy Efficient Control. In: Naiouf M., Chichizola F., Rucci E. (eds) *Cloud Computing and Big Data. JCC&BD 2019. Communications in Computer and Information Science*, vol 1050. Springer, Cham
An Overview on Edge Computing Research
[11] Keyan Cao, Yefan Liu, Gongjie Meng, Quimeng Sun “An Overview on Edge Computing Research” in *IEEE Access*, vol. 8, pp. 85714-85728, 2020.
[12] W. S. Shi, X. Z. Zhang, and Y. F. Wang, “Edge computing: State-of-the-art and future directions,” *J. Comput. Res. Develop.*, vol. 56, no. 1, pp. 1_21, 2019.
[13] X. Hong and Y. Wang, “Edge computing technology: Development and countermeasures,” *Chin. J. Eng. Sci.*, vol. 20, no. 2, p. 20, 2018.
[14] D. Evans. “The Internet of Things How the Next Evolution of the Internet is Changing Everything”. Available: <https://www.researchgate.net/publication/30612290>
[15] M. Satyanarayanan, “The emergence of edge computing,” *Computer*, vol. 50, no. 1, pp. 30_39, Jan. 2017.
[16] Y. Q. Gao, H. Bguan, and Z. W. Qi, “Service level agreement-based energy-efficient resource management in cloud data centers,” *Comput. Elect. Eng.*, vol. 40, no. 5, pp. 1621_1633, 2014.
[17] W. Shi, H. Sun, J. Cao, Q. Zhang, and W. Liu, “Edge computing-an emerging computing model for the Internet of everything era,” *J. Comput. Res. Develop.*, vol. 54, no. 5, pp. 907_924, May 2017.
[18] J. de Antueno, S. Medina, L. De Giusti and A. De Giusti, “Analysis, Deployment and Integration of Platforms for Fog Computing”, *Journal of Computer Science and Technology*, 20(2), e12, October 2020.
[19] F. Bonomi, R. Milito, J. Zhu and S. Addepalli, “Fog Computing and Its Role in the Internet of Things”, in *MCC '12: Proceedings of the first edition of the MCC workshop on Mobile cloud computing. Association for Computing Machinery, New York, NY, United States*, 2012.

- [20] M. Asemani, F. Jabbari, F. Abdollahei and P. Bellavista, "A Comprehensive Fog-enabled Architecture for IoT Platforms", *High-Performance Computing and Big Data Analysis. TopHPC 2019. Communications in Computer and Information Science, vol 891. Springer, Cham*, 2019.
- [21] W. S. Aung and S. Aung Nyein Oo, "Monitoring and Controlling Device for Smart Greenhouse by using Thinger.io IoT Server". *International Journal of Trend in Scientific Research and Development*, 2019.
- [22] Dr. S. K. Selvaperumal, W. Al-Gumaei, R. Abdulla and V. Thiruchelvam, "Integrated Wireless Monitoring System Using LoRa and Node-Red for University Building". *Journal of Computational and Theoretical Nanoscience, Volume 16, Number 8, American Scientific Publishers*, 2019.
- [23] M. Pi Puig, J. M. Paniego, S. Medina, S. Rodriguez Eguren, L. Libutti, J. Lanciotti, J. de Antueno, C. Estrebou, F. Chichizola and L. De Giusti, "Intelligent Distributed System for Energy Efficient Control", *Cloud Computing and Big Data. JCC&BD 2019. Communications in Computer and Information Science, vol 1050. Springer, Cham*, 2019.
- [24] Medina, S., Montezanti, D. M., Gomez D'Orazio, L., Compagnucci, E., De Giusti, A. E., & Naiouf, M. (2022). Incorporating Resilience to Platforms based on Edge and Fog Computing. In X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022).
- [25] Gomez D'Orazio, L., Medina, S., Montezanti, D. "Integración de una red de sensores con una plataforma IoT para control inteligente de aulas". Proceedings del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022) – EUDELAR, pp. 1002 – 1007. ISBN 978-987-1364-31-2 (2022). <http://sedici.unlp.edu.ar/handle/10915/149102>.
- [26] Suarez, Sergio Leonel. "Análisis de patrones de resiliencia en una arquitectura basada en microservicios". Tesina de Licenciatura en Sistemas. Universidad Nacional de La Plata. 2022. <http://sedici.unlp.edu.ar/handle/10915/149187>
- [27] Cloud Computing, Big Data & Emerging Topics: 9th Conference, JCC-BD&ET, La Plata, Argentina, June 22-25, 2021, Proceedings. Editores: M. Naiouf, E. Rucci, F. Chichizola, and A. De Giusti, Springer International Publishing, ISBN: 978-3-030-84824-8, doi. 10.1007/978-3-030-84825-5, 2021.
- [28] A. E. De Giusti, M. Naiouf, L. C. De Giusti, E. Rucci, and F. Chichizola, "Short papers of the 9th Conference on Cloud Computing, Big Data & Emerging Topics" Facultad de Informática (UNLP), ISBN: 978-950-34-2016-4, 2021

Soporte Serverless para Aplicaciones Móviles de Nueva Generación

Nelson Rodríguez¹, María Murazzo¹, Marcelo Moreno¹, Susana Chávez¹, Adriana E. Martín¹, Sergio Flores¹, Sara Zogbe¹, Adriana C. Martín¹, Leonardo Celador², Ciro Delgado²

¹Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

² Alumno Avanzado Licenciatura en Sistemas de Información y Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

nelson@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar, mpmoren@gmail.com,
schavez@iinfo.unsj.edu.ar, arianamartinsj@gmail.com, sergior@gmail.com,
sarazogbe@yahoo.com.ar, adrianamartin1@gmail.com, leonardomiguelcelador@gmail.com,
cirodelgado7@gmail.com

Resumen

Desde que realizó la primera comunicación celular móvil, el avance de las tecnologías asociadas ha sido continuo, en consecuencia se incrementó la cantidad de dispositivos y la mayoría de la población tiene acceso a los recursos por medio de estas aplicaciones en todo momento y desde cualquier lugar. Actualmente datos y una variedad de servicios para estas aplicaciones se ejecutan en el Back-end fundamentalmente en el Cloud. Como una mejora de este último y adecuado para aplicaciones interactivas o que no requieran atención permanente surge Serverless Computing. Esta arquitectura está basada en funciones, es conducida por eventos y es adecuada para integrarse con la computación móvil. Ante la nueva generación de aplicaciones móviles que incluirán estrategias de seguridad avanzadas, tecnologías inmersivas, super aplicaciones e inteligencia artificial, surgirán necesidades de infraestructura y soporte que logren una ejecución efectiva y eficiente, en concordancia con la sofisticación de estas aplicaciones. La presente propuesta de investigación tiene como objetivo relacionar ambas tecnologías, permitiendo la interoperabilidad e integración de la computación móvil con Serverless y

además analizar y solucionar las limitaciones existentes o que puedan surgir, de forma tal de adecuar el soporte Serverless para las aplicaciones móviles de nueva generación

Palabras clave: *Serverless Computing, Mobile Computing, Apps, Cloud Computing.*

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es la línea marco del proyecto de investigación Soporte Serverless para aplicaciones móviles de nueva generación, cuya propuesta está en etapa de evaluación para el período 2023-2024. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil, distribuida y de alta performance, desde hace más de 22 años. Como continuación del proyecto anterior: Computación Serverless para tratamiento de datos provenientes de dispositivos de IoT, se continúa el trabajo con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

La computación móvil ha avanzado notablemente en los últimos años, surgiendo a nivel de desarrollo millones de apps de la más variada utilidad, área de aplicación y costos. Muchas empresas y organizaciones de

gobierno ofrecen soluciones para sus usuarios que son potenciadas en nuevas versiones.

La pandemia aceleró el uso de los teléfonos inteligentes o smartphones, que pasaron a ser el aula de clase virtual, la biblioteca de PDFs, el cuaderno donde se hacen las actividades, el dispositivo para sacar turno para algún trámite, donde reproducir videos de estudio o complemento de recreación y muchos otros usos.

Si las aplicaciones móviles son robustas o sofisticadas, seguramente requieren que todos los recursos que se ejecutan en el dispositivo móvil deban interactuar con datos y otros recursos que están en servidores, centros de datos o simplemente en la nube o Cloud Computing.

Cloud Computing es un modelo de negocios y una arquitectura ampliamente utilizada, con varios años en el mercado. Dicho modelo ha evolucionado desde una arquitectura monolítica en el servidor, a una arquitectura de microservicios basada en contenedores y recientemente a Serverless Computing, lo que produce aplicaciones más versátiles y que permiten al desarrollador ocuparse solamente de la solución del problema y no de la infraestructura de hardware y software subyacente.

Esta conjunción de tecnologías, puede dar soluciones interesante a las nuevas aplicaciones emergentes, las cuales integrarán aplicaciones como Super Apps, pero todas en mayor o menor medida, realizarán requerimientos al Cloud, tanto las basadas en Inteligencia artificial, como las que necesiten requisitos estrictos de seguridad (como GDPR o similar), o realidad aumentada, virtual, metaverso o lo que pueda surgir en un futuro. La computación Serverless es la tecnología de soporte ideal para que las apps puedan interactuar con los recursos de la nube.

Según varios sitios de empresas consultoras o expertos, indican que ambas tecnologías resultan de sumo interés para los próximos años. Por ejemplo, el MTI College afirma que entre las tendencias en Cloud Computing para 2023, Serverless Computing ocupa un lugar destacado [1]. Además el sitio Digilab, sugiere que las Super Apps móviles se encuentran entre las tendencias tecnológicas más importantes para 2023 [2].

Computación Móvil

El desarrollo de aplicaciones móviles es un campo en constante evolución, Con el rápido ritmo de los avances tecnológicos y la creciente demanda de aplicaciones móviles, los desarrolladores deben adaptarse a estas tendencias para satisfacer las necesidades cada vez mayores de los usuarios.

Dichas tendencias de apps móviles para los próximos años, están bien definidas, aunque en tecnología, siempre pueden surgir nuevas opciones. Los dispositivos móviles han transformado los modelos operativos y los mercados. Por este motivo, las empresas que quieran utilizar el enorme potencial de las aplicaciones para captar y retener clientes, deben estar al día de las tendencias de este mercado para los años venideros, que evoluciona de forma vertiginosa. Dentro de este mercado, ya se están viendo algunas tendencias que vienen con fuerza y que previsiblemente, se quedarán durante mucho tiempo [3].

Entre las principales tendencias para el mercado de las aplicaciones móviles, se puede mencionar: 5G/6G, Super Apps, Aplicaciones móviles basadas en IA, IoT Mobile Apps, Mobile Cloud Computing, AR/RV Mobile Apps, desarrollo Low-Code, tecnología de Blockchain aplicada y otras [4].

Serverless

Siguiendo la evolución observada en la historia de la contenerización, los servicios en la nube se han adaptado para ofrecer contenedores de mejor ajuste que requieren menos tiempo para cargar (arranque) y proporcionar mayor automatización en el manejo (orquestación) de contenedores en nombre del cliente, [5], surgiendo de esta forma la Computación Serverless.

El modelo de computación de la arquitectura serverless es impulsado por eventos en el que los recursos informáticos se proporcionan como servicios escalables. En el modelo tradicional se cobra un costo fijo y recurrente por los recursos informáticos del servidor, independientemente de la cantidad de trabajo. Sin embargo, la implementación de la computación Serverless ha superado esta deficiencia, ya que permite a los clientes pagar solo por el uso del servicio y no se cobran costos ocultos asociados con el tiempo de inactividad.

Un modelo basado en funciones como Serverless, es adecuado para ráfagas, uso de CPU intensivo, cargas de trabajo granulares. Actualmente, los casos de uso de FaaS varían ampliamente, incluido el procesamiento de datos, el procesamiento de flujo, la computación de borde (IoT) y la computación científica [5].

Existen diversos desafíos, oportunidades y problemas a resolver, entre ellos la experiencia del desarrollador [6], Interoperabilidad, testing, composición de funciones, seguridad, administración del ciclo de vida, administración de requerimientos no funcionales, performance, optimización del overhead, ingeniería para costo-performance, entre otros [7].

Tanto la computación móvil, como la computación Serverless, forman parte de un

todo que tiene en uno de sus extremos a lo que se conoce como Edge que puede incluir en este caso a la Computación móvil, pero también puede incluir a IoT, la computación convencional o algún otro tipo de tecnología que permita la adquisición de datos. Y del otro lado a Cloud Computing y en particular para el objeto de estudio de la presente propuesta, Serverless.

Computación Móvil y Serverless

Se define MCC (Mobile Cloud Computing) a la integración de la computación en la nube con dispositivos móviles para mejorar las capacidades de los dispositivos móviles, como la capacidad de cómputo y el almacenamiento [8]. Por otro lado, A finales de 2014, Mobile Edge Computing (MEC) se introdujo como un medio para ayudar a resolver el problema de latencia que ocurre durante el proceso de offloading en MCC [9]

Teniendo en cuenta la parte del Cloud, la computación serverless es adecuada para aplicaciones de corta duración sin estado dirigidas por eventos, por ejemplo microservicios, backends IoT móviles, procesamiento de flujo modesto, bots e integración de servicios.

Desde una perspectiva funcional, serverless y las arquitecturas más tradicionales pueden usarse indistintamente. La determinación de cuándo usar serverless probablemente estará influenciada por otros requisitos no funcionales como la cantidad de control sobre las operaciones requeridas y el costo y las características de la carga de trabajo de la aplicación, por ejemplo.

Además, hay una mayor responsabilidad del operador que brinda más información y control. Los operadores seleccionan los recursos; despliegan y aprovisionan los

mismos; implementan y controlan el monitoreo del uso de recursos, la intensidad de la carga de trabajo y el comportamiento de la aplicación; y puede escalar automáticamente o migrar la aplicación, todo esto, sin mucho o nada de esfuerzo por parte del desarrollador.

Desde la perspectiva del modelo de programación, la naturaleza de operar “sin estado” de las funciones serverless, se presta a la estructura de aplicaciones similares a los que se encuentran en la programación reactiva funcional y en muchas aplicaciones móviles, aunque puede resultar también un inconveniente, que deberá ser solucionado por código.

Actualmente hay esfuerzos de investigación para que la red móvil sea conducida por el software. Entre algunas de las alternativas, las comunicaciones móviles 5G están trabajando hacia la introducción de una arquitectura totalmente basada en software como las soluciones Software Defined Networking (SDN) y Network Function Virtualization (NFV) [10]. En consecuencia, esta operación flexible y bajo demanda de la red se parece mucho a la operación actual de las plataformas Cloud, y también debe seguir principios de computación distribuida, de ahí el nombre de Arquitecturas móviles Serverless.

El paradigma Serverless apunta al aprovisionamiento de servicios más eficiente, ajustando con precisión los recursos desplegados en cualquier momento a la demanda real [11]. Esto claramente es adecuado para aplicaciones móviles, que no tienen un comportamiento estable en cuanto al momento en que se realizan los requerimientos, y por supuesto hay tareas que desarrollar al respecto, dado que para beneficiarse de este paradigma, es esencial estimar con precisión la demanda requerida

por un servicio y pronosticar el consumo de recursos previsto.

De alguna manera, Serverless es una continuación de estrategias para lograr un mayor nivel de abstracción, lo que permite a los desarrolladores centrarse en aplicaciones con menos preocupación por la infraestructura subyacente, que sigue los pasos de la virtualización y la contenedorización de servidores. Es un cambio significativo en la construcción y el funcionamiento de las aplicaciones. A diferencia de la virtualización o los contenedores, Serverless no es una herramienta, sino un modelo de ejecución.

Las aplicaciones móviles de nueva generación, se integrarán, debido a sus características, más adecuadamente con una arquitectura Serverless en el Cloud. Esta presenta mayor sencillez en el desarrollo de aplicaciones, algo muy solicitado por los desarrolladores de aplicaciones móviles. Por otra parte, hay que sumarle la escalabilidad como una ventaja muy importante, la facilidad de Serverless para adecuarse a productos ágiles y la posibilidad de que el código puede ejecutarse desde cualquier lugar, y en cualquier momento.

Se prevé que el mercado global de Mobile Edge Computing (MEC) aumente a un ritmo considerable durante el período de pronóstico, entre 2023 y 2029. En 2021, el mercado está creciendo a un ritmo constante y con la creciente adopción de estrategias por parte de los actores clave, se espera que el mercado aumente en el horizonte proyectado. [12]

Objetivos

Los objetivos del grupo de investigación en esta línea de conocimiento son los siguientes:

- Desarrollar estrategias a nivel de apps móviles para solucionar

problemas de latencia en la comunicación al Cloud.

- Analizar las necesidades de BasS para construir funciones serverless aplicadas a la Computación Móvil.
- Realizar pruebas de comportamiento de las aplicaciones móviles con acceso a Serverless.
- Proponer y evaluar nuevas estrategias para lograr eficiencia en el Edge móvil (almacenamiento, gestión de servicios, funcionalidad de red y cómputo en el extremo).

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de ocho investigadores que figuran en este trabajo de las universidades Nacional de San Juan y dos alumnos de grado. Además, el proyecto marco donde se está desarrollando esta propuesta ha establecido vínculos con investigadores de la Nacional de San Luis, de la Universidad Champagnat y de la Universidad Nacional de Salta y dos alumnos de grado más.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis de maestría en áreas afines y dos tesinas de grado en el área de Serverless computing, Concurrencia y Computación distribuida. Además se espera aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.

Referencias

[1] Jonas E, Schleier-Smith J, Sreekanti V, Tsai C-C, Khandelwal A, Pu Q, Shankar V, Carreira J, Krauth K, Yadwadkar N, Gonzalez JE, Popa RA, Stoica I, Patterson DA (2019) Cloud Programming Simplified: A Berkeley View on Serverless Computing. <http://arxiv.org/abs/1902.03383>. Accessed 6 Jan 2021

[2] S. Eismann et al., ‘Serverless Applications: Why, When, and How?’ (2021), IEEE Software, vol. 38, no. 1, pp. 32–39, Jan. 2021, doi: 10.1109/MS.2020.3023302.

[3] Fromm, K. (2012). <https://readwrite.com/2012/10/15/why-the-future-of-software-and-apps-is-serverless/>

[4] Mobile Computing: 12 Trends to Watch in 2023. In:

<https://www.techopedia.com/2/31446/trends/the-top-10-trends-in-mobile-computing>.

[5] S. Ghemawat, H. Gobioff and S. T. Leung, The Google file system, SOSP, 2003.

[6] Mohanty, S. K., Premsankar, G., & Di Francesco, M. (2018). An Evaluation of Open Source Serverless Computing Frameworks. In CloudCom (pp. 115-120).

[7] Gottlieb, N. (2016). State of the Serverless Community Survey Results.

<https://serverless.com/blog/state-of-serverless-community/>.

[8] Yan Zhang. Mobile Edge Computing. In: Simula SpringerBriefs on Computing. Open Access publication (2022), pp. 1-9. <https://doi.org/10.1007/978-3-030-83944-4>.

[9] M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, and A. Neal, “Mobile-edge computing introductory technical white paper,” White Paper, Mobile-edge Computing (MEC) industry initiative, vol. 29, 2014

[10] M. Condoluci and T. Mahmoodi: Softwarization and virtualization in 5g mobile networks: Benefits, trends and challenges, Computer Networks, vol. 146, pp. 65–84, Dec. (2018).

[11] Gramaglia, M., et al, “The case for serverless mobile networking”. IFIP Networking Conference 2020, 22-26 June 2020 (Virtual conference). IEEE, 2020, Pp. 779-784.

<https://ieeexplore.ieee.org/document/9142747>

[12] Mobile Edge Computing (MEC) Market 2023 Size, Share, Growth, Analysis, Trends and Forecast to 2028 (Feb. 2023). In: <https://www.marketwatch.com/press-release/mobile-edge-computing-mec-market-2023-size-share-growth-analysis-trends-and-forecast-to-2028-2023-02-27>.

Predicción de fenómenos naturales mediante paralelismo, algoritmos evolutivos y búsqueda por novedad

Caymes-Scutari Paola^{1,2}, Bianchini Germán¹, Strappa Jan^{1,2}

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido,
Departamento de Ingeniería en Sistemas de Información. Facultad Regional
Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

pcaymesscutari@frm.utn.edu.ar, gbianchini@frm.utn.edu.ar, jstrappa@frm.utn.edu.ar

RESUMEN

La predicción de diversos fenómenos, como las avalanchas, los incendios, las inundaciones, etc., constituye un proceso complejo en el cual interviene una gran cantidad de información a considerar y numerosas acciones, operaciones, y procedimientos a realizar sobre dicha información. Ello permite modelar el fenómeno, predecir su comportamiento a un cierto plazo, y tomar las decisiones que se consideren más adecuadas a fin de prevenir, mitigar, o paliar los efectos negativos que pueden provocarse en el ecosistema, la población, la sociedad, y la economía a raíz del fenómeno. Nuestra línea de investigación propone abordar la predicción de fenómenos de propagación a través de la integración de tres elementos: la potencia que ofrece el cómputo paralelo/distribuido, la capacidad de los algoritmos poblacionales para orientar la búsqueda de soluciones, y la oportunidad que ofrecen los algoritmos basados en búsqueda por novedad (NS o novelty search). Dichos elementos serán considerados para una familia de métodos predictivos denominada DDM-MOS (*Data Driven Methods with Multiple Overlapping Solutions*), la cual se trata de

una familia de métodos de ayuda a la decisión. Las características intrínsecas de la búsqueda basada en novedad privilegian la evolución u obtención de nuevos resultados (predicciones) en función de la novedad y no solamente por aptitud, como sucede en los algoritmos evolutivos puros. Dicha característica permitiría prevenir o disminuir la aparición de situaciones de convergencia prematura o estancamiento, en beneficio de la calidad de los resultados obtenidos al realizar la predicción.

Palabras clave: Metaheurísticas, Predicción, Novelty, Incertidumbre.

CONTEXTO

Los DDM-MOS consideran un modelo de simulación que toma como entrada un conjunto de parámetros, los cuales describen características que afectan a la propagación, y en base a la combinación de diferentes técnicas computacionales arrojan una predicción sobre el comportamiento del fenómeno. Esta línea se ha desarrollado a lo largo de los años en el LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido), en la UTN-FRM,

aplicándose al caso particular de la predicción de incendios forestales. Los métodos desarrollados por el grupo tienen como objetivo la reducción de incertidumbre en las simulaciones de esta naturaleza, dado que las limitaciones para proveer parámetros de entrada correctos para el modelo conducen a errores en la precisión de la predicción. En general, los métodos desarrollados involucran un componente o metaheurística evolutiva para reducir el espacio de búsqueda y a la vez orientar la búsqueda. En este contexto, la incorporación de búsqueda por novedad o *novelty search*, tiene como objetivo modificar el funcionamiento clásico de las metaheurísticas, otorgando importancia a la riqueza que puedan brindar al proceso aquellos puntos que resulten más novedosos, y así privilegiar no sólo a los sectores aparentemente más aptos, sino también poniendo en valor a los que resulten más diferentes y ofrezcan posibilidades renovadas de exploración. Actualmente, la línea de investigación cuenta con dos líneas de financiamiento homologadas, a través de un proyecto PID UTN que dará inicio en abril de 2023 (SITCME0009816TC), y un proyecto PICTO de FONCyT (UUMM-2019-00042).

1. INTRODUCCIÓN

Existen diversos simuladores cuyo objetivo es predecir la evolución de la línea de fuego a lo largo de un período de tiempo [1]. Para esto, los simuladores requieren un conjunto de parámetros de entrada, llamados también escenarios, que describen características que afectan a la propagación, tales como el tipo de material combustible, su humedad, pendiente y relieve del terreno, temperatura, y la velocidad y dirección del viento. Desde el punto de vista computacional, este problema de

predicción es desafiante debido a la complejidad de los modelos involucrados y a las fuentes de incertidumbre existentes en los datos de entrada. Este último aspecto es de gran importancia dado que las limitaciones para proveer parámetros de entrada correctos para el modelo conducen a errores en la predicción. Esta incertidumbre se debe a la dificultad o imposibilidad de obtener los valores de las variables, tanto por limitaciones de recursos (por ejemplo, cantidad de sensores), porque las mediciones son indirectas (por ejemplo, la medición de la humedad de la vegetación), o porque las variables tienen un comportamiento dinámico y no es viable su observación en tiempo real (por ejemplo, características del viento). Esto hace que la predicción basada en una solución única no sea confiable [1].

Para enfrentar esta incertidumbre, se han propuesto métodos que combinan resultados de múltiples simulaciones para mitigar los efectos negativos de la misma. Este tipo de métodos son los denominados DDM-MOS [1], dos de los cuales han sido desarrollados en los últimos años: ESSIM-EA [5,6] y ESSIM-DE [10]. Su funcionamiento involucra la combinación de una Etapa de Optimización, para obtener escenarios cuya simulación permita obtener buenas predicciones, con una Etapa de Análisis Estadístico, que construye una matriz de probabilidades a partir de la agregación de mapas obtenidos de las múltiples simulaciones paralelas, para así poder predecir el comportamiento de la línea de fuego. En el caso de ESSIM-EA, el método de optimización empleado es un algoritmo genético, mientras que en ESSIM-DE se utiliza un algoritmo de evolución diferencial [8]. Debido a la carga de cómputo de este proceso, ambos sistemas utilizan cómputo paralelo; concretamente, se basan en un modelo de

islas [9], utilizando una doble jerarquía de procesos que les permite abarcar mejor el espacio de búsqueda y acelerar la optimización. Es importante señalar que, además de emplearse para la prevención y predicción de incendios forestales, estos métodos pueden aplicarse a otros modelos de propagación, tales como inundaciones, avalanchas o corrimientos de terreno.

Dado que los mencionados métodos se basan en metaheurísticas, se pueden ver afectados por algunas limitaciones intrínsecas de los mismos, como son el estancamiento o la convergencia prematura. Para el caso particular de ESSIM-DE, se propusieron diferentes aproximaciones que permitieron sintonizar su comportamiento y lograr mejoras tanto en calidad de resultados como en tiempo de ejecución.

Más recientemente, se consideró la posibilidad de utilizar un paradigma alternativo al de las metaheurísticas tradicionales: la *búsqueda basada en novedad* o *Novelty Search* (NS) [2-4]. Este es un paradigma de búsqueda que ignora el objetivo como guía para la exploración y, en su lugar, recompensa a soluciones candidatas que presentan comportamientos novedosos (diferentes a los anteriormente descubiertos), con el fin de maximizar la exploración y evitar óptimos locales. La búsqueda basada en novedad es una alternativa prometedora frente a las limitaciones de las metaheurísticas aplicadas previamente al problema de la predicción, debido a que, en este paradigma alternativo, la búsqueda es dirigida por una caracterización del comportamiento de los individuos, cambiando así el *paisaje de la función de fitness* (*fitness landscape*), previniendo de manera directa los problemas de convergencia prematura y estancamiento.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los sistemas de predicción existentes han obtenido resultados satisfactorios para la reducción de incertidumbre en el problema de aplicación. En particular, ESSIM-EA demostró obtener predicciones de buena calidad, mientras que ESSIM-DE redujo significativamente los tiempos de respuesta, pero sin obtener mejoras en calidad. Por esta razón, posteriormente se ha trabajado en mejorar el rendimiento del método ESSIM-DE mediante estrategias denominadas de sintonización [7]. Se desarrollaron dos de estas métricas [11, 12], ambas orientadas a mitigar los problemas de convergencia prematura y estancamiento de la población presentes en el caso de aplicación del algoritmo, obteniendo mejoras en calidad y tiempos de respuesta.

Dadas las características del problema de la predicción y los resultados previos obtenidos mediante los DDM-MOS ya desarrollados, recientemente se ha llevado a cabo una primera aproximación para incorporar NS al proceso de reducción de incertidumbre y predicción guiado por Evolución Diferencial. Los resultados obtenidos son prometedores y señalan un camino a seguir para continuar explorando en las posibilidades que puede proveer, en este caso, a los DDM-MOS. Asimismo, los resultados han permitido confirmar que el enfoque NS es adecuado para este tipo de problema, dado que, por un lado, resuelve por diseño los problemas propios de algoritmos explotativos; y por otro lado, porque presenta múltiples oportunidades de paralelización y de hibridación con enfoques basados en *fitness*. También es posible adaptar el comportamiento del

algoritmo de acuerdo a las características del problema.

Otra de las líneas de investigación del grupo se centra en la sintonización de aplicaciones paralelas, motivo por el que también es factible plantear a un plazo más largo, el análisis y modelado del comportamiento de aquellos DDM-MOS que incorporen búsqueda por novedad, y así contribuir a lograr predicciones más precisas y/o en menor tiempo.

3. OBJETIVOS TEÓRICOS Y EXPERIMENTALES

El principal objetivo del proyecto consiste en considerar nuevas estrategias para mejorar la calidad de predicción alcanzada por los DDM-MOS, como métodos computacionales para la reducción de incertidumbre en la predicción de fenómenos de propagación. Más allá de la utilidad y beneficios que los algoritmos evolutivos poblacionales aportan a los DDM-MOS para orientar la búsqueda, son algoritmos que intrínsecamente traen aparejados ciertos problemas de rendimiento (como convergencia prematura o estancamiento) que afectan a las características de la población a lo largo del tiempo. Es por ello que se propone potenciar las capacidades de los métodos predictivos basados en algoritmos evolutivos poblacionales mediante la integración de características basadas en novedad que privilegien la evolución según este aspecto y no sólo por aptitud. La incorporación de la búsqueda por novedad permitirá, además, conocer mejor las potencialidades de las metaheurísticas para aplicarlas en la

predicción de diferentes fenómenos de propagación, como se mencionó anteriormente. Como parte de este objetivo general pueden señalarse algunos objetivos más específicos. Por un lado, identificar los beneficios de utilizar búsqueda por novedad, y las posibilidades de adaptarla e integrarla a algoritmos predictivos y de reducción de incertidumbre basados en metaheurísticas y paralelismo. Y, por otro lado, analizar los resultados obtenidos, comparándolos y contrastándolos con los métodos existentes, a fin de considerar las fortalezas a sostener, y las debilidades a tener en cuenta respecto de la utilización de NS. De esta manera, el principal aporte que se espera alcanzar es la integración de NS a los DDM-MOS, bajo la hipótesis de que por sus características, NS permitiría flexibilizar y mejorar la exploración del espacio de búsqueda, y con ello mejorar los resultados de las predicciones ya sea en términos de tiempo de ejecución o de precisión de la predicción. Alcanzar al menos uno de los dos objetivos constituirá un logro importante, dado que aportará una herramienta de ayuda a la decisión más fiable, que podría utilizarse para el manejo de incendios forestales, tanto a nivel local, como nacional o internacional, en caso de catástrofe.

4. FORMACIÓN DE RECURSOS HUMANOS

La temática propuesta por este plan de trabajo permite continuar con la formación de los distintos integrantes del grupo de trabajo, de forma complementaria a la formación adquirida hasta el momento en el campo de los

métodos predictivos, del paralelismo, y de reducción de incertidumbre. En el caso particular de Dr. Strappa, esta temática se relaciona con su plan de trabajo postdoctoral (se incorporó al laboratorio en 2021 con una beca postdoctoral de CONICET, dirigida por la Dra. Paola Caymes Scutari y codirigida por el Dr. Germán Bianchini,), lo cual constituye un aporte a su formación en investigación.

En el marco de la línea general de los DDM-MOS, en 2020 se defendieron dos tesis doctorales relacionadas con esta línea de investigación, desarrolladas por la Dra. Laura Tardivo (bajo la dirección de la Dra. Paola Caymes Scutari) y el Dr. Miguel Méndez Garabetti (bajo la dirección del Dr. Germán Bianchini y la codirección de la Dra. Paola Caymes Scutari).

El grupo también está integrado por estudiantes de grado que están comenzando a introducirse en los temas atinentes a la línea de investigación, y asimismo las puertas del grupo de trabajo siempre están abiertas a la incorporación de nuevos integrantes (de grado o postgrado) que deseen familiarizarse con las temáticas con las que se trabaja dentro del mismo.

5. BIBLIOGRAFÍA

- [1] Bianchini, Germán, Denham, Mónica, Cortés, Ana, Margalef, Tomàs, Luque, Emilio. (2010). "Wildland Fire Growth Prediction Method Based on Multiple Overlapping Solution". *Journal of Computational Science* Vol 1 Issue 4, pp. 229-237. Elsevier.
- [2] Lehman J. and Stanley K. O.. (2011). "Abandoning Objectives: Evolution Through the Search for Novelty Alone," *Evolutionary Computation*, vol. 19, no. 2, pp. 189-223.
- [3] Lehman J. and Stanley K. (2008). "Exploiting Open-Endedness to Solve Problems Through the Search for Novelty," *Artificial Life*, p. 8.
- [4] Lehman J. and Stanley K. (2013). "Evolvability Is Inevitable: Increasing Evolvability without the Pressure to Adapt," *PLoS ONE*, v. 8, 4, pp. 2-10.
- [5] Méndez Garabetti, Miguel, Germán Bianchini, María Laura Tardivo, and Paola Caymes Scutari. (2015). "Comparative Analysis of Performance and Quality of Prediction Between ESS and ESS-IM." *Electronic Notes in Theoretical Computer Science* 314 (June): 45-60.
- [6] Mendez Garabetti, Miguel, Bianchini, Germán, Caymes Scutari, Paola, Tardivo María Laura. (2016). "Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic". *Fire Safety Journal*, pp. 49-62. Elsevier.
- [7] Naono, K., Teranishi, K., Cavazos, J. y Suda, R. (2010). *Software Automatic Tuning: From Concepts to State-of-the-Art Results*, Springer, New York.
- [8] Price K., Storn R., Lampinen J. (2005). *Differential Evolution - A practical approach to global optimization*. Springer-Verlag New York.
- [9] Talbi, E. (2009). *Metaheuristics: From Design to Implementation*. John Wiley & Sons, Hoboken, New Jersey.
- [10] Tardivo, María Laura, Paola Caymes Scutari, Germán Bianchini, and Miguel Méndez Garabetti. (2017). "Hierarchical Parallel Model for Improving Performance on Differential Evolution" *Concurrency and Computation: Practice and Experience* 29 (10): e4087.
- [11] Tardivo M. L., Caymes Scutari P., Méndez Garabetti M., and Bianchini G. (2018). "Optimization for an Uncertainty Reduction Method Applied to Forest Fires Spread Prediction," in *Computer Science 2017*, vol. 790, A. E. De Giusti, Ed. Cham: Springer International Publishing. pp. 13-23.
- [12] Tardivo M. L., Caymes Scutari P., Bianchini G., and Méndez Garabetti M. (2019). "Sintonización Dinámica del Método Paralelo de Predicción de Incendios Forestales ESSIM-DE," *XXV Congreso Argentino de Ciencias de la Computación*, p. 10, 2019.

SIMULACIÓN Y ANÁLISIS EFICIENTE DE GRANDES VOLUMENES DE DATOS

Mercedes Barrionuevo, LIDIC, FCFMyN, UNSL

Alicia Castro, LIDIC, FCFMyN, UNSL

Gil Costa G. Veronica LIDIC, FCFMyN, UNSL

Danilo Labella, FCFMyN, UNSL

Fernando Loor LIDIC, FCFMyN, UNSL

A.Marcela Printista, LIDIC, FCFMyN, UNSL

Cristian Tissera, LIDIC, FCFMyN, UNSL

Guillermo Trabes LIDIC, FCFMyN, UNSL

CONTEXTO

En este trabajo se presentan líneas de investigación que combinan “Simulación y análisis eficiente de grandes volúmenes de datos”, las cuales pertenecen al proyecto PROICO “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” de la UNSL. El proyecto tiene como metas principales investigar nuevas metodologías para el procesamiento eficiente y eficaz de grandes volúmenes de datos y proponer nuevas tecnologías para la obtención de información útil. Por otro lado, pretende avanzar en temas de investigación actuales, incorporando el análisis y uso de técnicas de ciencia de datos. El procesamiento eficiente de grandes volúmenes de datos (Big Data) y la ciencia de datos son dos áreas de la informática que han crecido y llamado la atención de diferentes disciplinas, tanto académicas como de negocios. En este sentido será posible detectar anomalías, descubrir patrones en los conjuntos de datos, encontrar posibles relaciones entre los datos, etc. y por supuesto obtener análisis que sirvan como respaldo para la toma de decisiones.

RESUMEN

El crecimiento en el volumen de datos, así como en la heterogeneidad de los tipos de datos que existen en el mundo ha crecido exponencialmente en la última década. Esto implica dos grandes desafíos para la comunidad científica: (1) cómo procesar esos datos lo más rápido y eficiente posible y (2) cómo obtener información a través de dichos datos. Por otro lado, este gran volumen de datos brinda una oportunidad única en el área de modelado y simulación. Estos datos pueden ser utilizados para generar modelos de diferentes situaciones y aplicaciones que puedan ser validados y posteriormente escalados. En este trabajo, se presentan los objetivos, trabajo realizado y desafíos que aborda el grupo de investigación de la Universidad Nacional de San Luis, para abordar los temas que involucran el procesamiento eficiente de grandes volúmenes de datos, para realizarlo en forma eficiente de forma tal de reducir los tiempos de ejecución, evitar saturar los recursos, balancear la carga de trabajo y reducir consumos de energía, así como para obtener información efectiva de esos datos que puedan ser utilizados para la toma de decisiones utilizando técnicas de ciencias de datos y finalmente utilizar esos datos para generar modelos de simulación que permitan analizar situaciones que no son factibles de ser implementadas en

la realidad por temas de costos, disponibilidad de recursos, etc.

Palabras clave: *Grandes volúmenes de datos, procesamiento eficiente, análisis de datos, simulación, resiliencia.*

1. INTRODUCCIÓN

Durante la última década, Big Data (Grandes Datos) se ha convertido en un fuerte foco de interés mundial, que atrae cada vez más la atención de la academia, la industria, el gobierno y otras organizaciones. El término "Big Data" apareció por primera vez en las comunidades científicas a mediados de la década de 1990, y actualmente se sugiere como fuente predominante de innovación, competencia y productividad. El creciente flujo de datos, que proviene de los diferentes tipos de sensores, sistemas de mensajería y redes sociales, además de sistemas de medición y observación más tradicionales, ya ha invadido muchos aspectos de nuestra existencia cotidiana. Por otro lado, tanto en las ciencias como en la industria, los proyectos a gran escala representados como sistemas complejos se han convertido en una tangible realidad, siendo la naturaleza de los mismos descentralizada, no lineal y generalmente compuestos por una multitud de unidades autónomas, que pueden interactuar, evolucionar, y adaptar su comportamiento a cambios en el entorno; lo que implica que en la mayoría de los casos es muy difícil (sino imposible) conseguir soluciones analíticas resolubles que sean capaces de describir un sistema complejo de manera adecuada. Utilizadas de manera apropiada, las diferentes técnicas de simulación proporcionan un método poderoso e indispensable para analizar una diversidad de problemas, ya que permite una comprensión cualitativa y cuantitativa de muchos fenómenos que son demasiados complejos para ser tratados con métodos analíticos o cuyos experimentos tienen limitantes temporales, económicas o éticas. La combinación de estos dos aspectos, Simulación y Grandes Volúmenes de Datos, nos permitirá abordar el estudio de diferentes escenarios de una aplicación social, incluyendo un gran número de personas y una importante variedad de características ambientales, haciendo más sencillo el estudio de comportamientos complejos que surgen de la interacción de las multitudes. Nuestra propuesta podría ser utilizada para instaurar políticas de seguridad ante posibles desastres y contribuir con acciones de prevención para lograr formas eficientes de evacuar ciudades o espacios

cerrados a través del diseño de políticas activas que minimicen el tiempo de evacuación y pongan al resguardo rápidamente a las personas.

2. INVESTIGACIÓN Y DESARROLLO

2.1. MODELOS COMPUTACIONALES PARA EL PROCESAMIENTO EFICIENTE CON SoCs

Hoy en día, el desarrollo de algoritmos se focaliza en efectuar los cálculos de manera tal de conseguir un buen rendimiento, logrando a su vez una eficiencia energética. Tecnologías como arreglo de compuertas lógicas programables en campo (field programmable gate array - FPGA) y los sistemas en chip (SoC) basados en FPGA (FPGA/SoC) han demostrado que es factible la aceleración de aplicaciones computacionalmente intensivas, disminuyendo el consumo de potencia gracias a la capacidad del alto paralelismo que puede obtenerse con estas plataformas y a la reconfiguración de la arquitectura.

Como se ha mostrado en trabajos anteriores [M22, GC22, GC22_1, M21], las FPGAs pueden ser exitosamente utilizados para acelerar el procesamiento de grandes volúmenes de datos. En particular, los trabajos citados se centran en la ejecución de un algoritmo de re-ranking de documentos para la Web.

Sin embargo, estas arquitecturas tienen la limitación de la cantidad de recursos que pueden utilizarse para ejecutar los algoritmos y además, el tiempo de sintetización suele ser elevado. Con lo cual hace que estas tecnologías sean difíciles de ser ampliamente utilizadas. En otras palabras, los ciclos de diseño existentes para FPGA/SoC son largos debido a la complejidad de la arquitectura.

Para reducir los costos de desarrollo y despliegue de algoritmos sobre las FPGAs, una alternativa es utilizar un modelo computacional como un estimador de rendimiento e integral para la aceleración por hardware, el cual tiene como objetivo predecir eficazmente la utilización de recursos y la latencia para FPGA/SoC, reduciendo la brecha entre las aplicaciones y las arquitecturas FPGA/SoC. Integrar modelos de costo para la predicción del rendimiento y un motor de exploración del espacio de diseño permite proporcionar información de alto nivel a los desarrolladores de hardware.

La Fig. 1 muestra el diseño de un modelo para estimar el rendimiento integral de aplicaciones para la aceleración por hardware. El modelo trabaja con micro-benchmarks y restricciones (constraints) para ajustar los parámetros de la placa a utilizar, también utiliza diferentes enfoques como autómatas y algunas técnicas que pueden ser definidas por el usuario para realizar la exploración (DSE) de todas las directivas posibles que pueden ser utilizadas para mejorar la ejecución del algoritmo.

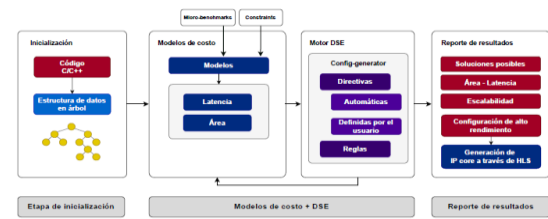


Fig. 1: Esquema general de modelo para estimar el rendimiento de aplicaciones en SoC.

2.2. CROWDSOURCING COMO BASE PARA EL PROCESAMIENTO ESCALABLE

Las plataformas de crowdsourcing han sido ampliamente utilizadas para realizar tareas en forma colaborativa. Cuanto mayor es el impacto de la situación o desastre ocurrido, mayor es la respuesta de la gente para cooperar [TW17]. A estas personas se las denomina voluntarios digitales. Los voluntarios digitales han ayudado a recopilar información pertinente y a procesarla mucho más rápido de lo que podrían hacerlo solos los funcionarios o personas a cargo de la coordinación de actividades en las situaciones particulares o, inclusive, que los algoritmos de aprendizaje que deben ser entrenados para tales fines. Es decir, en caso de que por ejemplo se deban identificar o taggear elementos de imágenes, videos o audio, los voluntarios digitales pueden hacerlo a medida que se generan los datos a procesar, mientras que los algoritmos de aprendizaje automático deben tener un conjunto de datos suficientemente grande y luego ser entrenados, lo cual puede llevar mucho tiempo.

En este contexto, se propone investigar una solución que combina técnicas de modelado y simulación discreta, análisis de sensibilidad y optimización dinámica de parámetros, para clasificar en forma eficiente imágenes y videos utilizando una plataforma de crowdsourcing que se construye sobre una red P2P.

En este sentido se ha avanzado en el diseño de la plataforma incluye un servidor conectado a una red de voluntarios a través de dispositivos set-top-box (STB) de TV digital, que constituyen una red P2P. El servidor envía tareas con imágenes e información adicional relacionada a desastres naturales a los usuarios de la red P2P y recupera los resultados. Las tareas consisten en una imagen y una lista de opciones o categorías, sobre las cuales los voluntarios deben enviar una respuesta. El servidor reúne las respuestas y evalúa si hay consenso para cada tarea, existiendo la posibilidad de reenviar las tareas en caso de ser necesario. La conexión entre la red P2P y el servidor se realiza mediante un proveedor de servicios de internet (ISP) [LOOR22][MAN19].

Además, la solución sobre la que se desea avanzar incluye: (I) El desarrollo de un simulador de una plataforma de crowdsourcing desplegada sobre una red P2P, (II) El análisis y la caracterización de los parámetros más relevantes a través de métodos de análisis de sensibilidad, (III) El diseño de algoritmos de

planificación de distribución de tareas y (IV) La optimización dinámica de parámetros de la plataforma para garantizar una ejecución eficiente. En particular la Fig. 2 muestra un esquema de cómo se pueden usar índices métricos para cumplir con este último punto. Sin embargo, también es posible llevarlo a cabo utilizando técnicas de aprendizaje por refuerzo u otras técnicas.

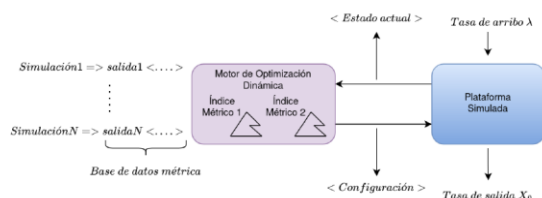


Fig. 2: Esquema general del método de ajuste de parámetros.

2.3 SIMULACIÓN DE EVACUACIONES EN CIUDADES COSTERAS

Existen diversas técnicas reportadas en la literatura para la simulación de evacuación de personas en situaciones de emergencias, como incendios, terremotos y tsunamis. La mayor parte de los trabajos presentan escenarios donde se estudian principalmente los tiempos de evacuación y densidad de personas en las zonas de encuentro [AG19]. En base a estos resultados, se crean, diseñan y evalúan los planes de evacuación que deberán ser socializados a los habitantes de la zona respectiva. Por lo tanto, la generación de escenarios con distintos parámetros, por ejemplo, diferentes tipos de población, intensidades de sismos y niveles de inundación, puede ser una herramienta de gran utilidad. Esto permitiría crear planes de evacuación efectivos y que se acerquen a la realidad, luego de analizar los resultados de las simulaciones de los distintos escenarios. En este sentido, se pretende investigar el desarrollo de un método para realizar un análisis de múltiples escenarios en forma eficiente utilizando herramientas como análisis geo-espacial y minería de datos para generar mapas de predicción que apoyen los respectivos planes de evacuación. El trabajo presentado en [AST22] presenta una primera aproximación para obtener esta metodología mediante un simulador basado en agentes denominado DEMPS. Este trabajo utiliza el protocolo ODD [GRIMM20] para describir el modelo desarrollado. El simulador desarrollado tiene un entorno visual que permite visualizar el avance de los agentes, definir las zonas seguras de evaluación, las zonas iniciales donde están los agentes y además reportar datos estadísticos como se muestra en la Fig. 3.

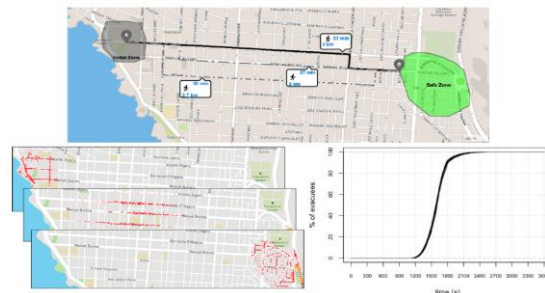


Fig. 3. Simulador basado en agentes.

2.4. SIMULACIONES DE EVACUACIONES EN AMBIENTES CERRADOS

Otra área donde el grupo de investigación tiene experiencia es en simulaciones de evacuaciones en ambientes cerrados. En este contexto, a pesar del surgimiento de numerosas reglamentaciones, aún existen numerosas edificaciones con capacidad de alojar cientos o miles de individuos en su interior, donde es posible observar que cualquier incidente podría comprometer la seguridad de sus ocupantes. La tarea de predecir la performance de una evacuación para grandes y complejas estructuras con cientos de individuos representa una labor compleja, la cual se ve incrementada cuando se considera la interacción entre los individuos y la presencia de fuego y humo en el ambiente. Realizar estudios preventivos que permitan analizar las consecuencias de futuros siniestros permitiría tomar las medidas de seguridad necesarias para evitar que tales catástrofes ocurran. Desde este punto de vista, las simulaciones de las dinámicas pedestres en situaciones de emergencia son una herramienta de vital importancia, ya que son capaces de mostrar un gran número de características existentes en el tráfico de peatones, permitiéndonos de esta manera tener una mejor comprensión de los principios fundamentales que rigen este fenómeno. Además, dan la posibilidad de explorar los efectos que tienen en el modelo el cambio de reglas o de la geometría del espacio físico, sin la necesidad de realizar costosos experimentos, los cuales incluso pueden ser peligrosos o hasta imposibles de realizar. Mediante la modelización, es posible estudiar nuevos diseños o políticas en el sistema sin la necesidad de construir los ambientes físicos y las situaciones por ellas modeladas, permitiendo detectar errores en los diseños. Además, cuentan con las ventajas de evitar los riesgos de experimentación, y de poder analizar su comportamiento y resultados de forma segura.

De acuerdo con lo expuesto, se pretende dar continuidad a la investigación y desarrollo de un nuevo framework de simulación tipo cloud “Software as a Service” (SaaS), basado en el modelo de interacción de autómatas celulares y agentes inteligentes, presentado en [TI14] que permite instanciar múltiples realidades o escenarios a ser estudiados mediante la experimentación simulada.

Se pretende además en este contexto, investigar tanto impacto como las relaciones de los paradigmas y técnicas

de simulación asociadas al campo de la ciencia de datos, haciendo especial hincapié en la utilización de las mismas en las etapas tempranas de preprocesamiento. Finalmente, se desea investigar el uso y la aplicación de técnicas y herramientas de visualización de datos para el análisis de información obtenida mediante simulación utilizando el framework propuesto.

2.5. LA VISUALIZACIÓN COMO HERRAMIENTA PARA LA INTERPRETACIÓN DE DATOS

Como se mencionó anteriormente, la ciencia de datos es un área que ha llamado la atención en muchas disciplinas y está siendo usada en diferentes ámbitos tanto académicos como industriales. Los mercados financieros, en particular, el mercado de las criptomonedas es un área de interés para la comunidad científica teniendo en cuenta dos puntos de vista [BARI]: tecnológico y económico/financiero.

La ciencia de datos permite usar técnicas para procesar los datos relacionados con el problema a estudiar, en tiempo real y para eso hay que buscar la forma de leer datos y extraer conocimiento en un tiempo razonable. Aquí es un punto crucial donde se une ciencia de datos con diseño de algoritmos y infraestructuras eficientes para grandes volúmenes de datos. Por ejemplo, desde el punto de vista económico y financiero, las investigaciones tienden a analizar todas las características financieras de las criptomonedas no sólo por su volumen de negociación sino también por su comportamiento distintivo frente a activos tradicionales y por el hecho de que las criptomonedas se comercializan las 24 horas del día, los 7 días de la semana, planteando nuevos desafíos sobre cómo procesar de manera eficiente un flujo continuo de grandes volúmenes de datos. Dada su reciente popularidad, estos conjuntos de datos son cada vez más grandes. La existencia de más de 1.000 tipos de criptomonedas en el mundo, y la certeza de que seguirán creciendo, tanto en número como en popularidad, es muestra de su auge y su entrada activa en el mercado financiero internacional.

Por lo tanto, el tratamiento de dichos datos involucra la aplicación de una gran variedad de herramientas para su análisis y exploración. Una de las técnicas de interés, dentro de las etapas de la ciencia de datos, es la visualización de datos, que permite transmitir información por medio de representaciones visuales interactivas de datos, siendo una herramienta adecuada para enfrentarnos a la gran cantidad de datos que se manejan actualmente como así también para su correcta interpretación. En este contexto, el objetivo general es contribuir al desarrollo de soluciones en torno al análisis visual de indicadores relevantes para el mercado de las criptomonedas.

En este sentido, se ha desarrollado un prototipo con interacciones dinámicas que busca responder preguntas iniciales que dan el puntapié inicial para seguir trabajando con dichos datos.

2.6. PRESTACIONES SOCIALES Y RESILIENCIA

Este proyecto integra el Consorcio de I+D+I en Cloud Computing-Big Data & Emergent Topics, el cual propone un modelo de trabajo inclusivo y orientado a adaptarse a las necesidades de una sociedad que requiere incorporar los Objetivos de Desarrollo Sustentable. En esa dirección, el presente proyecto se plantea dos objetivos específicos: a) analizar cómo las tecnologías de transformación digital que se utilizan en sus distintas líneas (tales como datos, información y conocimiento) tienen implicancia en las prestaciones sociales (human centric) que se abordarán; y b) analizar el grado de Resiliencia de las aplicaciones descritas en los apartados anteriores y desarrollar para éstas un *indicador de performance clave (KPI)* que muestre como un sistema, cuyo estudio se realiza por medio de simulación, está alcanzando sus objetivos.

3. RESULTADOS OBTENIDOS/ESPERADOS

Las líneas de investigación y las técnicas descritas en la sección anterior, involucran una serie de desarrollos individuales que en su conjunto logran obtener el objetivo planteado. Este objetivo contempla por un lado el modelado y diseño de sistemas complejos mediante diferentes herramientas, algunas de las cuales han sido desarrolladas por los integrantes del proyecto. Mediante estas herramientas se pretende facilitar el entendimiento del sistema, realizar *profiling* a la ejecución de los algoritmos y mejorar el análisis de resultados. Por otro lado, se propone utilizar técnicas de ciencias de datos para analizar los datos obtenidos y encontrar comportamientos o tendencias que puedan ser utilizadas en diferentes situaciones que requieran soporte a la toma de decisiones. Finalmente, el procesamiento de este gran volumen de datos debe realizarse en forma eficiente y en el menor tiempo posible. Esto implica investigar formas de despliegue de los algoritmos en diferentes plataformas y tecnologías que permitan lograr el objetivo deseado.

Los resultados obtenidos hasta el momento son:

- El equipo de investigadores y estudiantes, han implementado herramientas de simulación que se encuentran disponibles en GitHub.
- Se han obtenido publicaciones en revistas indexadas y congresos internacionales.
- Se han dictado cursos de posgrado relacionados con la temática (en Argentina y en Chile).
- Se ha participado en proyectos de cooperación internacional como Stic-Amsud y CNR-CONICET.
- Se obtuvo el premio: Industry Impact Award. ECIR 2022. Paper: Ensemble Model Compression for Fast and Energy-Efficient Ranking on FPGAs.
- En marzo del 2023 finalizó una co-tutela con la Universidad de Trieste, Italia, para una tesis de doctorado en Ciencias de la Computación.
- En 2022, finalizó una tesis de doctorado y una tesis de maestría en colaboración con la Universidad de Santiago de Chile.

Los resultados esperados son:

- Avanzar en formular nuevas metodologías y técnicas para la adquisición, tratamiento y análisis de datos masivos.
- Investigar el uso de simuladores eficientes y eficaces para evaluar el rendimiento de nuevos algoritmos para aplicaciones de sistemas de gran escala con datos masivos.
- Abordar problemas reales que involucren datos reales, desarrollando soluciones eficientes para el análisis de grandes volúmenes de datos.
- Continuar formando estudiantes de posgrado en la temática.
- Capacitar profesionales en la temática mediante talleres y cursos de posgrado.

4. FORMACION DE RECURSOS HUMANOS

Actualmente, se cuenta con tres doctores en ciencias de la computación realizando la investigación y dirección de los doctorandos involucrados en las líneas. Se cuenta con dos estudiantes de doctorado, los cuales se encuentran en la etapa final de su trabajo. El primer estudiante realiza su tesis en el desarrollo en el análisis, diseño y evaluación de factibilidad de plataformas distribuidas eficientes para el procesamiento de grandes cantidades de datos en el menor tiempo posible. En este trabajo se combinan técnicas de simulación y aprendizaje por refuerzo. El tema de estudio del segundo estudiante de doctorado, incluye el modelo, diseño e implementación eficiente de simuladores para procesar aplicaciones de gran escala en arquitecturas heterogéneas. También se cuenta con una integrante que ha concluido el trabajo final de la especialización en Inteligencia de Datos Orientada a Big Data utilizando técnicas de visualización para el tratamiento de datos financieros y se cuenta con un estudiante de la Lic. en Ciencias de la Computación.

Mediante este trabajo de investigación se podrán formar profesionales que puedan modelar, diseñar e implementar algoritmos eficientes (previamente evaluados y analizados con un mínimo porcentaje de error) que se ejecuten en sistemas complejos donde intervienen una gran cantidad de variables y requieren el procesamiento de datos masivos. Por otro lado, se investigará en el uso de técnicas de ciencia de datos para obtener información a partir de los datos procesados y que puedan ser utilizados en la toma de decisiones. Finalmente, como se comentó anteriormente, los datos obtenidos para el procesamiento eficiente y obtención de patrones mediante ciencia de datos, pueden ser utilizados para generar modelos y simular escenarios alternativos.

5. BIBLIOGRAFÍA

[AG19] Aguilar, L., Wijerathne, L., Jacob, S., Hori, M., & Ichimura, T. (2019). Mass evacuation simulation considering detailed models: behavioral profiles, environmental effects, and mixed-mode evacuation. *Asia Pacific Management Review*.

[AST22] Efficient simulation of natural hazard evacuation for seacoast cities Astudillo Muñoz, G., Gil-

Costa, V., Marin, M. *International Journal of Disaster Risk Reduction*, 2022

[GC22] V. Gil-Costa, F. Loor, R. Molina, F. Maria Nardini, R. Perego, S. Trani. Ensemble Model Compression for Fast and Energy-Efficient Ranking on FPGAs. *ECIR 2022*: 260-273

[GC22_1] V. Gil-Costa, F. Loor, R. Molina, F. Maria Nardini, R. Perego, S. Trani: Energy-Efficient Ranking on FPGAs through Ensemble Model Compression. *IIR 2022*

[LOOR22] F. Loor, M. Manriquez, V. Gil-Costa, M. Marín. Feasibility of P2P-STB based crowdsourcing to speed-up photo classification for natural disasters. *Clust. Comput.* 279-302 (2022)

[M21] R. Molina, F. Loor, V. Gil-Costa, F. Maria Nardini, R. Perego, S. Trani. Efficient traversal of decision tree ensembles with FPGAs. *J. Parallel Distributed Comput.* 155: 38-49 (2021)

[M22] R. Soledad Molina, V. Gil-Costa, M.L. Crespo, G. Ramponi. High-Level Synthesis Hardware Design for FPGA-Based Accelerators: Models, Methodologies, and Frameworks. *IEEE Access* (2022).

[MAN19] M. Manriquez, F. Loor, V. Gil-Costa and M. Marin. "A Digital Tv-based Distributed Image Processing Platform for Natural Disasters" *WinterSim 2019*

[GRIMM20] V. Grimm, S.F. Railsback, C.E. Vincenot, U. Berger, C. Gallagher, D.L. DeAngelis, B. Edmonds, J. Ge, J. Giske, J. Groeneveld, et al., The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism, *J. Artif. Soc. Soc. Simul.* (2020).

[BARI] I. Merediz-Solà and A. Bariviera. A bibliometric analysis of bitcoin scientific production. *Research in Int. Business and Finance*, 2019

[TW17] Twigg and Mosel (2017). Emergent groups and spontaneous volunteers in urban disaster response. *Environment and Urbanization*, 29(2), 443-458.

[TI14] (2014) Tissera C, Modelo Basado en Autómatas Celulares extendidos para Diseñar Estrategias de Evacuaciones en Casos de Emergencias. UNSL.

Autoescalado basado en Aprendizaje Profundo por Refuerzo de Workflows Científicos en la Nube

Elina Pacini^{1,2}, Carlos Catania¹, Yisel Garí¹ and Luciano Robino¹

¹ Facultad de Ingeniería, UNCuyo

² CONICET

elina.pacini@ingenieria.uncuyo.edu.ar, harpo@ingenieria.uncuyo.edu.ar, yisलगari@gmail.com,
luciano.ivan.robino@gmail.com

RESUMEN

Los workflows científicos representan una importante abstracción para el modelado y la ejecución de experimentos de gran envergadura en múltiples disciplinas. El paradigma de computación Cloud facilita el acceso elástico a la infraestructura computacional requerida para la ejecución de este tipo de aplicaciones. Las estrategias de autoescalado buscan explotar la elasticidad de Cloud para lograr ejecuciones eficientes de los workflow. El autoescalado es un problema de toma de decisiones en el cual es necesario establecer cuándo y cómo incrementar o reducir los recursos computacionales, y cómo ajustarlos a la carga de trabajo actual considerando algún criterio de optimización como pueden ser el tiempo y el costo económico asociado. El aprendizaje por Refuerzo (AR) aparece como un enfoque promisorio para el autoescalado en Cloud ya que permite aprender políticas adecuadas para el manejo de recursos que a su vez son transparentes (sin intervención humana), dinámicas (sin planes computados estáticamente) y adaptables (constantemente actualizadas). Luego, el Aprendizaje Profundo por Refuerzo (APR) permite además el manejo de espacios de estados y acciones de grandes dimensiones, una importante limitante en los enfoques más clásicos del AR.

CONTEXTO

El presente proyecto se desarrolla en el marco de la Facultad de Ingeniería dentro Laboratorio de sistemas inteligentes (LABSIN). El presente trabajo forma parte del proyecto de investigación B038-T1 que dio inicio en el mes de mayo de 2022 en el marco de los proyectos bienales de secretaria de Investigación, Internacionales y Posgrados (SIIP) de la Universidad Nacional de Cuyo.

1 INTRODUCCIÓN

Los workflows científicos [1] se han establecido como una importante abstracción para el procesamiento de datos y la ejecución de grandes y complejos experimentos científicos. A partir del modelado de los pasos de procesamiento y sus dependencias, los workflows permiten contar con una panorámica más clara para el análisis de las aplicaciones. Los workflows científicos usualmente se componen por cientos o miles de tareas de diferentes duraciones que pueden ir desde horas o días hasta semanas. Además, dichas tareas presentan diferentes requerimientos de recursos computacionales (CPU, RAM, ancho de banda). Por otro lado, las estructuras de dependencias entre las tareas de un workflow determinan una carga de trabajo variable durante la ejecución, con momentos donde muchas tareas pueden ser ejecutadas en

paralelo y momentos con largas ejecuciones secuenciales de tareas. Tal variabilidad determina momentos donde se requiere una infraestructura con mayor o menor capacidad computacional.

En este contexto, el modelo de computación Cloud, aparece como una alternativa ideal para la ejecución de workflows científicos, ya que facilita el acceso elástico a la infraestructura computacional requerida por este tipo de aplicaciones [2]. Particularmente, el modelo de servicios IaaS en Cloud permite a los usuarios crear y destruir instancias de Máquinas Virtuales (MV) de diferentes prestaciones bajo un esquema de pago por uso.

Luego, debido a la elasticidad provista por los modelos de computación Cloud, es posible ajustar dinámicamente la infraestructura ante las variaciones en la demanda que se presentan durante la ejecución de los workflows. Esta flexibilidad junto a la necesidad de lograr ejecuciones eficientes de los workflows, alienta el estudio y el desarrollo de estrategias para el autoescalado de workflows y otros tipos de aplicaciones científicas en Cloud [3,4].

Específicamente, las estrategias de autoescalado [5,6,7] resuelven periódicamente dos problemas de optimización interrelacionados, el escalado y la planificación. El escalado consiste en ajustar la capacidad computacional de infraestructura virtualizada en función de la demanda de la aplicación. La planificación consiste en asignar la ejecución de cada tarea a alguna de las instancias de MVs disponibles. Ambos subproblemas (escalado y planificación) son NP-duros por lo que suelen ser abordados con heurísticas o metaheurísticas.

Es importante considerar que existen varios elementos de incertidumbre en la ejecución de las aplicaciones que son inherentes al modelo de computación Cloud. Entre ellos se encuentran, la variabilidad en el

rendimiento de la infraestructura virtualizada, las fluctuaciones en los precios de algunos tipos de instancias y sus posibles fallos. Todo esto hace que sea más complejo poder determinar de antemano las mejores decisiones para lograr una ejecución eficiente.

El Aprendizaje por Refuerzo (AR) [19] constituye uno de los paradigmas del aprendizaje automático, como lo son también el aprendizaje supervisado y no supervisado. En el caso del AR se trata de un enfoque computacional que permite a un agente aprender, mediante la interacción con el entorno, un comportamiento adecuado para el logro de un determinado objetivo. Cuando es posible modelar la dinámica del entorno, el agente puede aprender en modo offline utilizando técnicas específicas para ello. Otra variante consiste en que el agente interactúe directamente con el entorno real aprendiendo en modo online. Esta última variante suele ser más utilizada ya que en muchos problemas no es posible contar con un modelo preciso de la dinámica del entorno o la misma puede cambiar en el tiempo.

Las estrategias basadas en AR están siendo ampliamente utilizadas y con muy buenos resultados en áreas como la teoría de juegos, lo que ha motivado su estudio y aplicación en otras áreas como es el caso del autoescalado en Cloud.

Varias investigaciones recientes proponen soluciones basadas en AR para resolver alguno de los subproblemas involucrados, el escalado [4,9,10] o la planificación [11,12]. El AR ofrece un marco de trabajo de propósito general para la toma de decisiones que ha demostrado un gran potencial en entornos estocásticos complejos. El AR propone un modo de aprendizaje basado en la interacción con el entorno y dirigido por objetivos. La idea consiste en que un agente puede observar el estado del entorno y tomar acciones que

influyen en las observaciones futuras. La retroalimentación de las acciones elegidas llega al agente como señales escalares denominadas recompensas, y el objetivo principal del agente es maximizar la recompensa total a largo plazo. El agente observa el entorno, actúa sobre él y evalúa el impacto de sus acciones. De esta manera puede aprender un comportamiento casi óptimo por prueba y error. Los elementos mencionados - estados, acciones y recompensas - son los componentes clave en la formulación de cualquier problema de AR, y el marco utilizado para definirlos se denomina Proceso de Decisión de Markov (PDM).

Este problema (conocido en inglés como *the curse of dimensionality*) es una importante limitante en los enfoques más clásicos del AR, en algunos casos afectando el tiempo de aprendizaje y en los casos más críticos volviéndose imposible de manejar. De esta manera, ante entornos complejos, el Aprendizaje Profundo por Refuerzo (APR) [13] ofrece un marco de trabajo más potente que el ofrecido por las estrategias clásicas del AR, ya que combina la potencialidad de dos ramas importantes del aprendizaje automático, esto es las Redes Neuronales Profundas (RNP) y el AR, para la resolución automática de problemas de toma de decisiones en entornos estocásticos altamente complejos.

Es importante destacar que la directora del proyecto cuenta con aproximadamente 10 años de experiencia en la implementación de técnicas metaheurísticas para la optimización de aplicaciones científicas en entornos Cloud. [1,14,15]. Además, en cuanto a la implementación de técnicas metaheurísticas para autoescalado en Cloud se pueden mencionar los siguientes trabajos [5,6,7]. Finalmente, cabe mencionar que el problema de autoescalado en Cloud basado en AR se ha avanzado hasta el momento en un exhaustivo estudio del estado del arte [16] y se ha trabajado en la implementación

de un autoescalador [4,17,18] que se tomará como base para extender el trabajo propuesto en el presente proyecto.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El AR ha demostrado un gran potencial para la resolución automática de problemas de toma de decisiones en entornos complejos y con incertidumbre, particularmente debido a su capacidad para considerar las consecuencias a largo plazo de las acciones disponibles. A partir del análisis del estado del arte [16] y de las limitaciones de los estudios previos realizados, se hace evidente que existe actualmente una larga lista de problemas abiertos, lo que a su vez significa que existe un amplio espectro de oportunidades de investigación con respecto a las técnicas de AR en el área de autoescalado en Cloud.

A continuación, se analizan 5 líneas de interés que guiarán luego a los objetivos y a las actividades específicas del presente plan de trabajo y que surgen de las limitaciones fundamentales encontradas en los trabajos del área [16].

a) Escalado y Planificación con APR. Es importante diseñar y desarrollar estrategias de autoescalado de workflows en Cloud que combinen un escalador y un planificador, ambos basados en APR, debido al impacto conjunto que tienen tanto las decisiones de escalado como las de planificación en la eficiencia de las ejecuciones de los workflows. Una solución integral permitiría expandir las ventajas (esto es, transparencia, dinamismo y adaptabilidad) de este enfoque de aprendizaje a todo el proceso de autoescalado. La política de escalado determinará cómo ajustar la infraestructura de forma dinámica según la demanda variable de la aplicación, mientras que la política de planificación determinará el recurso más adecuado para la ejecución de cada tarea, considerando información de las tareas y de la infraestructura disponible. Ambas

políticas serían aprendidas observando los efectos conjuntos de las acciones de escalado y planificación en la ejecución del workflow en una infraestructura Cloud.

b) *Elementos de los Entornos Cloud*. Es muy importante ampliar los estudios actuales para representar más elementos de los entornos Cloud reales (por ejemplo, más tipos de MVs y modelos de precios). Este tipo de extensión podría generar problemas de escalabilidad relacionados con la dimensión del espacio de estados y/o de acciones. En este sentido, se propone utilizar técnicas del APR que combinan las potencialidades del AR y de las RNP para resolver este tipo de problemáticas. En el área del autoescalado en Cloud, la mayoría de los enfoques aún utilizan las variantes más simples de los métodos de AR. Por lo tanto, todavía hay espacio para investigar más a fondo la sinergia de diferentes variantes de las estrategias básicas de AR y otros métodos de aprendizaje automático.

c) *Rendimiento Inicial y Convergencia*. Consiste en evaluar y proponer soluciones que garanticen un rendimiento inicial adecuado de las estrategias de autoescalado, así como implementar mecanismos para la aceleración de la velocidad de convergencia en el aprendizaje de las políticas. Ambos factores son de especial interés en el contexto de la ejecución de workflows en Cloud, ya que la reducción del tiempo de ejecución y del costo económico son objetivos primordiales que se ven afectados directamente cada vez que la estrategia de autoescalado no es capaz de tomar decisiones adecuadas de escalado y/o de planificación.

d) *Representación de Estados y Acciones*. La representación de los estados y las acciones es un elemento determinante en el aprendizaje y el desempeño de las estrategias. En AR, como en otras áreas del Aprendizaje Automático, estas opciones de representación, puede decirse que son más arte que ciencia. En el área del autoescalado en Cloud, es fundamental estudiar las implicaciones específicas de tales elecciones de representación (estados y acciones) y cómo impactan en el rendimiento

de las estrategias de autoescalado. Por ejemplo, preguntas interesantes para responder en este contexto son: (1) ¿Qué información de un entorno real Cloud es relevante para aprender adecuadamente una política? (2) ¿Cuál podría ser una representación adecuada de esta información para acelerar el proceso de aprendizaje?

3 RESULTADOS ESPERADOS

Al término de los dos años de duración del plan de trabajo se pretende que se haya logrado:

1) Contar con un análisis del estado del arte actualizado en el marco del tema desarrollado. Este análisis proveerá una además una guía para las futuras líneas de investigación fortaleciendo así el área.

2) Disponer de con una extensión del autoescalador que incorpore técnicas más promisorias basadas en APR para la resolución del problema estudiado.

4 FORMACIÓN DE RECURSOS HUMANOS

Se espera además impactar fuertemente en la formación de recursos humanos mediante la codirección de doctorado del Lic. Luciano Robino (integrante de este proyecto) y de la beca postdoc de Yisel Garí (Doctora en Ciencias de la Computación).

5 AGRADECIMIENTOS

Agradecemos el soporte financiero obtenido a través de los proyectos SIIP-UNCuyo B038-T1 y 06/B040-T1.

6 BIBLIOGRAFIA

- [1] E. Deelman, T. Peterka, I. Altintas et al.: The future of scientific workflows. The International Journal of High-Performance Computing Applications. 2018;32(1):159-175.
- [2] M.A.S. Netto, R.N. Calheiros, E.R. Rodrigues, R.L. F. Cunha, and R. Buyya: Hpc cloud for scientific and business applications:

Taxonomy, vision, and research challenges. *ACM Computing Surveys*, 51(1):1–29, 2018.

[3] D.A. Monge, Y. Garí, C. Mateos, and C. García Garino. Autoscaling scientific workflows on the cloud by combining on-demand and spot instances. *International Journal of Computer Systems Science and Engineering*, 32(4 Special Issue on Elastic Data Management in Cloud Systems), Jul 2017.

[4] Y. Garí, D.A. Monge, C. Mateos, and C. García Garino. Learning budget assignment policies for autoscaling scientific workflows in the cloud. *Cluster Computing*, 23:87–105, 2020.

[5] D.A. Monge, E. Pacini, C. Mateos, E. Alba and C. García Garino: CMI: An online multi-objective genetic autoscaler for scientific and engineering workflows in cloud infrastructures with unreliable virtual machines. *Journal of Network and Computer Applications* 149, (2020).

[6] D.A. Monge, E. Pacini, C. Mateos and C. García Garino, C.: Meta-heuristic based autoscaling of cloud-based parameter sweep experiments with unreliable virtual machines instances. *Computers & Electrical Engineering* 69, 364–377 (2018).

[7] V. Yannibelli, E. Pacini, D. Monge, C. Mateos and G. Rodriguez: A Comparative Analysis of NSGA-II and NSGA-III for Autoscaling Parameter Sweep Experiments in the Cloud. *Scientific Programming* 2020, Article ID 653204, 1-17 (2020).

[8] R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, USA, 2018.

[9] X. Dutreilh and S. Kirgizov. Using reinforcement learning for autonomic resource allocation in clouds: towards a fully automated workflow. In *7th International Conference on Autonomic and Autonomous Systems*, pages 67–74, 2011.

[10] E. Barrett, E. Howley, and J. Duggan. Applying reinforcement learning towards automating resource allocation and application scalability in the cloud. *Concurrency Computation Practice and Experience*, 2012.

[11] E. Barrett, E. Howley and J. Duggan. A learning architecture for scheduling workflow applications in the cloud. *Proceedings – 9th IEEE European Conference on Web Services, ECOWS 2011*, pages 83–90, 2011.

[12] Z. Peng, D. Cui, J. Zuo, Q. Li, B. Xu, and W. Lin. Random task scheduling scheme based on reinforcement learning in cloud computing. *Cluster Computing*, 18(4):1595–1607, 2015.

[13] F. Palau, C. Catania, J. Guerra, S. Garcia and M. Rigaki. Detecting DNS Threats: A Deep Learning Model to Rule Them All. *ASAI, Simposio Argentino de Inteligencia Artificial*, pages 90-101, 2019.

[14] E. Pacini, C. Mateos, and Carlos García Garino. A three-level scheduler to execute scientific experiments on federated clouds. *IEEE Latin America Transactions*, 13(10):3359–3369, 2015.

[15] E. Pacini, C. Mateos, C. García Garino, C. Careglio, and A. Mirasso. A bioinspired scheduler for minimizing makespan and flowtime of computational mechanics applications on federated clouds. *Journal of Intelligent & Fuzzy Systems*, 31(3):1731–1743, 2016. IOS Press.

[16] Y. Garí, D.A. Monge, E. Pacini, C. Mateos, C. García Garino: “Reinforcement Learning based Autoscaling of Workflows in the Cloud: A Survey”. *Engineering Applications of Artificial Intelligence*, Vol. 102, 2021.

[17] Y. Garí, D.A. Monge, C. Mateos, C. García Garino. Markov Decision Process to Dynamically Adapt Spots Instances Ratio on the Autoscaling of Scientific Workflows in the Cloud. *Latin American High-Performance Computing Conference*, pages 353-369, 2019.

[18] Y. Garí, D.A. Monge, C. Mateos. A Q-learning approach for the autoscaling of scientific workflows in the Cloud. *Future Generation Computer Systems* 127: 168-180, 2021.

ANÁLISIS DE RENDIMIENTO DE APLICACIONES Y SISTEMAS COMPLEJOS MEDIANTE COMPUTACIÓN PARALELA Y TÉCNICAS DE OPTIMIZACIÓN DE RECURSOS

Miguel Méndez-Garabetti^{1,2,3,4,5}, Eduardo Piray^{1,3,4}, Javier Roseinstein², Ignacio Picotto¹, Rodrigo Elgueta^{2,3,5}, Marcos Benegas², Natalia Anahí Magris¹, Andrea Cabrera², Manuel Battaglia², Agustín Giorlando^{3,5}, Gabriel Nasiff⁵, Jonathan Nicolás Guerra Moronta⁵

¹Universidad Siglo 21, Córdoba, Argentina.

²Facultad de Informática y Diseño, Universidad Champagnat, Mendoza, Argentina.

³Free and Open Source Software/Hardware Research Laboratory (FOSSHLab), Argentina.

⁴Departamento de Sistemas, Universidad CAECE, Mar del Plata, Argentina.

⁵Dirección de Posgrados, Facultad de Ingeniería, Universidad de Mendoza, Mendoza, Argentina.

mmendezgarabetti@gmail.com, edupiray@gmail.com, roseinstein.javier@uch.edu.ar,
ignaciopicotto@gmail.com, rodrigo.elgueta@uch.edu.ar, nmagris@outlook.com,
andreabcaabrera@gmail.com, manubattaglia94@gmail.com, ganasiff@gmail.com,
j.guerra@alumno.um.edu.ar

RESUMEN

El crecimiento exponencial de la potencia de las tecnologías de procesamiento, almacenamiento, comunicaciones ha permitido que los sistemas de computación de alto rendimiento aborden problemas científicos que son cada vez más grandes y complejos. Estas aplicaciones, que son diseñadas para ejecutarse en plataformas de hardware específicas, suelen tener diversos requerimientos donde la capacidad de procesamiento no necesariamente es el principal recurso que determina el rendimiento de éstas. En ocasiones otros recursos, como la capacidad de E/S, la memoria y el ancho de banda suelen ser protagonistas. En este sentido, esta línea de investigación plantea indagar en diferentes aristas de la computación paralela, tales como el diseño de técnicas de optimización de recursos basados de plataformas de HPC.

Palabras claves: Performance. Mejora de rendimiento. Optimización de recursos. Computación de alto rendimiento. Arquitecturas paralelas.

CONTEXTO

La presente línea de investigación consiste en una iniciativa encabezada por el FOSSHLab (Free and Open Source Software/Hardware Research Laboratory), institución que impulsa el desarrollo de proyectos centrados en tecnologías de software y hardware libre. Actualmente la línea cuenta con tres proyectos de investigación llevados a cabo en universidades diferentes: 1) Universidad Siglo 21, en el marco de la carrera Licenciatura en Informática (resolución en trámite), 2) Universidad Champagnat, entorno a la Licenciatura en Sistemas de Información (resolución 59-2022-UCH), 3) Universidad CAECE, en el marco de la Licenciatura en Sistemas, la Ingeniería en Sistemas y la Licenciatura en Gestión de Sistemas y

Negocios (resolución en trámite) y 4) Universidad de Mendoza, en el marco de la Ingeniería en Informática y la Maestría en Teleinformática (resolución 36/2021). Cada proyecto está financiado individualmente por su respectiva universidad.

1. INTRODUCCIÓN

Desde el origen de la computación ha habido un constante y arduo esfuerzo por conseguir mayores prestaciones computacionales. Este desarrollo ha generado grandes logros, materializados en avances significativos en el hardware y software de computadoras, haciendo cada vez más rápidos y potentes los sistemas de procesamiento, almacenamiento y comunicaciones.

Este incremento de poder computacional ha permitido abordar problemas cada vez más grandes y complejos, aunque, para una gran cantidad de éstos, aún sigue siendo necesario contar con mayor capacidad de procesamiento. Este tipo de requerimientos ha desembocado en la utilización de múltiples procesadores agrupados para trabajar de forma conjunta. Esta práctica se conoce como Computación de Alto Rendimiento (HPC, High Performance Computing), concepto también asociado al de Computación Paralela [1]. En términos generales, se puede decir que el HPC consiste en la utilización de determinada cantidad de elementos de procesamiento en la resolución de un problema de forma cooperativa, tratando de minimizar el tiempo de procesamiento y dar así una respuesta lo más rápido posible.

Aprovechar al máximo el potencial de estas plataformas es un gran desafío debido a las complejas interacciones que se producen entre el hardware y el software, sumado a la complejidad de las aplicaciones que se deben ejecutar en estos entornos, siendo que muchas de ellas no han sido diseñadas específicamente para ser ejecutadas en estas plataformas. Es aquí donde el análisis de rendimiento, revisión

de diseño y arquitectura de aplicaciones juega un rol central en el proceso de desarrollo de aplicaciones paralelas.

A medida que la tecnología avanza, el costo de los sistemas de HPC ha disminuido significativamente, lo que ha permitido que cada vez más organizaciones y empresas puedan acceder a estas capacidades. Sin embargo, el diseño y la optimización de aplicaciones para sistemas de HPC sigue siendo un desafío importante, ya que requiere conocimientos especializados tanto en el hardware como en el software.

En este contexto, es común observar que este tipo de soluciones requieren cada vez más capacidad de cómputo, comunicación y almacenamiento; si se desea dar respuesta en un plazo de tiempo durante el cual los resultados aún sigan teniendo validez. Es importante tener en consideración, que el aumento de recursos computacionales no implica de forma transparente un incremento en el rendimiento y, además, lograr mantener en funcionamiento plataformas HPC de gran escala generalmente implica un gran consumo de energía. Por lo tanto, es imprescindible además de esforzarse por mejorar el rendimiento, lograr optimizar la utilización de recursos. En otras palabras, resulta indispensable considerar la necesidad de abordar simultáneamente la consecución de dos objetivos: a) maximizar el rendimiento y b) minimizar el consumo energético.

En el desarrollo de aplicaciones paralelas es necesario contar con el manejo de aspectos propios del paralelismo, los que deben ser tenidos en cuenta para que una aplicación logre los objetivos principales de obtener el resultado esperado y el de alcanzar un desempeño acorde al entorno de ejecución [2]. Debido a esto, una vez que se ha implementado y testeado la aplicación, es necesario efectuar un análisis de rendimiento para determinar si coincide con el esperado. Si bien existen diferentes métricas para determinar el

desempeño de una aplicación paralela, (como el tiempo de ejecución, eficiencia, speed-up, balanceo de carga, entre otras) es necesario contar con herramientas que ofrezcan la información necesaria para determinar qué características deben ser mejoradas para lograr mejor rendimiento. La revisión y ajuste de una aplicación representa una parte importante y necesaria en el desarrollo de programas, ya que permite adaptar y mejorar el comportamiento de las aplicaciones mediante la detección de aquellos aspectos que impiden alcanzar el rendimiento esperado, y mediante la definición y aplicación de acciones tendientes a mejorar dichos aspectos [3].

En los clústeres de computadoras homogéneos la programación se lleva a cabo mediante la librería de paso de mensajes (MPI, Message Passing Interface) [4], además, dentro de cada nodo individual, generalmente con arquitecturas multicore se puede utilizar OpenMP [5], bajo el paradigma de memoria compartida. Ambos enfoques pueden utilizarse de forma conjunta o híbrida, con el propósito de incrementar el nivel de paralelismo en función del tipo de problema que se intente resolver y la plataforma de hardware subyacente disponible.

En esta línea, también es posible utilizar estos enfoques con otro basado principalmente en memoria compartida, que es compatible con el de memoria compartida-distribuida, lo que se conoce como computación heterogénea. Dicha implementación consta del uso de paralelismo mediante unidades de procesamiento gráfico de propósito general (GPGPU, General Purpose Graphics Processing Units) mediante CUDA (Compute Unified Device Architecture), la arquitectura de cálculo paralelo de NVIDIA [6].

En la literatura podemos observar numerosos antecedentes que dan cuenta que el uso de la computación paralela permite incrementar el rendimiento de una aplicación. En [7] se propone la paralelización de cuatro

aplicaciones científicas diseñadas con una programación totalmente secuencial. Se aplica una metodología sistemática para transformar un código base antiguo en código moderno, es decir, código paralelo y robusto.

En [8] [9], la utilización de la computación paralela/distribuida queda en clara evidencia su importancia, ya que es el medio que permite desarrollar sistemas de predicción de fenómenos naturales donde sus resultados lleguen a tiempo para poder ser tomados para la toma de decisiones.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se describen brevemente cada línea de investigación:

1. Una de estas líneas de investigación se enfoca en la optimización de metaheurísticas paralelas de forma adaptativa. Este enfoque busca mejorar la calidad de la búsqueda y la eficiencia del algoritmo mediante la adaptación de los parámetros de la metaheurística de acuerdo con las características del problema en cuestión. De esta manera, se espera obtener una mayor capacidad de exploración del espacio de búsqueda y una mayor robustez en la resolución de problemas complejos.
2. Otra línea de investigación se centra en el estudio de las arquitecturas de hardware paralelo y sus implicancias en el rendimiento de los modelos. En este sentido, se busca comprender cómo las características de las arquitecturas de hardware, como el número de núcleos, la memoria caché y la velocidad de la comunicación entre ellos, afectan el rendimiento de los modelos y cómo se pueden optimizar para obtener mejores resultados. Esta investigación puede tener importantes

implicancias en la selección de la mejor arquitectura de hardware para aplicaciones específicas.

3. Finalmente, otra línea de investigación está relacionada con la generación de un modelo de programación centrado en el rendimiento y la optimización de recursos. Esta investigación busca desarrollar nuevas técnicas y herramientas de programación que permitan a los desarrolladores crear aplicaciones paralelas más eficientes en términos de rendimiento y uso de recursos. De esta manera, se espera poder aprovechar al máximo el potencial de las arquitecturas de hardware paralelo y mejorar la eficiencia de los modelos y otras aplicaciones similares.

3. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la presente línea de I+D+i cuenta con la dirección del Dr. Miguel Mendez-Garabetti y la codirección del Mgter. Javier Roseinstein. Además, participan cuatro docentes investigadores por las diferentes universidades, uno de ellos con presentación en curso de plan de trabajo doctoral en el marco de la presente línea. En relación a estudiantes de grado, se cuenta con tres estudiantes de grado, y un estudiante de posgrado.

4. RESULTADOS OBTENIDOS/ESPERADOS

Existen diferentes resultados esperados relacionados a cada proyecto, los cuales confluyen al resultado esperado de la línea de I+D+i general.

Por un lado, se espera desarrollar metaheurísticas paralelas de forma adaptativa,

que permitan mejorar la calidad de la búsqueda y la eficiencia al ajustar los parámetros de acuerdo con las características específicas de cada problema. El resultado esperado es una mayor capacidad de exploración del espacio de búsqueda y una mayor capacidad de resolución de problemas complejos, lo que permitirá obtener soluciones más precisas y robustas en una variedad de aplicaciones prácticas.

En relación a la segunda línea de investigación, el resultado esperado consiste en obtener una comprensión detallada de cómo las características de las arquitecturas de hardware paralelo afectan el rendimiento y cómo se pueden optimizar para obtener mejores resultados. Además, se espera que este proyecto de investigación contribuya al desarrollo de nuevas arquitecturas de hardware que sean más eficientes y eficaces.

Y, por último, otro resultado esperado se relaciona con la propuesta de un modelo de programación centrado en el rendimiento y la optimización de recursos que crear aplicaciones paralelas más eficientes en términos de rendimiento y uso de recursos. Se espera que la investigación contribuya al desarrollo de nuevas técnicas y herramientas de programación que puedan ser utilizadas por la comunidad de desarrolladores de manera amplia y que permitan crear aplicaciones paralelas más eficientes y escalables en el futuro.

5. BIBLIOGRAFÍA

- [1] L. N. Long, "Parallel Computing", *Seminar*, p. 134, 2008.
- [2] J. J. Dongarra, *The sourcebook of parallel computing*, vol. 41. Morgan Kaufmann Publishers, 2002. doi: 10.5860/choice.41-0348.
- [3] R. Suda, K. Naono, K. Teranishi, y J. Cavazos, "Software automatic tuning: Concepts and state-of-the-art results", en *Software Automatic Tuning: From Concepts to State-of-the-Art Results*,

Springer New York, 2010, pp. 3–15. doi: 10.1007/978-1-4419-6935-4_1.

- [4] A. Grama, A. Gupta, G. Karypis, y V. Kumar, *Introduction to Parallel Computing*. Harlow, England ; New York, 2003.
- [5] B. Chapman, G. Jost, R. V. D. Pas, R. V. D. Pas, W. Gropp, y E. Lusk, *Using OpenMP – Portable Shared Memory Parallel Programming*. Cambridge, Mass: MIT Press, 2007.
- [6] S. Tsutsui y P. Collet, Eds., *Massively Parallel Evolutionary Computation on GPGPUs*. New York, 2013.
- [7] M. Aldinucci *et al.*, “Practical Parallelization of Scientific Applications with OpenMP, OpenACC and MPI”, *J. Parallel Distrib. Comput.*, vol. 157, jun. 2021, doi: 10.1016/j.jpdc.2021.05.017.
- [8] M. Méndez-Garabetti, G. Bianchini, y P. Caymes-Scutari, “HESS-IM: A Uncertainty Reduction Method that Integrates Remote Sensing Data Applied to Forest Fire Behavior Prediction”, *Commun. Comput. Inf. Sci.*, p. 17, 2021.
- [9] M. L. Tardivo, P. Caymes-Scutari, G. Bianchini, M. Méndez-Garabetti, A. Cencerrado, y A. Cortés, “A comparative study of evolutionary statistical methods for uncertainty reduction in forest fire propagation prediction”, *Procedia Comput. Sci.*, vol. 108, pp. 2018–2027, ene. 2017, doi: 10.1016/j.procs.2017.05.252.

Técnicas de Recuperación y Procesamiento de Grandes Volúmenes de Datos Sísmicos. Un Repositorio Público Basado en Serverless.

¹María Murazzo, ^{1,2}Marcelo Moreno, ¹Nelson Rodríguez, ²Ricardo Sifón, ²Valeria Nicolía, ³Ignacio Benemerito, ³Leonardo Celador

¹Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

²INPRES (Instituto Nacional de Prevención Sísmica)

³Alumno Avanzado Licenciatura en Sistemas de Información y Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

marite@unsj-cuim.edu.ar, nelson@iinfo.unsj.edu.ar, mmoreno@inpres.gob.ar, rsifon@inpres.gob.ar, vnicolia@inpres.gob.ar, ignacio.benemerito@gmail.com, leonardomiguelcelador@gmail.com

Resumen

El Instituto Nacional de Prevención Sísmica (INPRES), por Ley 19616, es el organismo público en argentina encargado, del monitoreo sísmico de todo el territorio Argentino. Como tal, es poseedor de un vasto e importante catálogo sísmico y de formas de onda, único en Argentina. Estos datos están disponibles en servidores locales, y cualquier usuario, puede solicitarlos. Cada pedido de datos supone un importante esfuerzo, ya que los requerimientos de ese tipo generalmente se refieren a grandes volúmenes de datos, y requiere de una también muy importante capacidad de cómputo, los que se deben realizar y satisfacer desde el mismo Centro de Datos dedicado a tareas de rutina. Dada esta problemática es que se plantea el uso de serverless computing con el objeto de usar una infraestructura cloud pública para alojar los datos y FaaS para implementar las técnicas de manipulación de los datos

Palabras clave: *serverless computing, cloud computing, grandes volúmenes de datos*

Contexto

El presente trabajo se encuadra dentro del área de I/D Sistemas Distribuidos y Paralelos y es una de las líneas de investigación internas, del proyecto: Soporte Serverless para aplicaciones móviles de nueva generación, cuya propuesta está en etapa de evaluación para el período 2023-2024. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil y distribuida desde hace más de 22 años.

Como continuación del proyecto anterior: Computación Serverless para tratamiento de datos provenientes de dispositivos de IoT, se continúa el trabajo con investigadores de otras universidades, lo cual favorece notablemente a todas las instituciones participantes.

Introducción

El Instituto Nacional de Prevención Sísmica (INPRES), por Ley 19616, es el organismo público en argentina encargado, entre otras tareas, del monitoreo sísmico de todo el territorio Argentino. Como tal, es poseedor de un vasto e importante catálogo sísmico y de formas de onda, único en Argentina. Todos los datos están disponibles en servidores locales de INPRES, y cualquier usuario, generalmente de organismos científicos, puede solicitar sus datos. Cada pedido de datos supone un importante esfuerzo, ya que los requerimientos de ese tipo generalmente se refieren a grandes volúmenes de datos, y requiere de una también muy importante capacidad de cómputo, los que se deben realizar y satisfacer desde el mismo Centro de Datos dedicado a tareas de rutina.

Además, dos factores de reciente irrupción, han aumentado la carga de trabajo del Centro de Datos de INPRES. Primero, la Ley 26.899 de creación de repositorios digitales institucionales de acceso abierto, que fue establecida en 2013, y que representa un enorme aporte para el avance del conocimiento científico-tecnológico en nuestro país, pero que a la vez, ha sumado presión a las capacidades de INPRES de poder responder a todos los requerimientos que la Ley ha

generado. En segundo lugar, los avances en la ciencia de datos ha alimentado el interés de terceros en el análisis de grandes conjuntos de datos Sísmicos. Temas tales como detección de eventos en volúmenes históricos, Machine Learning, estudios de réplicas, reproceso o refinamiento de localizaciones, etc. Todos ellos requieren disponibilidad de datos de forma de onda continua que abarcan meses o años e incluyen registros de decenas de estaciones.

En INPRES los datos de forma de onda se almacenan en un formato estándar denominado miniSEED, que es a la vez, un subconjunto del estándar para el intercambio de datos de terremotos (Standard for the Exchange of Earthquake Data - SEED), que es un formato de datos destinado al archivo e intercambio de datos de series temporales sismológicas y metadatos relacionados. El formato está definido por la Federación Internacional de Redes de Sismógrafos Digitales (FDSN) y está documentado en el Manual de SEED [1]. En miniSEED se incluyen metadatos muy limitados para la serie temporal, más allá de la identificación de la serie temporal. En particular, no se incluyen las coordenadas geográficas, la información de respuesta instrumental y otra información necesaria para interpretar los datos.

Las series de tiempo se almacenan como registros de datos de longitud fija independientes, cada uno de los cuales contiene un pequeño segmento de valores de series contiguos. Las longitudes de registro comunes son 512 bytes (para flujos en tiempo real) y 4096 bytes (para archivo). Además, existen numerosas librerías de código abierto para facilitar la lectura y escritura de datos miniSEED sin necesidad de conocer los detalles del formato.

La Red Nacional de Estaciones Sismológicas (RNES) [2] de INPRES se compone de cincuenta y cinco (55) estaciones sismológicas, de uno y tres canales cada una, transmitiendo en tiempo real al Centro de Datos de INPRES. Cada registro de una hora de duración para una estación típica a 100 muestras por segundo, tiene un tamaño aproximado de 3 Megabyte, con lo cual un día

completo representan 72 Megabytes, y 26 Gigabyte al año, para cada canal. Si la estación registra para tres canales (Vertical, Norte-Sur y Este-Oeste), se tiene que calcular que cada estación sísmica genera aproximadamente 80 Gigabyte de datos al año. Para las 55 estaciones hace aproximadamente 1.5 Terabytes anuales.

El INPRES comenzó sistemáticamente a registrar sus datos de onda en formato miniSEED desde el año 2010 con un comienzo de 20 estaciones, hasta llegar a las 65 actuales. El tamaño total de datos de forma de onda en formato miniSEED en INPRES asciende a 12 Terabytes. Muchos de esos datos están almacenados en dispositivos offline. La siguiente tabla muestra la evolución de estaciones y cantidad de datos almacenados en INPRES a febrero de 2023.

Año	Nº de Estaciones	Terabytes/Año
2010	20	0,52
2011	22	0,572
2012	19	0,494
2013	22	0,572
2014	20	0,52
2015	25	0,65
2016	30	0,78
2017	30	0,78
2018	38	0,988
2019	42	1,092
2020	42	1,092
2021	47	1,222
2022	50	1,3
2023	55	1,43
Tamaño total (en TB)		12,012

Infraestructura de soporte

En función de lo expresado en párrafos anteriores, se hace necesario contar con una solución que provea el acceso a los datos y al análisis de los mismos en forma eficiente, con mínimo esfuerzo por parte del usuario y con escasa administración y gestión de la plataforma de soporte.

El fenómeno de los grandes volúmenes de datos no es nuevo y ha sido abordado en forma transversal por múltiples áreas del

conocimiento. El tratamiento de estos datos en área de la sismología y en particular la problemática que la presente línea de investigación abordara confluye con cloud computing.

Cloud Computing se ha caracterizado por ser una tecnología centrada en ofrecer cómputo bajo demanda como cualquier otro servicio. Esto es una ventaja para montar aplicaciones donde es necesario el procesamiento intensivo, tales como aquellas aplicaciones que procesen y extraigan información de grandes volúmenes de datos [1].

Los proveedores cloud alegan muchas ventajas en la migración de aplicaciones para el tratamiento de grandes volúmenes de datos, como el acceso rápido a los recursos, costos más bajos y flexibilidad en la contratación y el aprovisionamiento de recursos. Un punto adicional es la seguridad, la cual, afirman es de alto nivel y en muchos casos muy difícil de implementar en la mayoría de los laboratorios, ya que tener personal de TI especializado en seguridad no es común.

Un aspecto más que lleva a la adopción del cloud como plataforma de despliegue de aplicaciones para grandes volúmenes de datos es, mejorar la colaboración científica, es decir, facilitar la investigación colaborativa y la innovación; este aspecto ha sido el foco de este proyecto desde hace varios años al incorporar investigadores de otras universidades del país.

Otro tema es el potencial ahorro de tiempo que ofrece el cloud a los usuarios finales. Estos usuarios no necesitan preocuparse por actualizaciones de software, compatibilidad o parches de seguridad, pues todo esto es aprovisionado de forma transparente [2].

Sin embargo, el cloud tiene dos grandes desventajas, la primera es la degradación de la performance de las aplicaciones al montarlas sobre arquitectura virtualizada, debido a que genera overhead en la contextualización de las máquinas virtuales; la segunda desventaja cuando se despliegan aplicaciones en el cloud, es que es responsabilidad de la organización mantener funcionando de forma correcta la infraestructura que se necesite para el despliegue de las aplicaciones, lo cual lleva a

cargar costos sobre el presupuesto para su mantenimiento y soporte [3].

A pesar de esto, en los últimos años, las necesidades de rendimiento y escalabilidad para el procesamiento de grandes volúmenes de datos se han realizado en forma exitosa mediante el uso de frameworks de infraestructuras distribuidas aprovisionados en el cloud. Sin embargo, los costos de personal operativo y los de infraestructura presentan factores económicos significativos para el procesamiento de datos a escala. Además, el procesamiento de grandes volúmenes de datos en el cloud requiere un amplio conocimiento para definir y ejecutar trabajos y para implementar, configurar y mantener la plataforma de infraestructura requerida para su ejecución. [4]

En este sentido, la aparición del Serverless Computing [5] y su core Function-as-a-Service (FaaS), logra que los desarrolladores no tengan que preocupar por el aprovisionamiento y escalado de la infraestructura, por lo que se pueden centrar en la lógica de sus aplicaciones. De esta forma es posible lograr la abstracción de la gestión de servidores (aprovisionamiento, configuración, escalado, etc.) para que los usuarios, en este caso desarrolladores, puedan enfocarse en la lógica de sus aplicaciones [6]. Esto se debe a que las funciones no cuentan con ningún tipo de dependencia de la infraestructura, por lo que se puede cargar en el entorno cloud y ser ejecutadas en cualquier momento. Una de las grandes ventajas de FaaS es que el negocio únicamente se preocupa de la lógica y desarrollo de la funcionalidad de las aplicaciones, mientras que la ejecución, seguridad y mantenimiento de la aplicación es tarea del proveedor cloud [7].

Líneas de Investigación, Desarrollo e Innovación

La presente línea de investigación se enmarca dentro del proyecto mencionado en el contexto y surge como resultado de la cooperación obtenida con el INPRES gracias a que un docente investigador perteneciente al grupo de trabajo desempeña tareas en dicha

institución. Además, se contara con la colaboración de dos profesionales del INPRES

Gracias a esta cooperación el grupo de trabajo contará con una importante fuente de datos para poder realizar las tareas necesarios para trabajar con un back end serverless.

Es por ello que la metodología a seguir será experimental deductiva, lo cual permitirá analizar cómo se comportan las aplicaciones en diferentes entornos de ejecución.

Objetivos

Los objetivos del grupo de investigación en esta línea de conocimiento son los siguientes: Desde el punto de vista institucional:

- Proveer un entorno en el cloud que permita dar cumplimiento con la Ley 26.899 de Creación de Repositorios Digitales Institucionales de Acceso Abierto, para los datos de forma de onda sísmica disponibles en INPRES.
- Proveer alta disponibilidad y escalabilidad de datos de forma de ondas sísmicas en la nube, que facilite investigaciones que utilizan procesamiento computacionalmente intensivo y grandes volúmenes de datos sísmicos.

Desde el punto de vista científico:

- Analizar la factibilidad de manipular grandes volúmenes de datos con FaaS.
- Estudiar las técnicas de analítica de datos sobre FaaS
- Evaluar la performance de serverless cuando procesa grandes volúmenes de datos.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de seis investigadores que figuran en este trabajo de las universidades Nacional de San Juan y un alumno de grado. Además, el proyecto marco donde se está desarrollando esta propuesta ha establecido vínculos con investigadores de la Nacional de San Luis, de la Universidad Champagnat y de la Universidad Nacional de Salta y dos alumnos de grado.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis

de maestría en áreas afines y dos tesinas de grado en el área de Serverless computing, Concurrencia y Computación distribuida. Además se espera aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.

Referencias

- [1] Bahrami, M., & Singhal, M. (2015). The role of cloud computing architecture in big data. *Information granularity, big data, and computational intelligence*, 275-295.
- [2] Amin, R., Vadlamudi, S., & Rahaman, M. M. (2021). Opportunities and challenges of data migration in cloud. *Engineering International*, 9(1), 41-50.
- [3] Sarmah, S. S. (2019). Cloud Migration-Risks and Solutions. *Science and Technology*, 9(1), 7-11.
- [4] Liu, X., Liu, Y., & Fei, Y. (2022). Computer big data analysis and cloud computing network technology. In *The 2021 International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy: SPIoT-2021 Volume 1* (pp. 517-522). Springer International Publishing.
- [5] Eismann, S., Scheuner, J., Van Eyk, E., Schwinger, M., Grohmann, J., Herbst, N., ... & Iosup, A. (2020). A review of serverless use cases and their characteristics. *arXiv preprint arXiv:2008.11110*.
- [6] Añel, J. A., Añel, J. A., Montes, D. P., Iglesias, J. R., & Romano. (2020). *Cloud and Serverless Computing for Scientists*. Springer International Publishing.
- [7] Fox, G. C., Ishakian, V., Muthusamy, V., & Slominski, A. (2017). Status of serverless computing and function-as-a-service (FaaS) in industry and research. *arXiv preprint arXiv:1708.08028*.

Aplicaciones para la salud y métricas de rendimiento en sistemas paralelos y distribuidos

Javier Balladini¹, Marina Morán¹, Claudio Zanellato¹, Rodrigo Cañibano¹, Belén Casanova¹, Mariano Conchillo¹, Cristina Orlandi², Enzo Rucci³, Armando De Giusti³, Remo Suppi⁴, Dolores Rexachs⁴, Emilio Luque⁴, Emmanuel Frati⁵

¹ Facultad de Informática, Universidad Nacional del Comahue
{javier.balladini, marina, claudio.zanellato, rcanibano}@fi.uncoma.edu.ar
{maria.casanova, mariano.conchillo@est.fi.uncoma.edu.ar}

² Hospital Francisco Lopez Lima - orlandi.mariacristina@gmail.com

³ Instituto de Investigación en Informática LIDI,
Universidad Nacional de La Plata - erucci@lidi.info.unlp.edu.ar, degiusti@lidi.info.unlp.edu.ar

⁴ Departamento de Arquitectura de Computadores y Sistemas Operativos,
Universidad Autónoma de Barcelona - {remo.suppi, dolores.rexachs, emilio.luque}@uab.es

⁵ Universidad Nacional de Chilecito, fefrati@undec.edu.ar

Resumen

Los grandes avances tecnológicos de los sistemas de cómputo paralelo y distribuido hacen viable nuevas soluciones a problemas. Por un lado, nos enfocamos en la mejora de métricas relacionadas al consumo de energía durante los periodos de recuperación de fallos en sistemas de cómputo de altas prestaciones de tipo cluster. Por otro lado, buscamos aplicar técnicas de cómputo paralelo y distribuido para brindar soluciones en el sector salud. En especial, nos orientamos a sistemas de alertas tempranas de gravedad. Una aplicación está destinada a Unidades de Cuidados Intensivos, que debe tratar con datos masivos en tiempo real, y otra aplicación tiene como fin la clasificación del nivel de gravedad de pacientes en salas de internación general, que presenta una arquitectura distribuida, resiliente y de simple administración. Los trabajos se desarrollan en colaboración con otras universidades, y un hospital público de Argentina. La formación de recursos humanos en estas líneas está orientada al nivel de tecnicatura, grado, maestría y doctoral.

Palabras claves: *computación de altas prestaciones, eficiencia energética, tolerancia a fallos, salud, big data, inteligencia artificial.*

Contexto

Las líneas de investigación aquí presentadas están enmarcadas dentro del proyecto de investigación 04/F017 "Cómputo paralelo y distribuido: métricas de rendimiento, aplicaciones de big data e inteligencia artificial", financiado por la Universidad Nacional del Comahue (UNComa), con inicio el 01/01/2021 y finalización el 31/12/2024, acreditado por el Ministerio de Educación de Argentina.

La mayoría de los temas de investigación surgieron en proyectos anteriores acreditados. El eje de aplicaciones para la salud se desarrolla en colaboración con el Hospital Francisco López Lima de General Roca (Río Negro) y con interés del Ministerio de Salud de la Provincia de Río Negro. El eje de métricas de rendimiento de los sistemas de cómputo paralelo, se desarrolla en colaboración con el Instituto de Investigación en Informática LIDI de la Universidad Nacional de La Plata, el grupo de investigación "High Performance Computing for Efficient Applications and Simulation" de la Universidad Autónoma de Barcelona, España, y con la Universidad Nacional de Chilecito, La Rioja.

1. Introducción

Numerosas aplicaciones informáticas están surgiendo gracias a la mejora y reducción de costos de las tecnologías de cómputo paralelo y distribuido. Nos interesamos principalmente en el sector de salud, un campo de gran impacto social, que puede verse altamente beneficiado con soluciones de software basadas en estas tecnologías.

A su vez, nos enfocamos en las métricas de eficiencia computacional y energética de los sistemas de cómputo paralelo y distribuido. Desde los inicios de la computación las métricas de interés se relacionaban únicamente con el rendimiento computacional. Sin embargo, desde hace algunos años, las métricas de rendimiento energéticas en muchos casos las han superado en relevancia. El consumo energético es muy importante ya que produce fuertes impactos en lo económico, medioambiental y social.

A continuación se detallan la motivación y problemática específica de cada línea de investigación desarrollada.

1.1 Aplicaciones para la salud

Unidad de Cuidados Intensivos

Una Unidad de Cuidados Intensivos (UCI) provee cuidados continuos y rigurosos a personas adultas críticamente enfermas que pueden beneficiarse de tratamiento, y da un buen morir a pacientes irrecuperables. Los datos de los pacientes involucran datos clínicos de baja frecuencia y flujos de datos fisiológicos de alta frecuencia generados por el equipamiento médico. En una UCI típica, los enfermeros completan manualmente datos en formularios, registrando datos clínicos y fisiológicos. Los datos fisiológicos se obtienen por observación de las pantallas del equipamiento médico a intervalos de tiempo que definen los médicos para cada paciente. El equipamiento médico emite alertas cuando hay riesgo en la salud del paciente, basándose en mediciones de

parámetros. Luego, los médicos analizan los datos de los formularios y dan a los enfermeros indicaciones de tratamientos.

Los principales problemas que ocurren en las UCIs típicas son: (a) gestión de la información es proclive a errores humanos, (b) pérdida de datos entre registros de enfermería, (c) detección tardía del deterioro de la salud de los pacientes. Las causas tienen origen en el tratamiento manual de la información (lento y propenso a errores), y cuyos efectos incluyen diagnósticos imprecisos o incorrectos o retrasados, inconsistencias en la información, menos datos para investigaciones médicas, y mayores necesidades de recursos humanos. En la mayoría de las UCIs del país, no hay médicos intensivistas durante las 24hs, y hay sobrecarga laboral, impactando negativamente en la salud de los pacientes.

Las soluciones se orientan a sistemas expertos de detección automática, temprana y progresiva del deterioro de pacientes, tal como los propuestos en [1, 2, 3].

Detección temprana de severidad en pacientes en internación general

La pandemia del SARS-COV-2 ha puesto de manifiesto la falta de preparación del sistema sanitario para afrontar este tipo de situaciones. Los centros de salud se vieron desbordados por el aumento repentino de la demanda. El recurso humano especializado se hizo escaso, agravado por contagios del propio personal, lo que provocó que dejaran de prestar servicios temporalmente.

En este contexto, la metodología tradicional de médicos y enfermeras debe optimizarse para permitir la atención de un mayor número de pacientes sin reducir la calidad de la atención. Para organizar y optimizar los recursos es posible utilizar el método de triaje, método que consiste en la selección y clasificación de pacientes en diferentes niveles de gravedad para lograr una correcta asignación de los cuidados. El triaje se aplicó tempranamente en China [4] durante la

emergencia sanitaria de pacientes con COVID-19 mediante la adaptación de uno de los sistemas de alerta temprana más conocidos y validados. El método mejora el rendimiento del personal, reduciendo los controles en pacientes estables, aumentándolos en pacientes graves y reduciendo la mortalidad hospitalaria inesperada en áreas de internación general.

Si bien el triaje se puede realizar mediante cálculos manuales, el proceso sería propenso a errores y aumentaría considerablemente la carga de trabajo del personal de salud, lo que limitaría el beneficio potencial del método. Para mejorar la calidad asistencial, nuestro grupo de investigación propone un sistema distribuido y resiliente que realice un triaje automático.

1.2 Métricas de rendimiento - eficiencia energética en HPC

La computación de alto rendimiento (HPC) sigue aumentando su rendimiento computacional y su eficiencia energética. Por ejemplo, el superordenador Frontier, la supercomputadora más rápida del mundo en el ranking Top500, presenta 1.102 PFlops y 52 GFlops/watts frente a los 442 PFlops y 15 GFlops/watt de su predecesor, Fugaku. Sin embargo, aunque la eficiencia energética aumenta, también se ha ido incrementando la demanda total de potencia debido al aumento del tamaño de las máquinas. En particular, Fugaku demanda casi 30 MW, aproximadamente lo mismo que consume una ciudad de 400.000 habitantes.

En estas máquinas, también se está aumentando en gran medida el número de unidades de procesamiento, lo que causa la disminución de la confiabilidad del sistema. Así, el consumo energético y la tolerancia a fallos se han identificado como los dos mayores desafíos a enfrentar en la actual era exaescala [5, 6, 7, 8], y es en estos dos desafíos que nos centramos.

2. Líneas de investigación

El eje central de nuestra investigación es desarrollar metodologías, modelos y soluciones informáticas para colaborar en la resolución de problemas de cómputo paralelo y distribuido, con alto impacto social en el campo de la salud, y en la reducción del consumo energético de sistemas de HPC.

2.1 Aplicaciones para la salud

UCI: Nuestro sistema intenta emular el comportamiento de un médico intensivista experto, dando recomendaciones para la toma de decisiones clínicas, con el objetivo de reducir la incertidumbre sobre el diagnóstico, las opciones de tratamiento y el pronóstico. La solución requiere la aplicación de técnicas de cómputo paralelo y distribuido para procesamiento en tiempo real de algoritmos de inteligencia artificial sobre grandes volúmenes de datos. A diferencia de otras alternativas, nuestro objetivo se orienta a la construcción de un sistema multihospitalario (con el fin de incrementar el volumen de datos y consecuente extracción de conocimiento) con soporte de telemedicina, y que integre componentes de software libre maduros, que nosotros optimizamos para nuestro dominio.

Detección temprana de severidad en pacientes en internación general: Nuestro objetivo es construir un sistema distribuido sobre dispositivos móviles, que implemente modelos computacionales de alertas tempranas de diferentes enfermedades, inicialmente basados en el conocimiento de personal experto en salud, y posteriormente mejorado mediante técnicas de aprendizaje automático. La complejidad de la arquitectura del sistema está en lograr un sistema seguro y resiliente (el sistema continúa brindando servicios en caso de fallos en la red y los dispositivos), y simple de implementar y mantener (no requiere el uso de servidores estándar en los hospitales).

2.2 Métricas de rendimiento - eficiencia energética en HPC

Nos centramos en el desarrollo de metodologías, modelos y construcción de software para administrar y gestionar el consumo de energía y prestaciones computacionales de sistemas de cómputo paralelo. Actualmente, nuestro principal objetivo es la gestión energética en mecanismos de tolerancia a fallos basados en checkpoints.

3 Resultados y objetivos

3.1 Aplicaciones para la salud

UCI: En [9] presentamos un análisis del estado general de la UCI del hospital Francisco López Lima, y la propuesta del diseño de alto nivel del sistema. En [10] presentamos el diseño de la infraestructura del sistema de procesamiento de reglas clínicas y un prototipo. En [11] propusimos la herramienta como mejora de la prestación de servicios integrados de cuidados intensivos de la salud. En [12] presentamos una optimización de la administración de datos de curvas fisiológicas. En el último año se terminó un software para análisis eficiente de la curva de electrocardiograma. Próximos objetivos: encontrar patologías en el electrocardiograma, continuar con el desarrollo de hardware y software para extraer datos del equipamiento médico, manejo de errores en datos fisiológicos, optimización de la infraestructura de datos masivos de tiempo real para el procesamiento de reglas clínicas, y el desarrollo de una aplicación para la interacción del sistema con médicos y enfermeros.

Detección temprana de severidad en salas de internación general: Se desarrolló una versión básica del sistema. Su diseño fue pensado para producir un desarrollo rápido, por lo que posteriormente comenzamos a trabajar en un nuevo diseño para superar limitaciones en aspectos de seguridad, y para permitir la extensión de su

funcionalidad, en especial para el soporte de nuevas enfermedades. Parte del nuevo diseño fue publicado en [13]. Se realizaron pruebas de concepto del nuevo diseño, que resta publicar.

Próximos objetivos: completar el nuevo diseño e implementarlo.

3.2 Métricas de rendimiento - consumo energético en HPC

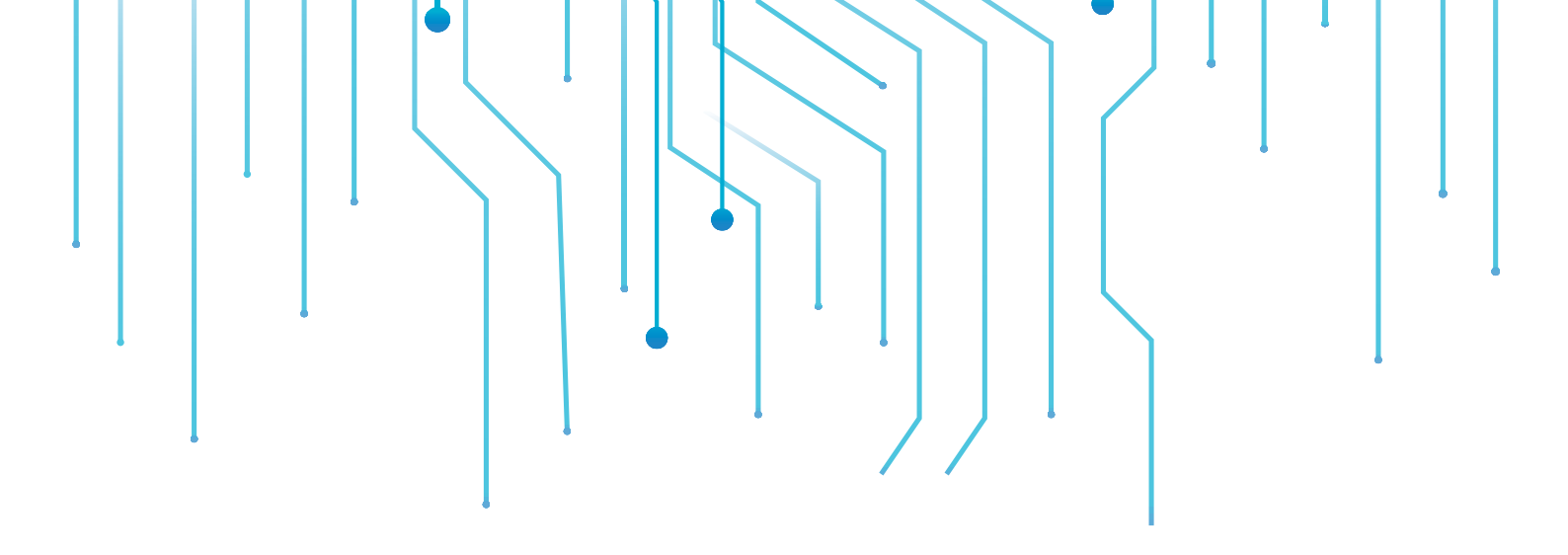
Los métodos de tolerancia a fallos tienen fuerte incidencia en el consumo energético de los sistemas de HPC, y resulta de suma importancia conocer, antes de ejecutar una cierta aplicación, el impacto que pueden producir los diferentes métodos y configuraciones del mismo. En [14], presentamos una metodología para predecir el consumo energético producido por el método de checkpoint coordinado remoto, y en [15] expusimos un análisis de los factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. En [16] propusimos un modelo para estimar el consumo energético de operaciones de checkpoint y restart, y un método para su construcción. En [17] presentamos estrategias para checkpoints no coordinados que, al momento de un fallo de un nodo, permiten gestionar y reducir el consumo energético de los nodos que no han fallado; se construyó también un modelo energético y un simulador que permite evaluar las estrategias. Se completó la funcionalidad con operaciones no bloqueantes y esperas en cascada causadas por dependencias entre procesos [18]. Próximos objetivos: desarrollar un planificador para un sistema GNU/Linux que implemente el modelo.

4 Formación de recursos humanos

El equipo de trabajo de la Universidad Nacional del Comahue tiene un Doctor y un Magíster, una estudiante de Doctorado en etapa de escritura de la tesis, y un estudiante de Maestría. En 2021 se finalizó una tesis de grado sobre el procesamiento eficiente de señales de electrocardiograma.

Referencias

- [1] S. Balaji, M. Patil, and C. McGregor, "A Cloud Based Big Data Based Online Health Analytics for Rural NICUs and PICUs in India: Opportunities and Challenges", in 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS) (2017), pp. 385-390.
- [2] Jagreet Kaur and Dr. Kulwinder Singh Mann, "AI based HealthCare Platform for Real Time, Predictive and Prescriptive Analytics using Reactive Programming", Journal of Physics: Conference Series 933 (2018), pp. 012010.
- [3] Fernando López-Martínez, Edward Rolando Núñez-Valdez, Vicente García-Díaz, and Zoran Bursac, "A Case Study for a Big Data and Machine Learning Platform to Improve Medical Decision Support in Population Health Management", Algorithms 13, 4 (2020).
- [4] Xuelian Liao, Bo Wang, and Yan Kang, "Novel coronavirus infection during the 2019--2020 epidemic: preparing intensive care units-the experience in Sichuan Province, ...", Intensive care medicine 46, 2 (2020), pp. 357--360.
- [5] Aurelien Bouteiller, Franck Cappello, Jack Dongarra, Amina Guermouche, Thomas Hérault, and Yves Robert, "Multi-criteria checkpointing strategies: Response-time versus resource utilization", in European Conference on Parallel Processing (2013), pp. 420--431.
- [6] Mohammed El Mehdi Diouri, Olivier Glück, Laurent Lefèvre, and Franck Cappello, "Ecofit: A framework to estimate energy consumption of fault tolerance protocols for hpc applications", in Proceedings of the 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (2013), pp. 522--529.
- [7] Yongpeng Liu and Hong Zhu, "A survey of the research on power management techniques for high-performance systems", Software: Practice and Experience 40, 11 (2010), pp. 943--964.
- [8] Franck Cappello, Al Geist, William Gropp, Sanjay Kale, Bill Kramer, and Marc Snir, "Toward Exascale Resilience: 2014 update", Supercomputing Frontiers and Innovations 1, 1 (2014).
- [9] Javier Balladini, Claudia Rozas, Emmanuel Frati, Nestor Vicente, and Cristina Orlandi, "Big Data Analytics in Intensive Care Units: challenges and applicability in an Argentinian Hospital", Computer Science and Technology (JCST) (2015).
- [10] Javier Balladini, Pablo Bruno, Rafael Zurita, and Cristina Orlandi, "An Automatic and early detection of the deterioration of patients in Intensive and Intermediate Care Units", Journal of Computer Science and Technology 18, 03 (2018), pp. e25.
- [11] Javier Balladini, Pablo Bruno, Rafael Zurita, Cristina Orlandi, Remo Suppi, Dolores Rexachs, and Emilio Luque, "A tool for improving the delivery of integrated intensive health care performance.", International Journal of Integrated Care 19, 4 (2019).
- [12] Rodrigo Cañibano, Claudia Rozas, Cristina Orlandi, and Javier Balladini, "Data Management Optimization in a Real-Time Big Data Analysis System for Intensive Care", in Conference on Cloud Computing, Big Data & Emerging Topics (2020), pp. 93--107.
- [13] Cañibano, R. S., Castagno, S., Conchillo, M., Chiarotto, G., Rozas, C., Zanellato, C., ... & Balladini, J. (2022). Towards a resilient e-health system for monitoring and early detection of severity in hospitalized patients during a pandemic. In X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022).
- [14] Javier Balladini, Marina Morán, Dolores Rexachs, and Emilio Luque, "Metodología para predecir el consumo energético de checkpoints en sistemas de HPC", XX Congreso Argentino de Ciencias de la Computación (CACIC 2014) (2014).
- [15] Marina Morán, Javier Balladini, Dolores Rexachs, and Emilio Luque, "Factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters", XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018) (2018).
- [16] M. Morán, J. Balladini, D. Rexachs, and E. Luque, "Prediction of Energy Consumption by Checkpoint/Restart in HPC", IEEE Access 7 (2019), pp. 71791-71803.
- [17] Marina Morán, Javier Balladini, Dolores Rexachs, and Enzo Rucci, "Towards Management of Energy Consumption in HPC Systems with Fault Tolerance", in 2020 IEEE Congreso Bial de Argentina (ARGENCON) (2020), pp. 1-8.
- [18] Morán, M., Balladini, J., Rexachs del Rosario, D., & Rucci, E. (2022). Some Issues to Consider in the Management of Energy Consumption in HPC Systems with Fault Tolerance. In X Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 2022).



PSSTR

Procesamiento de Señales y Sistemas en Tiempo Real

Avances en el desarrollo de un prototipo experimental de segmento terreno Satelital multiplataforma-multimisión UGS UNLaM

Rocío B. Fernández, Pablo Ferreira, Camila J. Forestiero, Santiago Mansfeld

Pablo Soligo , Jorge S. Ierache , Martin Becerra, Diego Sanz

Grupo de Investigación y Desarrollo de Software Aeroespacial (GIDSA)
Departamento Ingeniería e Investigaciones Tecnológicas (DIIT)
Universidad Nacional de la Matanza (UNLaM)

rociobefernandez@alumno.unlam.edu.ar; pablferreira@alumno.unlam.edu.ar;
cforestiero@alumno.unlam.edu.ar; smansfeldgimenez@alumno.unlam.edu.ar;
psoligo@unlam.edu.ar; jierache@unlam.edu.ar

RESUMEN

En este trabajo se presentan los avances en el desarrollo de un prototipo de segmento terreno multiplataforma-multimisión. Esto incluye el desarrollo de una estación terrena para la recepción de datos satelitales, su interpretación, almacenamiento y publicación.

Palabras clave: *Segmento Terreno, ...*

CONTEXTO

El desarrollo de un prototipo de segmento terreno multiplataforma-multimisión, denominado UGS (Unlam Ground Segment) [1], se encuentra enmarcado en el proyecto “PROINCE C-245: Estación Terrena Satelital Experimental UNLaM” llevado adelante por el Grupo de Investigación y Desarrollo de Software Aeroespacial de la Universidad Nacional de La Matanza (GIDSA) [2]. El GIDSA tiene como objetivo investigar e implementar prototipos de software de bajo costo basados en tecnologías ampliamente aceptadas, de probada madurez y con penetración en la industria de software de propósito general.

1. INTRODUCCIÓN

En trabajos previos [3] [4] [5] se evaluaron estrategias de desarrollo de un prototipo de segmento terreno basado exclusivamente en componentes de software de los denominados “de estantería”, con el objetivo de disminuir costos de desarrollo y mantenimiento.

Como resultado de estos desarrollos, el UGS utiliza en la actualidad:

- Un lenguaje de propósito general para el procesamiento de telemetría como para la generación de scripts de comandos.
- Un RDBMS para la definición de datos.
- Interfaces basadas completamente en HTTP/HTTPS.

Actualmente el UGS permite procesar datos crudos con independencia del satélite usando técnicas de reflexión de software. La mayoría de los datos son descargados de la red SatNOGS [6] aunque se han realizado pruebas prototipo logrando descargar y procesar balizas de satélites utilizando equipamiento propio. Este proceso ha sido manual, incluyendo el apuntamiento de la antena. Para poder descargar y visualizar los datos se necesita automatizar la estación terrena, incluyendo un rotor que permita el apuntamiento de la antena y paralelamente implementar cambios en las unidades de software que permitan distribuir y

visualizar las novedades en tiempo real. En el presente artículo se presentan resultados preliminares de estos desarrollos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de trabajo se centran en: a) Desarrollo estación terrena, mecanismos de guiado y software de control; que permita establecer una estación de radio diseñada para establecer comunicaciones con fuentes no terrestres: satélites, etc. Su diseño varía de acuerdo con la aplicación de esta, pero todas deben incluir necesariamente un procesador, equipos de radio, y antenas. Para evitar pérdidas de potencia en la recepción, cuando se usan antenas direccionales, es importante que exista una alineación constante entre la dirección de máxima ganancia de la antena y el satélite. De esto surge la necesidad de contar con un sistema capaz de seguir la trayectoria del satélite mientras este se encuentra visible. Pueden apuntarse las antenas manualmente, o incorporar sistemas con rotores. Esto último refiere a implementar un sistema de motores sobre el que se pueda montar la antena. Para el diseño de la estación terrena del UGS se siguen los lineamientos de la comunidad SatNOGS [6]. En la guía oficial se dictan algunos componentes necesarios y sugerencias sobre el software a usar en el controlador, dejando algunas libertades a la implementación. La versión 3 del rotor SatNOGS ofrece numerosas ventajas: bajo costo, ligereza, portabilidad, rigidez, durabilidad, facilidad tanto para construir como para reparar, blindaje electromagnético (reduciendo el ruido en la recepción) y alta precisión. b) Distribución y visualización de telemetría satelital: La capa de visualización del UGS está basada en el NASA OpenMCT, un *framework* de visualización de control de misión de código abierto de la NASA [8]. OpenMCT provee al desarrollador tableros predeterminados, almacenamiento local en el navegador y un módulo de complementos. Mediante la instalación de *plugins*, el entorno web UGS permite a su vez almacenamiento remoto a partir del inicio de sesión de usuarios oficiales, posibilitando el acceso en otros dispositivos de los *dashboards* u objetos (tales como relojes) creados.

El UGS actualmente dispone de un servicio web que permite recibir telemetría histórica de múltiples fuentes, tanto sea de una estación terrena, mediante adaptadores que consumen u obtienen datos crudos de internet, o directamente desde archivos compartidos por agencias espaciales vinculadas. Actualmente este servicio permite visualizar, de manera pública, telemetría satelital histórica procesada desde datos crudos provenientes principalmente de la red SatNOGS [7] y otras fuentes alternativas. Recientemente, se evaluaron implementaciones basadas en el uso de websockets para la comunicación de telemetría en tiempo real sobre la capa de visualización del UGS. La utilización de websockets no es obligatoria, pero está explícitamente recomendada por el equipo de desarrollo de OpenMCT.

Los websockets ofrecen una solución para establecer una conexión bidireccional entre un servidor y un cliente sobre internet. El servidor es responsable de administrar las conexiones y de decidir si debe informar a los clientes las novedades a las cuales están subscriptos. El cliente puede ser una instancia de OpenMCT la cual deberá suscribirse a las variables que se deseen visualizar. La Figura 1 muestra un dashboard ejemplo con datos del satélite de Satellogic BugSat1 (Tita). Accesible en

<https://ugs.unlam.edu.ar/#/browse/mine/demodash>

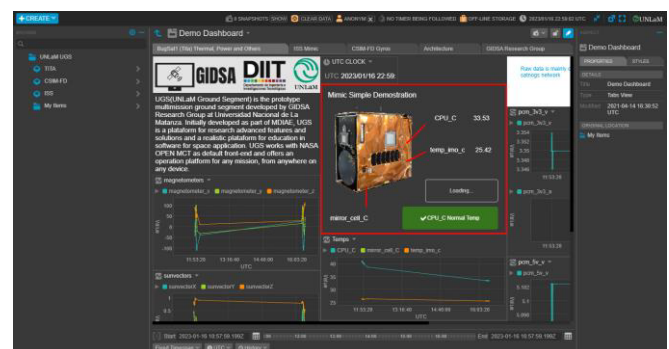


Figura 1 – UGS/NASA OpenMCT Front-end

3. RESULTADOS OBTENIDOS/ESPERADOS

Se realizaron pruebas de estrés sobre el back-end del UGS, en su última versión, utilizando telemetría simulada y múltiples clientes subscriptos a las distintas variables de telemetría de un único simulador de satélite.

Durante las pruebas se detectaron demoras importantes, se observó que los principales problemas radican la serialización, deserialización y filtrado de novedades correspondientes a cada cliente. Las limitaciones impuestas por el GIL (Global Interpreter Lock) [9] en el uso de hilos obligan a desarrollar procesos separados si se quieren aprovechar los múltiples núcleos disponibles. Estos procesos deben comunicarse mediante datos serializados lo que agrega un enorme costo al procesamiento ya que el back-end no solo debe serializar/deserializar para comunicarse con cada cliente, sino que también para procesar y filtrar las variables de telemetría en función de las subscripciones.

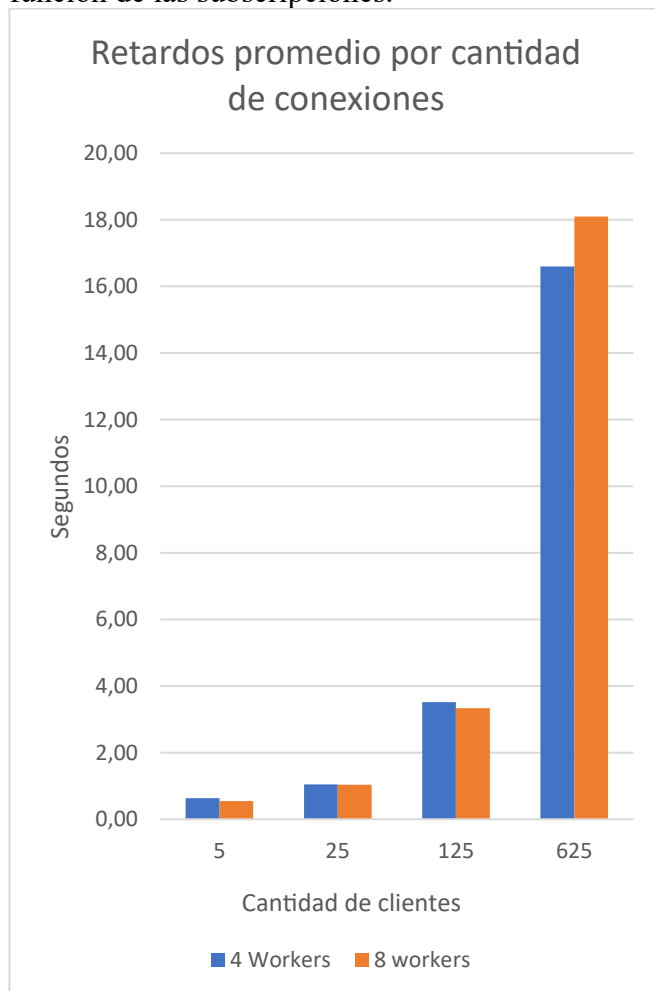


Figura 2 – Tiempos de retardo promedio según cantidad de conexiones clientes

El uso del tipo integrador “set” demostró mejor rendimiento al momento de determinar la intersección entre las novedades recibidas y las que

deben ser informadas a cada cliente, aunque los resultados (ver Figura 2 – Tiempos de retardo promedio según cantidad de conexiones clientes) no permiten pensar en un sistema escalable de bajo costo. Con respecto a la serialización y deserialización de los datos se están explorando alternativas como reducir el tamaño del paquete de información a serializar y los filtrados, reduciendo la carga de trabajo. Así como también el uso de bibliotecas optimizadas para esta finalidad (serialización/deserialización).

En cuanto a la estación terrena, esta se encuentra en la fase de ensamblado mecánico, siguiendo los lineamientos ya mencionados de SatNOGS. El único cambio destacable respecto a la implementación por defecto de SatNOGS es la utilización de una placa Arduino UNO junto con una placa CNC Shield como controlador del rotor.

4. FORMACIÓN DE RECURSOS HUMANOS

Los prototipos en desarrollo presentan una plataforma realista de experimentación. Le permiten a investigadores y estudiantes probar soluciones de software, obtener límites, comparar alternativas y establecer criterios de decisión.

Actualmente el grupo de investigación este compuesto por un investigador formado, cuatro investigadores en formación, tres alumnos investigadores y un alumno becario BIC (Beca de Investigación Científica UNLaM).

5. BIBLIOGRAFÍA

- [1] GIDSA, «UNLaM Ground Segment,» [En línea]. Available: <https://ugs.unlam.edu.ar>.
- [2] GIDSA, «Grupo de Investigación y Desarrollo de Software Aeroespacial de la Universidad Nacional de La Matanza,» [En línea]. Available: <https://gidsa.unlam.edu.ar>.
- [3] P. Soligo, G. Merkel y J. S. Ierache, «Software de segmento terreno de próxima generación,» de *Congreso Argentino de Ciencias de la Computación*, Tandil, 2018.

- [4] P. Soligo, G. Merkel y J. S. Ierache, «Investigación, desarrollo y publicación de un prototipo de segmento terreno satelital,» de *Workshop de Investigadores en Ciencias de la Computación*, 2021.
- [5] P. Soligo, G. Merkel y J. S. Ierache, «Telemetría de altas prestaciones sobre base de datos de serie de tiempo,» *REDDI*, vol. 5, nº 2, 2020.
- [6] SatNOGS, «SatNOGS,» 2023. [En línea]. Available: <https://satnogs.org/>. [Último acceso: 20 Enero 2023].
- [7] D. J. White, I. Giannelos, A. Zissimatos, E. Kosmas, D. Papadeas, P. Papadeas, M. Papamathaiou, N. Roussos, V. Tsiligiannis y I. Charitopoulos, «SatNOGS: Satellite Networked Open Ground,» *Engineering Faculty Publications*, 2015.
- [8] «OPENMCT Getting Started,» [En línea]. Available: <https://github.com/nasa/openmct-tutorial>. [Último acceso: 7 Enero 2023].
- [9] D. Beazley, «Understanding the python gil,» de *PyCON Python Conference. Atlanta, Georgia*, 2010.
- [10] J. G. Filgueira, D. Melgarejo Rodao y M. G. Mullukian Panosian, Construcción y Operación de Estación Terrena para el Seguimiento de Satélites, Montevideo:

Universidad de la República Facultad de Ingeniería, 2019.

ACRÓNIMOS

GIDSA	
Grupo de Investigación y Desarrollo de Software Aeroespacial .	1
GIL	
Bloqueo de Intérprete Global, del inglés Global Interpreter Lock	
.....	3
HTTP	
Protocolo de transferencia de hipertexto, del inglés Hypertext Transfer Protocol.....	1
HTTPS	
Protocolo de transferencia de hipertexto seguro, del inglés Hypertext Transfer Protocol Secure	1
NASA	
Administración Nacional de Aeronáutica y el Espacio, del inglés, National Aeronautics and Space Administration.....	2
OpenMCT	
Software de control de misión abierto, del inglés Open Source Mission Control Software	2
PROINCE	
Programa de Incentivos a Docentes Investigadores de la Secretaría de Políticas Universitarias del Ministerio de Educación de la Nación	1
RDBMS	
Sistema de administración de base de datos relacional, del inglés Relational Database Management System	1
UGS	
UNLaM Ground Segment.....	1, 2

Partición de Visual SLAM para su uso con celulares

Alejandro Silvestri, Jorge Eterovic, Alesio Sinopoli

Departamento de Ingeniería e Investigaciones Tecnológicas Universidad Nacional de La Matanza

Florencio Varela 1903 (B1754JEC), San Justo, (5411) 4480-8900

{jasilvestri; eterovic}@unlam.edu.ar; asinopoli@alumno.unlam.edu.ar

RESUMEN

Un sistema de *Visual SLAM* obtiene en tiempo real la localización espacial de un dispositivo con cámara a partir del video que capta, en un recinto previamente mapeado. Esta capacidad se denomina **autolocalización**, imprescindible para la conducción autónoma.

Visual SLAM es una tecnología experimental que promete masificar la capacidad de autolocalización para **AMR** (*Autonomous Mobile Robot*), pequeños vehículos autónomos de carga para espacios controlados.

Es de sumo interés ejecutar *Visual SLAM* en celulares, pero la alta demanda computacional de estos sistemas lo hace inviable.

Este trabajo desarrolla un sistema partido de autolocalización visual basado en *Visual SLAM*, que ejecuta una parte en el celular y otra en una PC de escritorio.

Palabras clave: *VSLAM*, *STELLA_VSLAM*, *VISUAL SLAM*, *BIORRUEDA*, *AMR*, *ORB-SLAM2*

CONTEXTO

Un AMR es un pequeño vehículo autónomo de carga, que circula muchas veces a la par de las personas, en recintos controlados - no en la vía pública - como depósitos, plantas industriales, campus y predios en general.

Los AMR comerciales usan un sistema de autolocalización basado en LIDAR, un sensor efectivo pero muy costoso.

El proyecto PROINCE - Programa de Incentivos a Docentes Investigadores

Secretaría de Políticas Universitarias (2022-2023) de la Universidad Nacional de La Matanza consiste en el desarrollo de un AMR autolocalizado por *Visual SLAM*. El proyecto es multidisciplinario, con partes mecánicas, electrónicas, de control e informáticas.

El presente trabajo se concentra en el componente informático de autolocalización del AMR que se está desarrollando en el proyecto PROINCE de la Universidad Nacional de La Matanza.

1. INTRODUCCIÓN

1.1. Sobre Visual SLAM

Sensores como Lidar 2D y 3D hacen un relevamiento espacial del entorno. Montados sobre un vehículo, a medida que se desplaza sucesivas lecturas de estos sensores permiten relevar grandes áreas, llevando a cabo lo que se denomina **mapeo** del área. La composición en tiempo real de estas lecturas no es trivial, y se denomina *SLAM* (*Simultaneous Localization And Mapping*).

Un sistema SLAM mapea en tiempo real un área mientras el vehículo se desplaza, y a la vez localiza la posición del vehículo en ese mapa. Esta capacidad de autolocalización sobre un mapa relevado es la clave para conducir de manera autónoma al vehículo a través de una trayectoria hacia un destino.

Los Lidares son sensores costosos. En 2007 dos ensayos fundacionales lograron los

primeros sistemas de *Visual SLAM*, que reemplazaron el Lidar por una cámara de video. Estos sistemas son PTAM [1] y monoSLAM [2]. En pocos años el costo de una cámara para PC se redujo enormemente haciendo posible su masificación, poniendo a *Visual SLAM* como una tecnología candidata a sustituir los sistemas de SLAM con Lidar.

A estos trabajos pioneros de *Visual SLAM* le siguieron una sucesión de papers que marcaron una evolución acelerada durante 10 años. En 2014 y 2015 los ensayos LSD-SLAM [3] y ORB-SLAM [4] establecieron el hito máximo en cuanto a performance y precisión, sin alcanzar el “santo grial”: la posibilidad de poder ejecutarse en un celular.

En 2017 *Visual SLAM* se consideró un trabajo terminado, con ORB-SLAM2 [5] representando el estado del arte hasta la fecha. Varios ensayos posteriores han contribuido en mejorar diversos aspectos, entre los que se destacan el uso de cámara estéreo, omnidireccional, RGBD y combinación con acelerómetros, pero con apenas mejoras menores en la performance.

Cabe mencionar el surgimiento de una rama de investigación en 2018, denominada *Spatial AI* [6], que procura desarrollar sistemas de *Visual SLAM* con inteligencia artificial. Esta rama, con muchos méritos propios, no ha aportado a la mejora de la performance (sino todo lo contrario), divergiendo significativamente del objetivo de este trabajo.

En 2019 el ensayo OpenVSLAM [7] puso a disposición de la comunidad un sistema de Visual SLAM basado en ORB-SLAM2 pero con código abierto con licencia BSD (uso libre e irrestricto), con un código totalmente reescrito aplicando buenas prácticas de programación, y corrigiendo una infinidad de bugs que impiden el uso práctico de los otros sistemas publicados.

En 2021 OpenVSLAM se retiró por presuntas violaciones de licencias, y fue sustituido por **stella_vslam** [8], un proyecto abierto dedicado a corregir todas las posibles violaciones de licencia de OpenVSLAM, convirtiéndose en la

única solución de código abierto para proyectos de ingeniería y productos comerciales.

1.2 Visual SLAM con celulares

Hay dos antecedentes de uso de *Visual SLAM* en celulares, ninguno de los cuales satisface la necesidad de un AMR.

En 2015 LSD-SLAM demostró que puede ser usado en celulares para autocalización sobre mapas relevados, desactivando el módulo de mapeo que es el de mayor demanda computacional. A priori esto satisface la necesidad de autocalización para AMR, pero falla en otros aspectos, principalmente en el enorme tamaño de los mapas en memoria, y en la incapacidad de guardarlos en un archivo para ser reutilizados en otro momento.

Entre 2018 y 2023 se han repetido con éxito varios intentos de compilación de ORB-SLAM2 (entre otros) en celulares de alta gama, pero obteniendo siempre una performance muy pobre.

La idea de partir el procesamiento entre un celular con cámara montado en el AMR y una PC fija conectada por WIFI claramente no elimina la necesidad de la PC, pero flexibiliza enormemente el despliegue del sensor (la cámara) en el vehículo experimental, habilitando incluso la experimentación con plataformas muy baratas como los pequeños vehículos robot Arduino.

El planteo simplista consiste en usar el celular como cámara IP y transmitir en tiempo real el video por WIFI a la PC con *Visual SLAM*. Lamentablemente la compresión del video imprescindible para acomodarse en el ancho de banda disponible introduce un retardo superior a 1 s que hace inviable su uso en navegación autónoma.

En Diciembre de 2022 alumnos de grado de ingeniería informática de la Universidad Nacional de La Matanza, participando en el proyecto PROINCE, presentaron con éxito su trabajo de grado UNSLAM [9], la primera aplicación con Visual SLAM partido, con

performance adecuada para la navegación autónoma de un AMR.

Este trabajo demuestra experimentalmente la viabilidad del concepto, logrando un retardo de 200 ms en un celular antiguo de bajo poder computacional, suficiente para la navegación autónoma, y muy mejorable con el uso de celulares de más potentes.

Más notablemente, el sistema no se implementó en una App, sino en una página web, lo que le confiere amplia compatibilidad y fácil despliegue.

UNLaM es la prueba de concepto del módulo de navegación del proyecto PROINCE. Adopta *stella_vslam* como sistema de *Visual SLAM* ejecutando en una PC.

El *pipeline* de *stella_vslam* es complejo, pero a los efectos de la exposición este trabajo lo dividió en dos partes: el preprocesamiento en celular de la imagen obtenida por la cámara, y el posprocesamiento en PC de la información extraída al combinarse con el mapa.

El preprocesamiento es el conjunto de todas las operaciones que se realizan sobre la imagen. Al finalizar, la imagen se descarta, y el sistema sólo utiliza la información extraída de ella, que ocupa una fracción menor que la imagen completa.

Ésta es la parte que se ejecuta en el celular:

1. captación de la imagen color y conversión a monocromática
2. generación de pirámide
3. detección de *keypoints* [10] en todos los niveles de la pirámide
4. selección de *keypoints* representativos por medio de quadtree
5. extracción de descriptores [11, 12]

Este preprocesamiento está desarrollado en C++ utilizando la biblioteca OpenCV [13]. Su implementación en una página web requirió:

- ☒ aislar los módulos de OpenCV utilizados
- ☒ compilarlos como bibliotecas estáticas

- ☒ adoptar un soporte de paralelismo OpenMP [14] con una biblioteca para Web Assembly
- ☒ compilar el código C++ del pipeline para Web Assembly con emscripten [15] y llvm [16], vinculando las bibliotecas estáticas
- ☒ elaborar la página web que capte la imagen con javascript y la preprocese con el módulo anterior en Web Assembly
- ☒ enviar los resultados al servidor: la opción directa, fetch y JSON, se descartó para favorecer una más performante que evita el parseo: web sockets para enviar los datos de manera binaria

De lado de la PC se desarrolló una variante de *stella_vslam* que admite como entrada los datos preprocesados por el celular en lugar de una imagen.

Además se implementaron *bindings* de *stella_vslam* para Python, con dos propósitos principales:

- ☒ permitir al servidor web que recibe los datos del preprocesamiento, enviarlos al *stella_vslam*
- ☒ obtener el resultado buscado: la pose de localización

1.3. Sistema de referencia real

En general los sistemas de Visual SLAM presentan la información de pose en términos de una matriz de 4x4 en coordenadas homogéneas que representan una transformación euclidiana 3D.

Durante el mapeo los sistemas de *Visual SLAM* establecen un sistema de referencia espacial virtual, con un centro y una orientación de ejes arbitraria y desconocida por el usuario. En particular, los sistemas monoculares - aquellos que usan una sola cámara, en contraste con los estereoscópicos y los de profundidad - adoptan también una escala arbitraria, ya que carecen de referencia de escala real.

La segunda parte del presente trabajo fue diseñar e implementar una estrategia de vinculación entre el sistema de referencia virtual arbitrario del *Visual SLAM*, y un

sistema de referencia real definido por el usuario. Los pasos para la vinculación de estos sistemas de referencia son:

1. el usuario posiciona el celular en el punto que se elige como origen (0;0;0) del sistema de referencia, que suele ser en el suelo; el sistema registra las coordenadas virtuales asociadas obtenidas por Visual SLAM, y el acelerómetro del celular para relevar la dirección vertical
2. a continuación el usuario posiciona el celular desplazándolo 1 m sobre el eje X, a las coordenadas (1;0;0); el sistema registra esas coordenadas virtuales y la dirección vertical

Se asume que el suelo es horizontal, el eje Y es vertical y el eje Z es perpendicular a los otros dos.

Con las dos mediciones el algoritmo calcula la matriz de transformación que convierte del sistema virtual arbitrario de *Visual SLAM* al sistema métrico real elegido por el usuario.

Durante la operación, para cada imagen procesada el sistema de *Visual SLAM* informa sus coordenadas virtuales, y el código propio usa la matriz anterior para convertirlas a coordenadas métricas.

2. LÍNEAS DE INVESTIGACION Y DESARROLLO

Este trabajo es parte del proyecto PRINCE de la Universidad Nacional de La Matanza, dedicado al diseño de un AMR versátil de bajo costo como prototipo para uso académico y para uso comercial.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos se describieron en la Introducción, y se resumen aquí:

- ☒ prueba de concepto de Visual SLAM partido, con una medición de performance adecuada para AMR

- ☒ método de vinculación entre sistemas de referencias virtual y real
- ☒ el sistema UNLaM completo y funcionando publicado en GitHub

Los resultados esperados son:

- ☒ Aislación del autolocalizador partido como un módulo para su disponibilidad en otros proyectos, incluyendo PROINCE
- ☒ Modificación y documentación del módulo
- ☒ Tabla comparativa con pruebas de performance con varios celulares
- ☒ Sistema de navegación que indique la trayectoria a seguir a partir de la autolocalización.

4. FORMACIÓN DE RECURSOS HUMANOS

Cinco alumnos de ingeniería informática se recibieron de ingenieros con su trabajo de grado UNLaM.

El proyecto promueve la formación en proyectos de I+D de 4 docentes ingenieros.

5. BIBLIOGRAFIA

- [1] PTAM, Parallel Tracking and Mapping for Small AR Workspaces, Klein -Murray, ISMAR . 2007.
- [2] MonoSLAM: Real -Time Single Camera SLAM, Davidson et al, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, V OL. 29, NO. 6, JUNE 2007.
- [3] LSD -SLAM: Large -Scale Direct Monocular SLAM, Engels et al, Technical University Munich. 2014.
- [4] ORB -SLAM: a Versatile and Accurate Monocular SLAM System, Mur Artal et al, IEEE TRANSACTIONS ON ROBOTICS . 2015.
- [5] ORB -SLAM2: a n Open-Source SLAM System for Monocular, Stereo and RGB-D

Cameras, Mur Artal et al, IEEE TRANSACTIONS ON ROBOTICS . 2017.

[6] FutureMapping: The Computational Structure of Spatial AI Systems, Davidson et al. 2018.

[7] OpenVSLAM : A Versatile Visual SLAM Framework, Sumikura et al, Arxiv . 2019.

[8] https://github.com/stellacv/stella_vslam

[9] <https://github.com/UNLaM>

[10] Faster and better: a machine learning approach to corner detection, Rosten et al, 2008.

[11] BRIEF: Binary Robust Independent Elementary Features, Calonder et al, 2011.

[12] ORB: an efficient alternative to SIFT or SURF, Rublee et al, 2012.

[13] <https://docs.opencv.org/4.7.0/>

[14] <https://www.openmp.org/>

[15] <https://emscripten.org/>

[16] <https://llvm.org/>

Implementación y Simulación de Sistemas de Tiempo Real

Fernando Romero¹, Diego Encinas¹, Armando De Giusti^{1,2}, Santiago Medina¹, Lucas Maccallini, César Estrebou¹, Alan Castelli¹, Horacio Villagarcía^{1,3}, Fernando G. Tinetti^{1,3}

¹Instituto de Investigación en Informática LIDI (III-LIDI)³
Facultad de Informática – Universidad Nacional de La Plata – Centro Asociado CIC

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas.

³Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{fromero, dencinas, degiusti, smedina, cesarest, hvw, fernando}@lidi.info.unlp.edu.ar,
{lucas.maccallini, alanfabcast}@gmail.com

Resumen

Esta línea de investigación se dedica al desarrollo y análisis de Sistemas de Tiempo Real. Dentro de esta temática se trabajó sobre cuatro sublíneas: 1) Modelado y Simulación de Sistemas de Tiempo Real. 2) Sistemas Operativos de Tiempo Real (SOTR), hardware de comunicaciones y procesamiento utilizados en nodos de redes de Sistemas Distribuidos de Tiempo Real 3) Posicionamiento, tanto a través de GPS en exteriores como de ultrasonido, señales de infrarrojos y de WIFI para posicionamiento en interiores 4) Inteligencia artificial aplicada al reconocimiento de patrones en tiempo real.

Las simulaciones se realizan tanto de sistemas de hardware y cloud computing como de situaciones tales como circulación de personas en ambientes hospitalarios, transmisión de enfermedades y de virus informáticos.

Contexto

Esta línea de Investigación forma parte del proyecto Computación de Alto Desempeño y Distribuida: Arquitecturas, Algoritmos, Tecnologías y Aplicaciones en HPC, Fog-Edge-Cloud, Big Data, Robótica, y Tiempo Real del Instituto de Investigación en Informática LIDI acreditado por la UNLP.

Palabras Claves: Tiempo Real, Simulación, Comunicaciones, Redes de Sensores, Microcontroladores, Cloud Computing, Inteligencia Artificial.

1. Introducción

Un Sistema de Tiempo Real (STR) es aquel en el cual hay restricciones de tiempo en los plazos en que, ante una determinada entrada, deben producir una salida. Además, interactúan con el mundo físico, o sea sus entradas provienen de información producida por sensores y detectores y sus salidas están conectadas a actuadores [3] [4] [7] [8] [17]. Estas restricciones temporales a los plazos dependen fundamentalmente del medio

físico a controlar. Para respetar esto debe haber una sincronización entre el tiempo que transcurre en el mundo físico o real y el procesamiento dentro del sistema de cómputo. Para ello deberá contar con un reloj de tiempo real sincronizado con algún estándar de tiempo físico. Los sensores pueden ubicarse a distancia del sistema que procesa la información conformando una red de sensores, tanto de conexión con conductores como inalámbricas. También estos sensores pueden ser móviles, lo que trae el problema de saber su ubicación actual [6] [14] [23] En redes de sensores se utilizan placas de desarrollo basadas en microcontroladores como, Arduino, NodeMCU, CIAA [11] [18] y Computadoras de Placa Simple como, Raspberry Pi, Raspberry Pi Zero utilizando diferentes SOTR (Linux RT-Preempt, FreeRTOS, MQX, OSEK-OS, Zephyr, Raspbian, etc.) [5]. Se realizan pruebas de alcance, integridad y funcionalidad de redes de sensores inalámbricas [19] [20] [21] [33] principalmente utilizando módulos WiFi y LoRa [16]. También se estudiaron sistemas para conectar los nodos de una red a plataformas y servicios del Cloud [1] [13]

En el campo del modelado y simulación [10] [9] [15], se trata de predecir el comportamiento y la eficiencia de distintos sistemas ante diferentes escenarios, siendo calibrados con datos reales

2. Resultados y Objetivos

Se han desarrollado tareas sobre los temas antes expuestos tales como:

- Construcción y estudio de redes de sensores inalámbricas basadas en WiFi y LoRa [1].

- Modelado y simulación de transmisión de enfermedades intrahospitalarias [22].
- Simulación de la propagación del dengue[23].
- Paralelización y aceleración en la ejecución de simulaciones [2] [27].
- Posicionamiento 2D mediante ultrasonido [24][25]
- Sistema de seguimiento utilizando GPS y conectado por MQTT[6]
- Posicionamiento por RSSI(received signal strength indicator)[12]
- Detector de Frutas en mal estado basado en Inteligencia Artificial[17, 29] utilizando una biblioteca propia de código abierto para TinyML [30]
- Simulación del movimiento de pasajeros en una estación de tren[26]
- Rendimiento de comunicaciones en tiempo real para aplicaciones 3D [28]

3. Formación de Recursos Humanos

Se desarrollan trabajos de alumnos en la Convocatoria a Proyectos de Desarrollo e Innovación de la Facultad de Informática de la UNLP.

Además, se encuentran en desarrollo 2 Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Computación y Analista en TICs, orientadas a las redes de sensores.

De postgrado, investigadores del grupo están desarrollando un trabajo final de especialización, una tesis de Maestría y una tesis de Doctorado.

5. Referencias

- [1] Integration of Sensor Networks with Cloud ComputingS. Medina, F. Romero, and F. G. Tinetti, Short papers of the 8th Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET

2020), ISBN: 978-950-34-1927-4, págs. 2-5, 2020.

[2] R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose and R. Buyya "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms" Published online 24 August 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/spe.995.

[3] Burns, A. A., Wellings. "Real-Time Systems and Programming Languages: Ada, Real-Time Java and C/Real-Time POSIX", Addison-Wesley Educational Publishers Inc., 2009.

[4] Buttazzo, G. C., "Hard RealTime Computing Systems", Third edition, Springer, 2011.

[5]"FreeRTOS - market leading RTOS (real time operating system) for embedded systems supporting 34 microcontroller architectures". [http:// www.freertos.org/](http://www.freertos.org/).

[6]<https://simcom.ee/modules/gsm-gprs-gnss/sim868/>

[7]Kopetz. H., "Real-Time Systems, Design Principles for Distributed Embedded Applications". Second Edition. Springer. 2011.

[8]Liu, J. W. S. Liu, "Real Time Systems", Integre Technical Publishing Co., Inc., 2000

[9]C. Macal, M. North, Tutorial on agent-based modeling and simulation part 2: how to model with agents, in: Proceedings of the Winter Simulation Conference, 2006.

[10]PHILLIP A. LAPLANTE, SEPPO J. OVASKA. REAL-TIME SYSTEMS DESIGN AND ANALYSIS Tools for the PractitionerFourth Edition. A JOHN WILEY & SONS, INC., PUBLICATION. IEEE PRESS. 2012.

[11]<http://www.proyecto-ciaa.com.ar/devwiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp>

[12]R.S. Rosli, M.H. Habaebi, R. Islam On the analysis of received signal strength

indicator from ESP8266, Bulletin of Electrical Engineering and Informatics · September 2019

[13] Análisis de una plataforma de simulación para Cloud Computing. Un caso de estudio. Tomás Rosales, Julián Spinelli, Marcos Di Nardo, Román Bond, Daniel Rosatto, Diego Encinas, Fernando Romero XXVI Congreso Argentino de Ciencias de la Computación (CACIC 2020) La Matanza

[14] F. G. Tinetti and O. C. Valderrama Riveros, "Unmanned Vehicles: Towards Heterogeneous Hardware Approaches," 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2018, pp. 919-924.

[15] Fernando G. Tinetti, Oscar C. Valderrama Riveros, Fernando L. Romero, "Unmanned Vehicles: Real Time Problems in Drone Receivers", Conf. on Computational Science & Computational Intelligence (CSCI'19), Las Vegas, Nevada, USA , 2019, pp. 1081-1085.

[15] D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.

[16] LoRa <https://www.lora-alliance.org/> 2017

[17] Moreda, G. P., Muñoz, M. A., Barreiro, P. Inspección de frutas y hortalizas mediante

videocámara: estado actual de la técnica.

<https://www.interempresas.net/Horticola/Articulos/56223-Inspeccion-de-frutas-y-hortali>

zas-mediante-videocamara-estado-actual-de-la-tecnica.html

[18] NodeMcu <http://www.nodemcu.com/> 2017

[19] Akyildiz, Ian F., and Mehmet Can Vuran. "Wireless sensor networks" Vol. 4. John Wiley & Sons, 2010.

[20]Lewis, Franck L."Wireless sensor networks." Smart environments: technologies, protocols, and applications 11 (2004): 46.

- [21] Raghavendra, Cauligi S., Krishna M. Sivalingam, and Taieb Znati, eds. "Wireless sensor networks" Springer, 2006.
- [22] Maccallini, L. ., Encinas, D. O., & Romero, F. . (2021). An Approach to the Modeling and Simulation of Intra-Hospital Diseases. *Journal of Computer Science and Technology*, 21(2), e14. <https://doi.org/10.24215/16666038.21.e14>
- [23] Oviposition activity and seasonal pattern of a population of *Aedes* (*Stegomyia*) *aegypti* (L.) (Diptera: Culicidae) in subtropical Argentina
Autores: Micieli, María Victoria | Campos, Raúl E. *Memórias do Instituto Oswaldo Cruz*; vol. 98, no. 5. 2003
- [24] Triangulation based Fusion of Ultrasonic Sensor Data. 1998. O. Wijk, P. Jensfet, H.I. Christensen.
- [25] PGA460 Array of Ultrasonic Transducers for Triangulation and Tracking. 2019. Akeem Whitehead and Bharat Aravamudhan
- [26] Uri Wilensky (29 de septiembre de 2022). NetLogo User Manual version 6.3.0. Disponible en Internet en: <https://ccl.northwestern.edu/netlogo/docs/>
- [27] Martín Paradiso, Lucas Maccallini, Agustina Vericat, Fernando Romero, Diego Encinas. An Approach to Parallelization of Respiratory Disease Spread Simulations in Emergency Rooms. 2022 International Conference on Computational Science and Computational Intelligence (CSCI'22). En prensa.
- [28] Mauro Santos, Diego Encinas. Análisis de plataformas de Computación en la Nube para implementación de protocolo de comunicaciones con una aplicación móvil 3D. XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022) La Rioja
- [29] Gutti, V., Karthi, R. (2022). Real Time Classification of Fruits and Vegetables Deployed on Low Power Embedded Devices Using Tiny ML. Third International Conference on Image Processing and Capsule Networks. ICIPCN 2022. Lecture Notes in Networks and Systems, vol 514. Springer, Cham. https://doi.org/10.1007/978-3-031-12413-6_27
- [30] Estrebou, C.A., Fleming, M., Saavedra, M.D., Adra, F., De Giusti, A.E. (2022). Lightweight Convolutional Neural Networks Framework for Really Small TinyML Devices. *Smart Technologies, Systems and Applications. SmartTech-IC 2021. Communications in Computer and Information Science*, vol 1532. Springer, Cham. https://doi.org/10.1007/978-3-030-99170-8_1

ORCID autores:

Fernando Romero: 0000-0002-1498-3752

Diego Encinas: 0000-0002-6948-9786

A. De Giusti: 0000-0002-6459-3592

Santiago Medina: 0000-0001-6852-7165

Horacio Villagarcía:

Fernando G. Tinetti:

César Estrebou: 0000-0001-5926-8827

CAPTURA DE VARIABLES PARA ALERTA TEMPRANA DE INCENDIOS FORESTALES Y SU ALMACENAMIENTO PARA INTEGRACIÓN A SISTEMAS DE PREDICCIÓN MEDIANTE UTILIZACIÓN DE REDES INALÁMBRICAS DE SENSORES

Rodrigo Atilio Elgueta^{1,2}, Miguel Mendez-Garabetti^{1,2}

¹Universidad de Mendoza, Dirección de Posgrado, Facultad de Ingeniería
rodrigo.elgueta@um.edu.ar, miguel.mendez@um.edu.ar

²Free and Open Source Software/Hardware Research Laboratory (FOSSHLab), Argentina.

RESUMEN

Los incendios forestales, generan grandes pérdidas y daños alrededor del mundo generando un impacto negativo significativo en el medio ambiente, la economía y la sociedad. La utilización de redes inalámbricas de sensores (WSN), permite recopilar datos en tiempo real sobre las condiciones ambientales, como la temperatura, la humedad, la velocidad del viento, la dirección del viento, la concentración de gases peligrosos, etc. Estos se utilizan para enviar alertas en tiempo real a los sistemas de gestión de emergencias en caso de detección de niveles anormales de humo, calor o gases peligrosos, y permiten tomar acciones de contención más efectivas y planear medidas preventivas para contener el incendio forestal antes de que se propague a áreas más grandes. Además, se pueden utilizar para alimentar modelos de predicción de incendios, lo que mejora la precisión y la eficacia de estos modelos, proporcionando datos certeros respecto a las variables obtenidas en la medición. Esto puede mejorar la capacidad de los sistemas de predicción de incendios forestales y también puede ayudar en la gestión de emergencias en tiempo real para contener y extinguir el incendio de manera más efectiva.

Palabras clave: Red Inalámbrica de Sensores, Teledetección de Incendios

CONTEXTO

El presente trabajo de I+D se desarrolla como proyecto de Investigación de tesis de posgrado presentado y aprobado en la Maestría en Teleinformática, Dirección de Posgrado, perteneciente a la Facultad de Ingeniería de la Universidad de Mendoza, (Ciudad, Mendoza).

INTRODUCCIÓN

El Servicio Nacional de Manejo del Fuego del Ministerio de Ambiente y Desarrollo Sostenible de la República Argentina, define a incendio forestal como un fuego que se propaga libremente con efecto no deseado para la vegetación y sin estar sujeto a control humano. Cuando afecta zonas no boscosas ni aptas para la forestación, se incluye el término incendio rural. [1]. Generalmente, es producido por descuidos humanos, o de forma ocasional, producido naturalmente. Si encuentra condiciones apropiadas para su expansión, puede recorrer extensas superficies causando importantes pérdidas ecológicas, económicas y sociales. [2] Este fuego es la reacción rápida producto de la unión del oxígeno del aire, la cobertura vegetal como combustible y una fuente de calor. El mismo transita cinco etapas: Fuera de control, cuando se propaga libremente. Detenido, o contenido cuando la propagación se ha detenido en uno o más sectores del incendio. Circunscripto cuando la contención abarca todo el perímetro pero no está

definitivamente terminado. Controlado cuando existe una línea de control establecida definitivamente, anclada y asegurada. Extinguido cuando el incendio no muestra signos de actividad en ninguna de sus partes [3].

Durante años se ha trabajado en el desarrollo de diferentes herramientas para la prevención, cuyo objetivo es lograr que los incendios no alcancen a materializarse, la detección, a fin determinar la ubicación de los focos de incendio antes de que se pierda el control sobre ellos y la predicción de incendios forestales. Cada una de éstas corresponde a distintas fases del proceso de lucha contra incendios. [4] [5] [6] [7]

Los métodos de predicción son técnicas utilizadas para predecir eventos futuros basados en datos históricos y patrones previamente identificados. En el caso de los incendios forestales, los métodos de predicción se utilizan para identificar áreas de riesgo y anticiparse a posibles incendios. En este sentido, tiene dos acepciones: predicción de incendios forestales y predicción del comportamiento de incendios forestales. La primera de ellas intenta predecir la ocurrencia de incendios antes de que estos sucedan, mientras que la segunda, busca determinar el posible comportamiento de un incendio forestal una vez que éste ya se ha iniciado, permitiendo tomar decisiones acertadas en el plan de acción a tomar [8].

Los métodos de predicción utilizados en incendios forestales incluyen: Métodos estadísticos, basados en el análisis de datos históricos de incendios forestales, identificación de patrones y tendencias. utilizados para predecir la probabilidad de que ocurra un incendio forestal en un área determinada. Basados en modelos físicos que se estructuran en la comprensión de los procesos físicos involucrados en la combustión y en la propagación del fuego utilizando

modelos matemáticos para simular el comportamiento del fuego y predecir su propagación. Basados en tecnología de sensores remotos que utilizan tecnología de sensores remotos para obtener datos sobre las condiciones ambientales en áreas forestales que se utilizan para predecir la probabilidad de que ocurra un incendio forestal. Basados en sistemas de información geográfica (SIG) que utilizan datos geográficos y de manejo forestal para identificar áreas de riesgo de incendios forestales.

Las implementaciones computacionales de dichos modelos suelen realizarse con simuladores de comportamiento de incendios. Debido al tamaño del conjunto de datos y la complejidad de las operaciones que deben efectuarse, puede requerirse la utilización de sistemas de alto rendimiento y procesamiento paralelo (HPC, High performance Computing) para resolver el problema en el menor tiempo posible [9].

Las redes inalámbricas de sensores, además de utilizarse como sistema de detección de incendios en tiempo real, permiten el almacenamiento de valores medidos a fin complementarse con los sistemas de predicción integrando variables reales a los mismos con fin de reducir la brecha de incertidumbre del modelo y evitar cálculos en base a estimaciones indirectas.

Las Redes Inalámbricas de Sensores (WSN, Wireless Sensor Networks) consisten en sensores conectados a nodos que utilizan canales de radio para comunicarse y transmitir datos. Dichos Nodos, son capaces de obtener datos de su entorno, procesarlos localmente, y comunicarlos a través de enlaces inalámbricos. Cada nodo actúa como elemento de la infraestructura de comunicaciones al reenviar los mensajes transmitidos por nodos más lejanos hacia un centro de recolección de información. [10].

El presente trabajo desarrolla un proyecto de utilización de una WSN con componentes de bajo costo para la recolección de datos necesarios para detección de incendios forestales a fin de generar alertas tempranas de ocurrencia de incendios así como el almacenamiento de dichos valores para permitir su utilización por modelos de predicción.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Identificación y selección de elementos a monitorear

Para establecer los elementos a monitorear, es necesario realizar previamente un estudio de la zona que se desea monitorear ya que, tanto la detección de incendios como los datos necesarios para integrar a sistemas de predicción dependen de: Las condiciones meteorológicas: temperatura, humedad relativa, velocidad y dirección del viento, precipitaciones, entre otros. Las características del combustible: tipo, humedad, densidad, altura y continuidad, entre otros. La topografía y características del terreno: pendiente, orientación, elevación, tipo de suelo, vegetación, entre otros. La presencia de fuentes de ignición: actividad humana, rayos, etc. Y del historial de incendios forestales en la zona [11].

Definición de infraestructura general

La infraestructura general planteada en el presente, se puede descomponer en los siguientes elementos: WSN para la adquisición de datos. Una puerta de enlace para conexión y envío de datos a internet. La utilización de servicios específicos para el tratamiento de la información. Particularmente, en el presente se utilizó contenedores para los siguientes servicios: EMQ para la administración de mensajes MQTT. MariaDB como base de

datos para el almacenamiento de variables obtenidas, Python como lenguaje para generar interfaces y adaptación de variables. Se optó por la utilización de contenedores en función de su adaptabilidad, portabilidad y posibilidad de brindar escalabilidad. En función a que cada sensor enviará datos a la red y los mismos serán almacenados y procesados a fin de obtener información, se puede establecer que los mismos, son parte de la IoT (Internet de las cosas) [12].

Definición de WSN

En el presente trabajo, se realizó un laboratorio para la captura de datos utilizando elementos WiFi convencionales mediante el uso de dispositivos ESP8266. Se utilizó la capacidad del procesador para: Crear una WSN utilizando una topología en malla dinámica, donde los dispositivos se conectan entre sí cambiando las rutas de envío de información según la necesidad o pérdida de alguno de ellos. Procesar los datos recibidos por los sensores directamente conectados como entradas al dispositivo.

Asimismo, como se destacó anteriormente, gran parte del armado de la WSN, dependerá exclusivamente de los sensores a utilizar para la obtención de datos. En general, se requiere monitorear diversas variables relacionadas con el combustible, como la carga de combustible, la densidad aparente, la altura de la capa muerta, la humedad del combustible, la humedad del aire, la velocidad del viento y la pendiente del terreno. Para medir estas variables, se pueden utilizar diversos sensores, como: Humedad del suelo: para medir la humedad del combustible. Humedad relativa: para medir la humedad del aire. Anemómetro: para medir la velocidad del viento. Inclinación: para medir la pendiente del terreno. Temperatura: para medir la temperatura del aire. Monóxido de carbono: para medir la existencia de gases. Para obtener valores

precisos de estas variables, es necesario colocar los sensores en ubicaciones estratégicas, teniendo en cuenta factores como la altura y la distancia a la que se encuentran del material combustible. Asimismo, debido a la extensión de las áreas donde pueden ocurrir incendios forestales, agregar un sensor a cada elemento combustible sería una tarea muy difícil e impráctica debido al gran número de elementos combustibles presentes en un área forestal. En su lugar, se podrían utilizar sensores distribuidos estratégicamente en diferentes puntos de la zona forestal, capaces de proporcionar una imagen global de las condiciones ambientales y de los materiales combustibles presentes. Para ello, si bien en el presente sólo se utilizó elementos WSN basados en WiFi 802.11, sería útil contar con infraestructuras WSN híbridas.

Protocolos y primitivas de conexión

Para la WSN existen diferentes tecnologías, entre ellas las más importantes son: IEEE 802.15.4, ZigBee e IEEE 802.11 b/g/n [13]. Asimismo, para la presente, por motivos de costos, se realizó un laboratorio utilizando componentes con el estándar IEEE 802.11 b/g/n de bajo consumo configurados con una topología en mesh con una lógica jerárquica donde se designan nodos como “cluster-head” con mayor responsabilidad para controlar a otros nodos. Asimismo, como se mencionó con anterioridad, para la WSN sería útil contar con infraestructuras WSN híbridas a fin de obtener mayor alcance para el muestreo.

En cuanto a la transmisión de los datos obtenidos, se utilizó el protocolo MQTT. Éste permite la comunicación de dispositivos de Internet de las cosas (IoT) que generalmente tienen que transmitir y recibir datos a través de una red con recursos restringidos y un ancho de banda limitado. Además, MQTT admite la mensajería entre dispositivos a la nube y la nube al dispositivo [14].

Almacenamiento de datos

Como se mencionó con anterioridad, mediante el almacenamiento de los datos recolectados, se puede obtener estadísticas y variables que pueden ser utilizadas por sistemas de predicción. En este trabajo en particular, se utilizó MariDB su almacenamiento [15]. Asimismo, en este sentido, es necesario contar con una interfaz que reciba los mensajes enviados mediante MQTT e inserte los registros en las tablas correspondientes de la base de datos. Para ello, se utilizó Python como lenguaje de programación [16] y se realizó dicho modelo acorde al laboratorio desarrollado en el presente trabajo.

RESULTADOS OBTENIDOS

El presente proyecto, si bien sólo ha sido probado en laboratorio, posee el potencial para brindar la flexibilidad necesaria que le permita implementarse tanto en equipos de monitoreo estáticos como en aquellos desplegados sobre frentes de incendio. Además, por el uso del protocolo MQTT para el envío de mensajes con los datos recolectados, puede integrarse y formar parte de otras tecnologías y topologías de red. Un ejemplo de ello sería la formación de pequeñas células WSN distribuidas en los extremos de otra red de mayor alcance para la recolección de datos. Respecto a la incorporación de datos a sistemas de predicción, el almacenamiento de datos permite que a partir de los mismos, se realicen conversiones de medidas y se utilicen como entradas de variables en sistemas de predicción de comportamiento de incendios. Por lo tanto, el despliegue del presente proyecto, podría colaborar en gran medida en la generación de alertas de ocurrencia de incendios forestales así como en el planeamiento estratégico de la lucha contra el fuego, generación de estadísticas y facilitaría la incorporación de

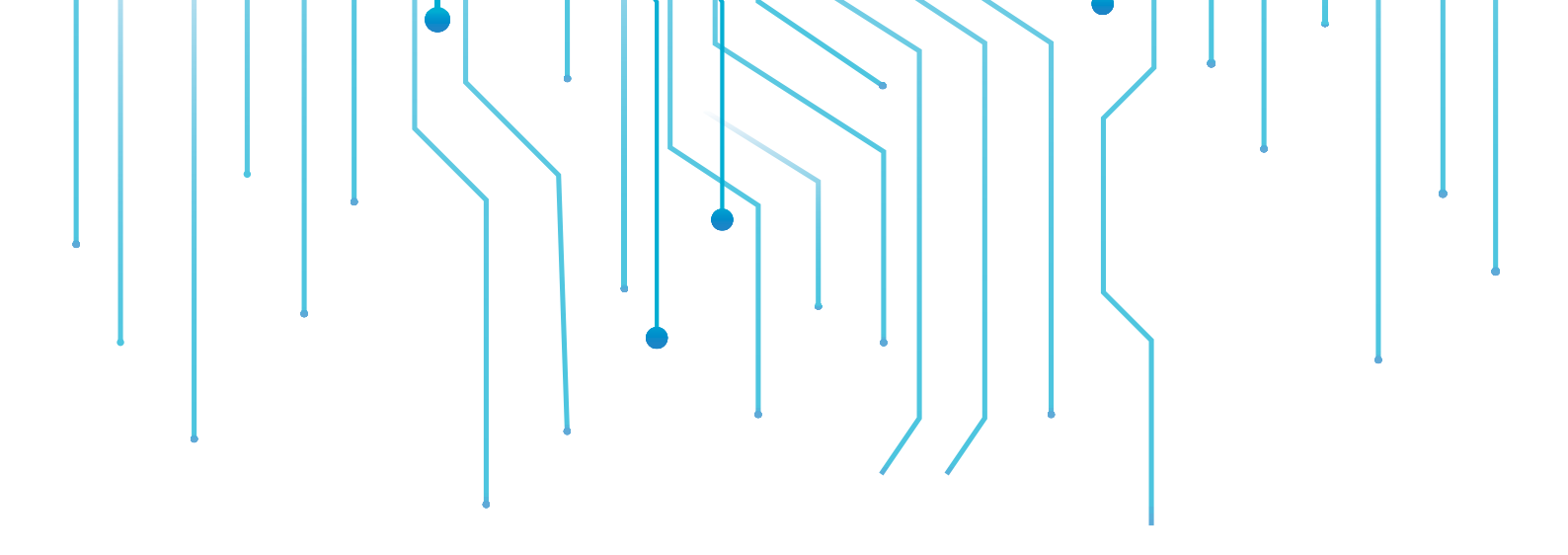
datos reales en sistemas de predicción de comportamiento de incendios forestales.

FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesis de posgrado por parte del Ing. Rodrigo Atilio Elgueta, quien es estudiante de la Maestría en Teleinformática de la Universidad de Mendoza y dirigida por el Dr. Miguel Mendez-Garabetti.

BIBLIOGRAFÍA

- [1] «¿Qué es un incendio forestal?», *Argentina.gob.ar*, 15 de marzo de 2018. <https://www.argentina.gob.ar/ambiente/fuego/conocemas/incendioforestal> (accedido 26 de febrero de 2023).
- [2] J. E. G. Mora, «INCENDIOS FORESTALES: causas e impactos», *El Antoniano*, vol. 135, n.º 1, Art. n.º 1, 2020, doi: 10.51343/anto.v135i1.866.
- [3] «script-tmp-inta_-_manual_de_combate_de_incendios_forestales_y_ma.pdf». Accedido: 26 de febrero de 2023. [En línea]. Disponible en: https://inta.gob.ar/sites/default/files/script-tmp-inta_-_manual_de_combate_de_incendios_forestales_y_ma.pdf
- [4] P. N. Omi, *Forest fires: a reference handbook*. Santa Barbara, Calif: ABC-CLIO, 2005.
- [5] E. A. Johnson y K. Miyanishi, Eds., *Forest fires: behavior and ecological effects*. San Diego, Calif: Academic Press, Inc, 2001.
- [6] J. G. Flores Garnica y D. A. Rodríguez Trejo, *Incendios forestales : definiendo el problema, ecología y manejo, participación social, fortalecimiento de capacidades, educación y divulgación*. México: Mundi-Prensa, 2006.
- [7] J. Lowe, *Wildland firefighting practices*. Africa ; Albany, N.Y: Delmar Thomson Learning, 2001.
- [8] M.-G. Miguel, T. M. Laura, B. Germán, y C.-S. Paola, «Predicción del Comportamiento de Incendios Forestales mediante un Método de Reducción de Incertidumbre basado en HPC y Evolución Diferencial», 2014.
- [9] J. Strappa, P. Caymes-Scutari, y G. Bianchini, «A Parallel Novelty Search Metaheuristic Applied to a Wildfire Prediction System». arXiv, 23 de julio de 2022. Accedido: 26 de febrero de 2023. [En línea]. Disponible en: <http://arxiv.org/abs/2207.11646>
- [10] R. Azzam, T. M. Fernández-Steeger, C. Arnhardt, H. Klapperich, y K.-J. Shou, *Monitoring of landslides and infrastructures with wireless sensor networks in an earthquake environment*. Santiago, 2011.
- [11] D. Saha, B. K. Shaw, S. Bose, y S. Paik, «WSN based Forest Fire Detection System», *Adv. J. Eng.*, pp. 12-17, 2022, doi: 10.55571/aje.2022.04013.
- [12] J. S. Rueda y J. M. Talavera Portocarrero, «Similitudes y diferencias entre Redes de Sensores Inalámbricas e Internet de las Cosas: Hacia una postura clarificadora», *Rev. Colomb. Comput.*, vol. 18, n.º 2, pp. 58-74, dic. 2017, doi: 10.29375/25392115.3218.
- [13] J. Garbarino, *Protocolos para redes inalámbricas de sensores : conciencia de energía y técnicas de disseminación en protocolos de red para redes inalámbricas de sensores / Jimena Garbarino*. Alemania: Editorial Académica Española, 2012.
- [14] «MQTT - The Standard for IoT Messaging». <https://mqtt.org/> (accedido 26 de febrero de 2023).
- [15] «MariaDB Foundation - MariaDB.org». <https://mariadb.org/> (accedido 26 de febrero de 2023).
- [16] «Welcome to Python.org», *Python.org*, 15 de febrero de 2023. <https://www.python.org/> (accedido 26 de febrero de 2023).



RCCI

Redes de Cooperación Científica Internacionales

Maestría en Gestión y Tecnología de Ciudades Inteligentes

Resultado del Proyecto CAP4CITY

Patricia Pesado¹, Ariel Pasini¹,
Pablo Thomas¹, Rocío Muñoz¹, Armando De Giusti¹

Elsa Estevez^{2,4}, Pablo Fillottrani^{2,3}, Karina Cenci^{2,3}, Gabriela Andrea Díaz^{2,4}

¹ Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 - La Plata, Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 - La Plata, Buenos Aires
{ppesado, apasini, pthomas, rmunoz, degiusti} @lidi.info.unlp.edu.ar

² Departamento de Ciencias e Ingeniería de la Computación – Universidad Nacional del Sur (UNS)

³ Laboratorio de Ingeniería de Software y Sistemas de Información (LISSI), Centro Asociado CIC

⁴ Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Av. San Andrés 800 – Campus de Palihue - Bahía Blanca, Buenos Aires
526 e/ 10 y 11 - La Plata, Buenos Aires
{ece, prf, kmc, gabriela.diaz} @cs.uns.edu.ar

RESUMEN

El proyecto CAP4CITY, cofinanciado en el marco del Programa Erasmus+ de la Unión Europea entre los años 2018 y 2022, para la construcción de capacidades en la educación superior, tuvo como objetivo fortalecer y desarrollar la capacidad de instituciones académicas en América Latina y Europa para mejorar la calidad de la educación superior en el campo de ciudades inteligentes sostenibles. Como uno de los resultados de este proyecto, se creó la Maestría en Gestión y Tecnología de Ciudades Inteligentes, formulada de forma conjunta entre la Universidad Nacional de La Plata (UNLP) y la Universidad Nacional del Sur (UNS) de Argentina, actualmente en proceso de acreditación por CONEAU, para ser dictada en modalidad a distancia.

Palabras Clave:

Ciudades Inteligentes – Sostenibilidad - Maestría Interinstitucional - Colaboración Internacional – Construcción de Capacidades – Transformación Digital

CONTEXTO

Uno de los objetivos del proyecto “Strengthening Governance Capacity for Smart Sustainable Cities (CAP4CITY)” fue desarrollar cursos universitarios utilizando nuevas herramientas de enseñanza y aprendizaje, y nuevos planes de estudio en todos los niveles del proceso educativo para el desarrollo de Ciudades Inteligentes Sostenibles. El Proyecto, financiado como parte del programa Erasmus+ de la Unión Europea, fue ejecutado por un consorcio integrado por 12 universidades; cuatro de ellas europeas – Donau Universität für Weiterbildung (DUK) en Austria, Tallinn University of Technology (TUT) en Estonia, Delft University of Technology (TU Delft) en los Países Bajos, y Gdańsk University of Technology (GUT) en Polonia; y ocho universidades en la región de América Latina – Universidad Nacional de La Plata (UNLP) y Universidad Nacional del Sur (UNS) en Argentina; Pontificia Universidade Católica do Rio Grande do Sul (PUCRS) y Faculdade Meridional (IMED) en Brasil; Universidad Técnica Federico Santa María (UTFSM) y Universidad Católica del Norte (UCN) en

Chile; y Universidad Externado de Colombia (UEC) y Escuela Colombiana de Ingeniería (ECI) en Colombia.

En Argentina existen distintos niveles formales de educación de grado y postgrado. Para la creación de una nueva Maestría a distancia es necesario definir las competencias del egresado y seguir la reglamentación del SIED (Sistema de Educación a Distancia). Es importante destacar que las carreras interinstitucionales y los convenios entre universidades deben ser acreditados por la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU).

1. INTRODUCCION

La Maestría en Gestión y Tecnología de Ciudades Inteligentes, propuesta por la Universidad Nacional de La Plata (UNLP) y la Universidad Nacional del Sur (UNS) se enfoca en la formación de recursos humanos capacitados para la gestión de ciudades inteligentes y sostenibles, con un conocimiento del estado actual de la tecnología y su aplicación, en el contexto de una sociedad digital.

El Programa se concibe como un espacio de educación continua, actualización tecnológica, científica y metodológica en el campo de gobierno digital aplicado a ciudades inteligentes y sostenibles. Se trata de crear capacidades en los recursos humanos que trabajen o colaboren con instituciones públicas para que las mismas sean más eficientes, transparentes y que pongan al ciudadano en el centro de sus agendas.

Cuando se habla de Ciudades Inteligentes y Sostenibles se hace referencia a aquellas ciudades innovadoras que utilizan las Tecnologías de la Información y la Comunicación (TIC), y otros medios, para mejorar la calidad de vida, la eficiencia de la operación y los servicios urbanos, y la competitividad, al tiempo que se garantiza

que satisfaga las necesidades de las generaciones presentes y futuras con respecto a los aspectos económicos, sociales, ambientales y culturales.

Los cursos definidos para la Maestría desarrollan las competencias del futuro egresado. Para cada competencia a desarrollar para el futuro egresado se definen uno o más cursos que abordan su desarrollo.

La carrera propone una modalidad educativa híbrida que combina la realización de encuentros presenciales y/o por videoconferencia y el trabajo mediado por tecnologías digitales, en particular por un entorno virtual de enseñanza y aprendizaje (EVEA). Acorde a la temática y metas de la carrera, los cursos fueron diseñados orientándose al diálogo y a la participación activa de los alumnos, a partir de la realización de actividades que promuevan el intercambio de ideas y experiencias en entornos digitales, la colaboración, la reflexión, el análisis, la lectura crítica, la búsqueda de información en medios digitales, y la transferencia de los conocimientos a los propios contextos de trabajo. De esta manera, los alumnos no sólo tienen oportunidad de analizar los marcos teóricos que sustentan los temas abordados, sino que también realizan experiencia, y vivencian determinadas estrategias metodológicas a partir de las actividades propuestas en los cursos de la carrera. Los encuentros sincrónicos y el tipo de actividades a realizar en cada caso, son estratégicos y se planifican acorde a los contenidos, objetivos y competencias que se proponen.

La carrera es de tipo estructurado, comprendiendo 9 cursos teórico-prácticos obligatorios y una Tesis de Maestría. A los cursos pueden agregarse actividades complementarias tales como cursos optativos, trabajos, tutoriales, seminarios, etc., buscando completar la formación e información de los alumnos.

ASIGNATURAS	CARGA HORARIA
1.Introducción a la gestión de Ciudades Inteligentes (UNS)	75 Hs.
2.Aplicaciones en Ciudades Inteligentes (UNLP)	75 Hs.
3.Administración y Estrategias para la Transformación Urbana (UNS)	75 Hs.
4.Aspectos Legales para Ciudades Inteligentes (UNS)	75 Hs.
5.Taller de Metodologías de Investigación (UNLP)	75 Hs.
6.Gobernanza y Gestión de Tecnologías de la Información (UNLP)	75 Hs.
7.Gobernanza de Datos y Gestión de la Información (UNS)	75 Hs.
8.Diseño y Co-Creación de Servicios (UNLP)	75 Hs.
9.Materia optativa:	75 Hs.
<i>Análisis Inteligente de datos en entornos de Big Data (UNLP)</i>	
<i>Comunicación e Interacción en Redes de Trabajo (UNS)</i>	
<i>Educación Digital (UNLP)</i>	
<i>Transformación Digital (UNS)</i>	
10.Tesis de Maestría (UNLP-UNS)	225 Hs.
TOTAL DE HORAS	900 Hs.

La Maestría en Gestión y Tecnología de Ciudades Inteligentes será una maestría profesional que otorgará una titulación conjunta de ambas universidades.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La Maestría incursiona sobre distintos aspectos que representan los temas de interés de las unidades de investigación que sustentan la carrera, como por ejemplo:

- comprender los requerimientos de gestión y tecnología de las ciudades inteligentes, así como su proyección futura en función de las estrategias de desarrollo en los diferentes países y en el cambio tecnológico en curso;
- comprender al gobierno digital como proyectos estratégicos transversales de las administraciones públicas que sirven para simplificar las interacciones del Estado con los ciudadanos;
- analizar modelos regulatorios, organizacionales y de gobernanza para gobierno digital;

- identificar modelos de gestión de las tecnologías digitales que contribuyan a transformar las estructuras y procesos de las instituciones públicas a fin de diseñarlos para que den respuestas a las necesidades de los ciudadanos y facilitar la adecuada toma de decisiones;
- desarrollar capacidades para la aplicación directa de la tecnología al servicio del ciudadano (aplicaciones móviles, servicios WEB, consumo energético, seguridad, tránsito, etc.);
- desarrollar habilidades blandas, como la comunicación, pensamiento crítico, motivación, negociación, y toma de decisiones que permitan la planeación, implementación y monitoreo de proyectos de gobierno digital.

3. RESULTADOS ESPERADOS

Durante 2023 se comenzarán a dictar cursos correspondientes a la Maestría en el contexto de los programas de doctorado de ambas unidades académicas. Estos cursos serán acreditables para la Maestría una vez aprobada la carrera por CONEAU.

4. FORMACION DE RECURSOS HUMANOS

Los egresados de la Maestría serán profesionales con habilidades para gestionar estrategias de tecnologías digitales que contribuyan a la implementación y uso de servicios digitales para facilitar las condiciones de vida de los ciudadanos y promover al desarrollo socio-económico, con el foco puesto en el desarrollo actual y futuro de ciudades inteligentes.

En particular, los egresados de la Maestría estarán capacitados para:

- conocer las tecnologías actuales aplicables en ciudades inteligentes;
- diseñar servicios públicos digitales en base a las necesidades de los ciudadanos;
- conocer cómo generar valor público a través de iniciativas de gobierno digital;
- dirigir cambios organizacionales facilitados por la tecnología;
- conocer herramientas informáticas para mejorar la construcción de políticas públicas a través de la participación ciudadana;
- alinear el desarrollo estratégico de tecnología con la misión, objetivos organizacionales y las necesidades de los ciudadanos;
- diseñar soluciones de gobierno digital interoperables y sostenibles, en el marco de políticas públicas / prioridades establecidas por el país;
- identificar barreras para la innovación de servicios públicos y el modo de superarlas desde la educación digital (que incluye la educación formal e informal);
- impulsar la capacitación de las personas ciudadanas en las tecnologías digitales y los servicios derivados de las mismas, de modo de mejorar su calidad de vida;

- impulsar una estrategia comunicacional que garantice la comunicación efectiva con todos los interesados;
- conocer nuevas tecnologías y su aplicación en cuestiones de gobierno;
- asesorar sobre cuestiones de seguridad de la información y otros activos digitales.

5. AGRADECIMIENTOS

Los autores agradecen al Proyecto CAP4CITY cofinanciado por el Programa Erasmus+ de la Unión Europea (Id 598273-EPP-1-2018-1-AT-EPPKA2-CBHE-JP), así como a las universidades miembro y a sus representantes en el Consorcio del Proyecto.

6. BIBLIOGRAFIA

- Strengthening Governance Capacity for Smart – Final Report
- CAP4CITY Deliverable 5.3 New Executive Master Programme on SSC.
- La biblioteca de Cursos CAP4CITY Para Desarrollar Capacidades en Ciudades Inteligentes y Sostenibles
- UNLP – Disposición R N° 424 sobre Expediente 3300-4663-21-001 – Creación de la Maestría Interinstitucional. Diciembre 2021
- UNS – Resolución AU-06/22 Sobre Expediente 3995/21 – Creación de la Maestría Interinstitucional. Septiembre 2022
- Procedimiento de acreditación de carreras de postgrado
<https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-160-2011-192733/actualizacion>

Avances del Consorcio para la colaboración en I+D+I en Temas de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET)

III-LIDI – Instituto de Investigación en Informática LIDI (UNLP – Argentina)
LISSI – Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información (UNS – Argentina)
VyGLab – Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (UNS – Argentina)
LIDIC – Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (UNSL – Argentina)
HPC4EAS – High Performance Computing for Efficient Applications and Simulation (UAB – España)
SMILe – Soft Management of Internet and Learning (Universidad de Castilla-La Mancha – España)
ArTeCS – Group of Architecture and Technology of Computing Systems (UCM – España)
LITRP – Laboratorio de Investigaciones Tecnológicas en Reconocimiento de Patrones (UCM – Chile)
LCG – Laboratorio de Computación Gráfica (UNSL – Argentina)
ITIC – Instituto para las Tecnologías de la Información y las Comunicaciones (UNCu – Argentina)
DisCo – Grupo de I+D+I en Computación Distribuida (Universidad de Zaragoza – España)
& Investigadores Asociados al CCC- BD&ET

Resumen

El Consorcio de I+D+i en *Cloud Computing, Big Data & Emerging Topics* (CCC-BD&ET) [1] es una iniciativa para fomentar y formalizar la colaboración existente entre grupos de investigación de varias universidades en temáticas vinculadas a Cloud Computing, al Análisis Masivo de Datos y a Tópicos Emergentes, como las tecnologías 4.0, entre otros. Estas temáticas, y su integración, han adquirido creciente importancia por su aplicación en dominios de alto impacto como las ciudades inteligentes, la transformación digital, los sistemas de E-health y los basados en tecnologías tales como IoT, IIoT y Blockchain.

Los integrantes del consorcio, provenientes mayoritariamente de Argentina, Chile y España han tenido, a lo largo de los años, diversas experiencias de trabajo conjunto que fueron consolidadas a partir de la organización y realización de las *Jornadas de Cloud Computing-Big Data & Emerging Topics* (JCC-BD&ET) llevadas a cabo en la Universidad Nacional de La Plata (Argentina). La constitución de este Consorcio, reafirma y formaliza estas líneas de colaboración proponiendo acciones de cooperación académica vinculadas con la formación de recursos humanos, la formulación y ejecución de proyectos conjuntos, y la vinculación con

empresas y organismos relacionados con la industria informática, entre otras.

Este trabajo presenta avances del consorcio en el proyecto integrador inicial que tiene como eje la Resiliencia para la Transformación Digital, así como nuevas líneas de cooperación que se han consolidado en 2022.

Palabras clave: Objetivos de Desarrollo Sostenible, Cloud Computing, Big Data, Sistemas Inteligentes, Aplicaciones Sociales, Resiliencia.

Contexto

El Consorcio de I+D+i en *Cloud Computing, Big Data & Emerging Topics* (CCC-BD&ET) resulta de la cooperación, llevada a cabo durante varios años, entre grupos de investigación, desarrollo e innovación de Universidades vinculadas con las realización anual de las *Jornadas de Cloud Computing-Big Data & Emerging Topics* (JCC-BD&ET), organizadas por la Universidad Nacional de La Plata (UNLP), en Argentina.

I. Introducción

Las Jornadas de Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET) son un encuentro anual de intercambio de ideas,

proyectos, resultados científicos y aplicaciones concretas en diferentes áreas relacionadas con Cloud Computing, Inteligencia de Datos, Big Data y Tecnologías Emergentes.

Las Jornadas integran ponencias científicas con experiencias de desarrollos y aplicaciones, fomentando la interacción entre la academia y los sectores productivos/industriales, en el área temática de Cloud Computing, Big Data y Tecnologías Emergentes.

En el marco de las Jornadas también se desarrollan Conferencias, Paneles y Cursos de Posgrado específicos.

Las JCC-BD&ET son organizadas por el Instituto de Investigación en Informática LIDI (III-LIDI) y la Secretaría de Posgrado de la Facultad de Informática de la UNLP en colaboración con Universidades de Argentina y del exterior. Cuentan con diferentes auspicios de organismos de Ciencia y Tecnología de Argentina y de sectores representativos de la industria del Software de Argentina. Se iniciaron en 2013 con foco en los temas de Cloud Computing. A partir del 2015, pasaron a ser las *Jornadas de Cloud Computing & Big Data* y, en el año 2020, se constituyeron en JCC-BD&ET.

A partir de estos encuentros fueron surgiendo distintas actividades en colaboración, que no sólo se formalizaron mediante acuerdos entre los distintos actores. Los resultados de las colaboraciones se encuentran reflejados en las publicaciones detalladas en las páginas mencionadas en las referencias.

Como corolario de las distintas actividades de colaboración realizadas por los participantes a las JCC-BD&ET es que surge la propuesta de conformar un Consorcio de I+D+i en Cloud Computing, Big Data & Emerging Topics que permita afianzar y proyectar a futuro las relaciones existentes [1]. En este contexto de colaboración, las JCC-BD&ET seguirán constituyendo un foco anual de encuentro,

independientemente del trabajo académico y científico que se realice durante el año para la concreción de los objetivos del Consorcio.

Objetivos

Uno de los objetivos iniciales del Consorcio fue formular un proyecto de investigación que integre capacidades de los grupos de investigación que lo componen y que a la vez sea un punto de partida para la formulación de posibles presentaciones conjuntas de financiamiento en áreas relacionadas [2].

Actualmente la mirada se ha ampliado, considerando a las relaciones del Consorcio como un grafo que puede dar lugar a subgrafos de cooperación y proyectos entre parte de los grupos del Consorcio, así como se ha ampliado el núcleo de asociación con otros investigadores y grupos de I+D+I, dando lugar a nuevas áreas de cooperación, en particular en Postgrado.

El Consorcio propone avanzar con una visión prospectiva que dé respuestas a problemas o desafíos estructurales de la sociedad del futuro. De este modo, no sólo se limita a definir un objetivo sino también a delinear una metodología de trabajo que compromete a nuestras áreas de investigación y desarrollo a abordar con rigor científico y con mirada social el camino de la transformación digital que está atravesando nuestra sociedad del siglo XXI.

Esta metodología se enfocará en las posibilidades de incidencia de algunos de los temas transversales de dominio del Consorcio en diferentes retos de los Objetivos de Desarrollo Sostenible (ODS) [3] y en la aplicación de herramientas y estrategias para la resolución de los mismos.

II. Un primer Proyecto Integrador

En 2021 se formuló el Proyecto Integrador: *“Transformación Digital en la incorporación de la Resiliencia como un Key Performance Indicator de Prestaciones Sociales (KPIS)”*, buscando potenciar las capacidades de los grupos que forman el Consorcio y volcarlas en un objetivo concreto.

El contexto lo dan los ODS que conforman una iniciativa de la ONU y se definen como 17 retos cuyo fin es erradicar la pobreza, proteger el planeta y garantizar que todas las personas del mundo, sin distinción, gocen de paz y prosperidad.

La Transformación Digital es el proceso que surge como una alternativa para generar un cambio/innovación positiva en una organización, a partir de la incorporación inteligente de diferentes tecnologías digitales en todos los niveles y funciones de ésta.

Todo modelo de innovación propone como punto de partida el conocimiento y la asimilación de lo que ya se conoce del problema planteado. Luego, en base a las posibles soluciones, se incorpora un proceso de evaluación y comparación de alternativas, que debe hacerse a través de ciertos valores medibles que demuestren efectivamente cómo el sistema está alcanzando los objetivos seleccionados. Este proceso introduce una métrica denominada *“Key Performance Indicator (KPI)”* cuya especificación depende del ámbito de aplicación.

En el proyecto se ha trabajado sobre Índices/Indicadores específicos relacionados con prestaciones sociales (KPIS) y su impacto sobre los Objetivos de Desarrollo Sostenible (ODS).

A modo de ejemplo, la siguiente tabla muestra ámbitos transversales que forman parte de la *expertise* del Consorcio y los retos seleccionados de los ODS en los que puede impactar una transformación digital.

Aplicaciones de interés social	ODS
En Educación	4,5,11 y 16
Gobierno Digital	5,11,15 y 16
Industriales y Agrícolas	9 y 15
En Economía	11,15 y 16
En Salud	4,5 y 11
En Ciudades Inteligentes y Sostenibles	4,11 y 16
Móviles y su evolución a las Web progresivas (WPA)	4,9 y 16

En todos los casos, el Consorcio propone el empleo de diferentes tecnologías con foco en estos objetivos/aplicaciones, dado que una transformación digital no subyace en una única herramienta sino que existen múltiples procesos claves involucrados para afrontar la misma:

Tecnologías de la transformación digital	Herramientas
Datos, Información y Conocimiento	Modelos y Simulación - IA/Inteligencia de Datos Big Data - Análisis Visual - Realidad Extendida - Resiliencia
Computación Avanzada	Cloud Computing – HPC, Arquitectura y Aceleradores - IoT – Edge/Fodge Computing

Finalmente, la metodología presentada en el Proyecto Integrador, propone incorporar la

capacidad del sistema u organización bajo transformación para soportar y sobreponerse a circunstancias de adversidad y desastres, como un KPI de las prestaciones sociales (KPIS).

Este indicador, denominado Resiliencia, evalúa la capacidad de recuperar un estado operativo cuando ha cesado la perturbación a la que había estado sometido.

III. Cooperaciones actuales en I+D+I en el Consorcio

La evolución del Consorcio se ha reflejado en diferentes acciones complementarias al Proyecto Integrador, las cuales se expusieron en el Taller de trabajo de Febrero de 2023, Entre ellas podemos mencionar:

- *IA aplicada a la ingeniería inversa para la certificación medioambiental, búsqueda de índices de sostenibilidad / sustentabilidad.*
- *Administración pública e Inteligencia Artificial. Aplicaciones en Contrataciones del Estado.*
- *Computación de alto desempeño e Inteligencia computacional en la resolución de modelos complejos.*
- *Identificación biométrica en base al análisis de las venas de la mano.*
- *Servicios inteligentes para el ciudadano.*
- *Edge computing y robótica en aplicaciones de tiempo real.*
- *Educación digital, a partir del empleo de nuevas tecnologías.*

IV. Resultados Obtenidos/ Esperados

Los resultados esperados están relacionados con el fortalecimiento de las capacidades de los grupos de investigación del Consorcio, en términos de sus RRHH, redes de trabajo, acceso a recursos organizacionales y mejoramiento de la calidad de los resultados producidos, producto de la colaboración interdisciplinaria e inter-universitaria

promovida por el Consorcio. La definición de un proyecto integrador ordena y permite integrar los aportes de los participantes del proyecto.

Por otra parte en las exposiciones de los responsables de los grupos de I+D+I del Consorcio se mostraron avances significativos en las líneas específicas del proyecto, relacionadas con resiliencia en diferentes sistemas y aplicaciones.

En particular la UAB y la UNSL presentaron una evolución conceptual del proyecto integrador y el III-LIDI mostró la incorporación de los temas del Proyecto integrador en diferentes áreas de las líneas de investigación del Instituto, así como en trabajos de alumnos avanzados de Informática.

V. Formación de RRHH

En lo concerniente a la formación de recursos humanos los distintos integrantes están abocados a la formación de recursos humanos tanto a nivel de grado como de posgrado, en temáticas afines al Consorcio y al Proyecto en particular que en muchos casos se realiza en colaboración entre distintos grupos participantes.

A modo de ejemplo, en el Taller de Febrero de 2023 se mencionaron:

- *1 Tesis de Maestría y Doctorado conjunto UCM y UNLP.*
- *4 Tesis de Doctorado con colaboraciones UAB, UNLP y alumnos de UNCOMA, UNJU y UNGS.*
- *2 Tesis doctorales UCLM con UTN y UNLP.*
- *1 Tesis doctoral en co-tutela URV y UNLP.*
- *Tesis de Maestría y Doctorado con direcciones UNS, UNSL y UNLP-.*
- *Especializaciones UNLP y UN Chilecito.*

Por otro lado, fruto de las interacciones entre investigadores en diferentes temas, han surgido cursos e incluso nuevas carreras de

Postgrado que tienen intervención de grupos de I+D+I del Consorcio.

Entre ellas:

- *Maestría en Gestión y Tecnología de Ciudades Inteligentes.*
- *Especialización en BioInformática.*
- *Posible Maestría en Transformación Digital.*

Asimismo los grupos tienen reuniones de intercambio de posibilidades de apoyo en actividades de postgrado para alumnos de Maestría y Doctorado.

Esto incluye financiamiento para estadías de investigación y dictado de cursos abiertos a las Universidades del Consorcio, así como la definición de Ciclos y Estadías Postdoctorales.

VI. Integrantes del Consorcio

El Consorcio está integrado tanto por grupos de investigación como por investigadores asociados. A continuación se detallan los mismos.

Grupos que conforman el CCC- BD&ET

III-LIDI – Instituto de Investigación e Informática LIDI (Universidad Nacional de La Plata – Argentina) [4]
Directora: *Lic. Patricia Pesado*

LISSI – Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información (Universidad Nacional del Sur – Argentina) [5]
Director: *Dr. Pablo Fillottrani*

VyGLab – Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (Universidad Nacional del Sur – Argentina) [6]
Directora: *Dra. Silvia Castro*

LIDIC – Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (Universidad Nacional de San Luis – Argentina) [7]
Director: *Dr. Marcelo Errecalde*

HPC4EAS – High Performance Computing for Efficient Applications and Simulation (Universidad Autónoma de Barcelona – España) [8]
Director: *Dr. Emilio Luque*

SMILe – Soft Management of Internet and Learning (Universidad de Castilla-La Mancha – España) [9]
Director: *Dr. José A. Olivas Varela*

ArTeCS – Group of Architecture and Technology of Computing Systems (Universidad Complutense de Madrid – España) [10]
Director: *Dr. Francisco Tirado*

LITRP – Laboratorio de Investigaciones Tecnológicas en Reconocimiento de Patrones (Universidad Católica de Maule – Chile) [11]
Representante para el Consorcio: *Dr. Ricardo Barrientos*

LCG – Laboratorio de Computación Gráfica (Universidad Nacional de San Luis – Argentina) [12]
Director: *Mg. Roberto Guerrero*

ITIC – Instituto para las Tecnologías de la Información y las Comunicaciones (Universidad Nacional de Cuyo – Argentina) [13].
Director: *Dr. Carlos García Garino*

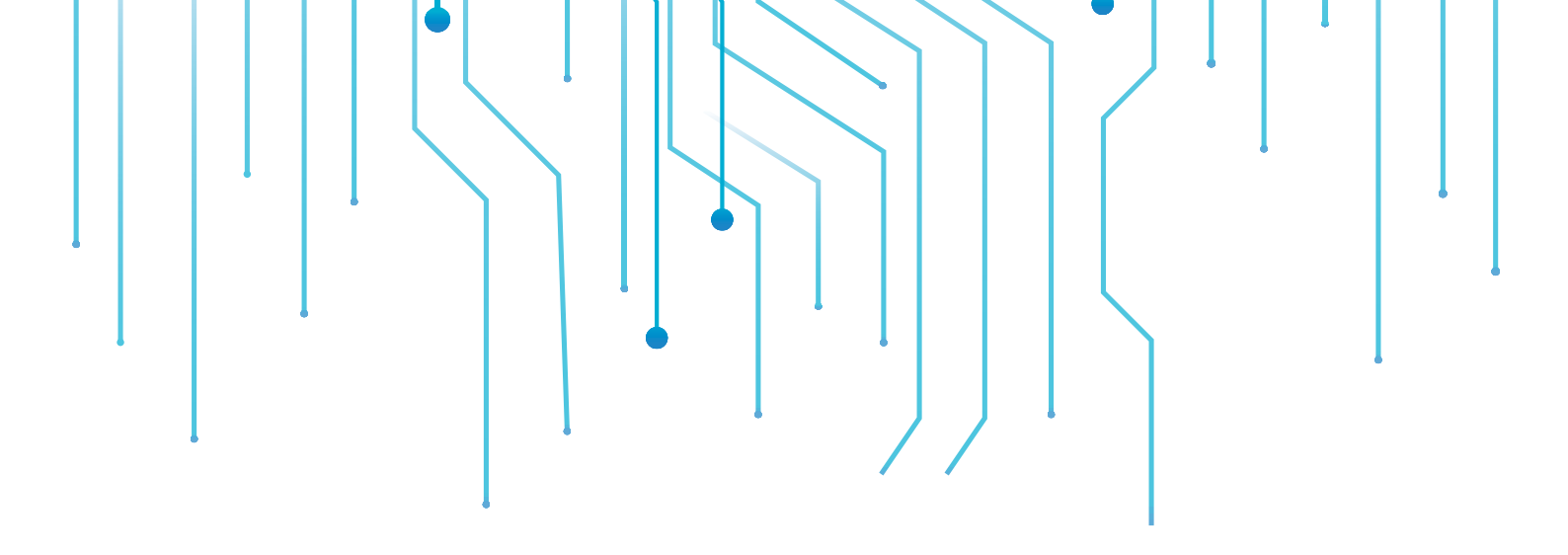
DisCo – Grupo de I+D+I en Computación Distribuida (Universidad de Zaragoza – España) [14].
Coordinador: *Dr. Pedro Javier Álvarez Pérez-Arados*

Investigadores Asociados al CCC- BD&ET

Dr. Aurelio Fernández (Universidad Rovira i Virgili – España).
Dr. Emmanuel Frati (Universidad Nacional de Chilecito – Argentina).
Dr. Javier Balladini (Universidad Nacional de Comahue – Argentina).
Dra. Adriana Gaudiani (Universidad Nacional de General Sarmiento – Argentina).
Dra. Mónica Denham (Universidad Nacional de Río Negro – Argentina).
Lic. Nelson Rodríguez (Universidad Nacional de San Juan – Argentina).

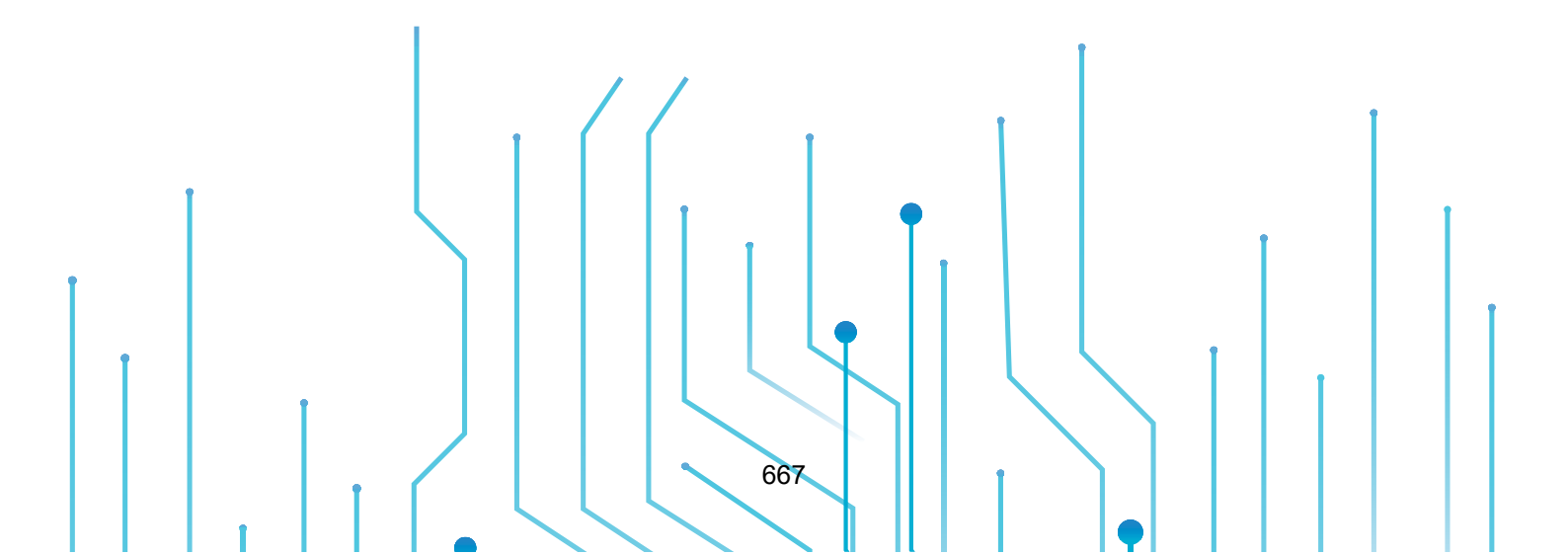
Referencias

- [1] Consorcio de I+D+I en *Cloud Computing, Big Data & Emerging Topics* (CCC-BD&ET), <https://jcc.info.unlp.edu.ar/consorcio-de-idi-en-cloud-computing-big-data-emerging-topics/>.
- [2] Consorcio para la colaboración en I+D+I en temas de Cloud Computing, Big Data y Emerging Topics (CCC-BD&ET)- WICC 2021. 978-987-24611-3-3; 978-987-24611-4-0. Páginas: 778-783.
- [3] La UNESCO y los Objetivos de Desarrollo Sostenible: <https://es.unesco.org/sdgs>
- [4]<http://weblidi.info.unlp.edu.ar>
- [5]<https://lissi.cs.uns.edu.ar>
- [6]<http://vyglab.cs.uns.edu.ar>
- [7]<http://lidic.unsl.edu.ar>
- [8] <https://grupsderecerca.uab.cat/hpc4eas/>
- [9]<http://smile.esi.uclm.es>
- [10]<https://artecs.dacya.ucm.es>
- [11]<http://www.litrp.cl>
- [12]<http://www.lcg.unsl.edu.ar/>
- [13] <http://itic.uncu.edu.ar/>
- [14] <http://webdiis.unizar.es/DISCO/>



SI

Seguridad Informática



Estudio de las propiedades criptológicas de los Tokens No Fungibles Inteligentes para asegurar dispositivos de Internet de las Cosas

Jorge Eterovic; Marcelo Cipriano; Edith García; Luis Torres;

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{marcelo.cipriano; jorge.eterovic; edith.garcia; luisantonio.torres}@usal.edu.ar

RESUMEN

El Token No Fungible (NFT) es una solución creada para permitir representar objetos con cualidades únicas, irrepetibles e indivisibles dentro de una blockchain. Es un activo digital que representa objetos del mundo real como arte, música y videos. Se compran y venden en línea con criptomonedas y además se construyen utilizando el mismo tipo de programación, como Bitcoin o Ethereum.

Los datos únicos de los NFT facilitan la verificación de su propiedad y la transferencia de tokens entre propietarios. Un token no fungible inteligente (Smart NFT), podría representar dispositivos de Internet de las cosas (IoT), que son activos físicos inteligentes y lo identificaría como de propiedad de un usuario. Le asignaría una Dirección de Cuenta de Blockchain (BCA) para participar activamente en las transacciones pudiendo establecer canales de comunicación seguros entre propietarios y usuarios y operar de manera segura.

Con esta solución, los dispositivos IoT podrían probar no solo que su hardware es de confianza sino también su software, porque ejecutarían un arranque seguro y llevarían a cabo procesos de autenticación mutua con los propietarios y usuarios.

Este proyecto de investigación se centra en analizar las posibilidades de asegurar los datos de los dispositivos IoT usando los Smart NFT.

Palabras Clave:

Seguridad en IoT. Blockchain. Contrato Inteligente. Ethereum. Tokens No Fungibles.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndola como un servicio a la comunidad y entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias Unidades Académicas se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándolos a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e internacionales, como así también, apoyo y orientación de recursos para la investigación.

A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto denominado “Estudio de las propiedades criptológicas de los Tokens No Fungibles Inteligentes para asegurar dispositivos de Internet de las Cosas.”, con una duración de 2 años (2023-

2024) y que ya ha sido evaluado y aprobado para su realización.

1. INTRODUCCIÓN

Si buscamos una tecnología que impactará y beneficiará nuestras vidas en los próximos años, es el Internet de las cosas. Los automóviles, electrodomésticos, teléfonos inteligentes, medidores de servicios públicos, sensores incorporados al cuerpo, indumentaria y casi cualquier cosa que podamos imaginar estarán conectados a Internet y serán accesibles desde cualquier parte del mundo [1]. La revolución que generará IoT será inigualable, algunos autores dicen que será similar a la construcción de carreteras y ferrocarriles que impulsaron la Revolución Industrial de los siglos XVIII al XIX [2], y será transversal a todos los sectores de la sociedad y todas las industrias, desde educación, salud, hogar y ciudad inteligente, hasta manufactura, minería, comercio, logística y vigilancia, solo por mencionar algunas [3].

Internet de las cosas implica directa o indirectamente la generación de cantidades significativas de información y un grupo dinámico de partes interesadas con distintos niveles de derechos de acceso a ella. Además, el alcance de la información relacionada con IoT variará según los requisitos del dominio de la aplicación y el contexto de los dispositivos.

Las aplicaciones de IoT involucran a muchas partes interesadas, con diferentes roles y funcionalidades que acceden a distintos tipos de información con distintos niveles de acceso, múltiples identidades y condiciones particulares de seguridad para cada una de ellas. Administrar todos estos activos de manera eficiente, segura e interoperable es un problema desafiante. Se analizará si la tecnología Blockchain y los contratos inteligentes de NFT pueden desempeñar un papel importante en este sentido [4].

Una cadena de bloques mantiene una colección, o libro mayor, de transacciones de manera descentralizada y distribuida. El libro mayor es inmutable e irreversible, lo que significa que las transacciones pasadas no

pueden ser modificadas por ninguna entidad que registre transacciones en la Blockchain, y se comparte y sincroniza en todos los nodos participantes. De esta manera, la cadena de bloques garantiza que el libro mayor no puede ser manipulado, y que todos los datos que posee la Blockchain son confiables [5].

Una cadena de bloques puede ser pública [6] o estar restringida solo a usuarios autorizados [7]. La Blockchain se considera una forma democrática de mantener transacciones [8] y se prevé que proporcione mecanismos de seguridad novedosos, que contribuyan a la sostenibilidad de las aplicaciones de IoT y permitan nuevos modelos de confianza [9].

Un contrato inteligente es una aplicación distribuida que vive en la cadena de bloques [6]. Esta aplicación es, en esencia, una clase de lenguaje de programación con campos y métodos. Los usuarios pueden interactuar con los campos y métodos públicos de esta clase enviando transacciones a su dirección en la cadena de bloques.

Cada vez que un usuario interactúa con un contrato inteligente, todos los nodos de la red Blockchain ejecutan todas las operaciones de manera determinista y confiable y uno de estos nodos se selecciona para almacenar el resultado de la ejecución de los contratos, si corresponde, en la cadena de bloques. Los contratos inteligentes pueden verificar las identidades y firmas digitales de los usuarios de la Blockchain, realizar cálculos de propósito general e invocar a otros contratos [10].

El código de un contrato inteligente es inmutable y no puede ser modificado ni siquiera por su propietario [11]. Además, todas las transacciones enviadas a un contrato se registran en la cadena de bloques, por lo que es posible obtener todos los valores históricos de una variable del contrato.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación propone analizar las posibilidades de asegurar los datos de los

dispositivos de Internet de las Cosas usando los Tokens No Fungibles Inteligentes.

Internet de las cosas (IoT) es el paradigma en el que cualquier cosa (dispositivos, objetos, sujetos, etc.) puede interconectarse a través de Internet con la capacidad de interactuar, recopilar, procesar y compartir datos de manera inteligente. El rápido crecimiento en la cantidad de cosas conectadas a Internet ha llevado a la necesidad de buscar soluciones de seguridad.

La mayoría de las soluciones existentes se centran en el manejo y el intercambio seguro de datos, afrontando riesgos que van desde el uso de datos personales sin el consentimiento o conocimiento del propietario hasta el acceso o manipulación de datos por parte de terceros no autorizados [1].

Las tecnologías Blockchain se han utilizado recientemente con IoT ya que proporcionan una cadena de bloques de datos distribuida y criptográficamente segura, lo que permite una trazabilidad de los datos inmutable, que garantiza la propiedad de los mismos y la privacidad del usuario [2]. Cada bloque de una cadena de bloques se identifica unívocamente y se vincula con el anterior mediante una función hash criptográfica. Se agrega un nuevo bloque a la cadena si los participantes en la cadena de bloques con el rol de mineros demuestran que el nuevo bloque es seguro y la mayoría de los mineros lo validan, aplicando un algoritmo de consenso.

Si un atacante intenta cambiar los datos de un bloque, por ejemplo, los datos capturados por los sensores de un dispositivo IoT, el hash del bloque cambia y los datos del siguiente bloque no coinciden. El atacante tendría que cambiar todos los bloques posteriores de la cadena para no ser detectado [3].

Los datos en los bloques pueden ser datos capturados por los dispositivos IoT, así como datos sobre transacciones y los participantes involucrados en las transacciones. El intercambio de moneda digital o activos genéricos generalmente involucra transacciones entre participantes de blockchain. En cadenas de bloques públicas,

como Ethereum, los acuerdos entre los participantes se formalizan a través de contratos inteligentes representados por scripts.

Estos scripts son validados como parte de las transacciones por el algoritmo de consenso y, por lo tanto, una vez validados, los contratos inteligentes son inviolables como los demás datos registrados en la cadena de bloques [4]. Además, se aplican protocolos para lograr la ejecución segura de los contratos inteligentes [5]. Todos los participantes ejecutan el código de un contrato inteligente de la misma manera y pueden comprobar si se cumplen las condiciones establecidas por el mismo.

Por lo tanto, además de la integridad, blockchain permite la transparencia porque cualquier participante puede consultar los datos registrados, incluidos los contratos inteligentes. Esto se ha aprovechado para integrar blockchain con la gestión de la cadena de suministro, lo que, en el caso del ecosistema IoT, permite garantizar el origen y la confiabilidad de los datos sin necesidad de intermediarios [6]. Los dispositivos IoT pueden ser totalmente autónomos para intercambiar directamente sus datos con terceros a través de contratos inteligentes que rigen todas las políticas de comercio de datos y derechos de propiedad [7].

La información de propiedad se puede incluir en contratos inteligentes por medio de tokens. Dado que un token es la representación digital de un activo en la cadena de bloques, existen dos tipos principales de tokens, tokens fungibles y no fungibles (NFT), según el activo representado.

Los tokens fungibles son tokens idénticos e intercambiables, como una moneda fiduciaria, que permiten transacciones de contabilidad y facturación. Los tokens no fungibles son tokens únicos y no intercambiables, como instrumentos notariales o coleccionables de obras de arte, que permiten la trazabilidad de posesiones físicas o no físicas únicas. Un ejemplo de posesión no física en el contexto de IoT es la capacidad de acceder a recursos o servicios.

Un dispositivo IoT que actúe como proveedor de datos puede representar mediante NFT los recursos o servicios proporcionados. [8, 9]. Otro uso de los NFT es para representar productos físicos. Cada producto fabricado por un servicio en la nube puede representarse mediante un NFT, de modo que su precio, procedencia y propiedad pueden rastrearse mediante una cadena de bloques [10].

Aplicando una tokenización más fina, cada componente de un producto puede representarse mediante un NFT de modo que la creación de un producto a partir de sus componentes se registre en la cadena de bloques como la creación de un nuevo token a partir de los tokens que representan sus componentes. De esta forma, no solo se puede rastrear el origen sino también la posterior transformación de un producto [11].

La tecnología blockchain se aplica no solo para garantizar la confiabilidad de la fabricación del dispositivo de hardware, sino también para garantizar la seguridad de su software. De lo contrario, toda la seguridad falla [12]. En cuanto al software, se presta especial atención al arranque seguro del microcontrolador en el hardware del dispositivo IoT, cuyo objetivo es comprobar que el código a ejecutar es el esperado, y a la comunicación segura con propietarios y usuarios, que pueden llegar a cambiar el software del dispositivo. En éste trabajo de investigación usaremos la cadena de bloques Ethereum ya que es una de las cadenas de bloques públicas más extendidas.

En la comunidad Ethereum, se desarrollaron los estándares ERC (Ethereum Request for Comments) [13], el estándar ERC-20 define los tokens fungibles y el estándar ERC-721 que define los tokens no fungibles. Un atributo importante de un ERC-721 NFT es su propietario, que se identifica únicamente por su Dirección de Cuenta de Cadena de Bloques (BCA). Un participante de la cadena de bloques interactúa con la cadena de bloques a través de una BCA, que se compone de un par de claves criptográficas (pública y privada) y una dirección única derivada de la clave pública. Además de un propietario, hay un

identificador único que está asociado con un ERC-721 NFT.

En las propuestas donde los tokens ERC-721 representan bienes físicos, los identificadores de los tokens suelen estar relacionados con los atributos lógicos del producto, como su código de barras o QR [10], o a los datos de información del producto, como las instrucciones de eliminación y las fechas de caducidad [11]. El problema es que estos identificadores no están relacionados con algo intrínseco del producto, es decir, no son verdaderos identificadores de producto ya que pueden ser modificados, copiados o transferidos a otro producto.

Para evitar este problema, se propusieron los anclajes criptográficos como un vínculo estrecho entre los dominios digital y físico [14]. Las Funciones Físicas no Clonables (PUF) de silicio, que representan las variables del proceso de fabricación de los semiconductores [15], pueden tomarse como anclajes criptográficos para productos electrónicos, como los dispositivos IoT.

El enlace de ERC-721 NFT con dispositivos IoT físicos que usan PUF se propuso en un trabajo presentado en la Conferencia Internacional sobre Criptografía Aplicada y Seguridad de Redes realizada en Alemania en el año 2020 [16]. Ese trabajo incluyó, como un nuevo atributo en ERC-721 NFT, la dirección de cuenta de blockchain (BCA) asociada con el dispositivo IoT. Se propuso el uso de PUF en el hardware del dispositivo IoT para reconstruir la clave privada de la que se deriva la dirección BCA del dispositivo en su token asociado.

Además, ese trabajo también incluyó, como un nuevo atributo en ERC-721 NFT, la dirección de la cuenta blockchain (BCA) asociada con el usuario del dispositivo IoT, de modo que no solo se puede rastrear la propiedad sino también el uso del dispositivo mediante la cadena de bloques con el mismo token.

Sin embargo, ese trabajo no permitió detectar si el dispositivo no está funcionando correctamente, si el vínculo entre el dispositivo y el NFT está roto, o si el compromiso con el

propietario y el usuario se pierde en algún momento. Para evitar estas fallas de seguridad, un trabajo reciente presenta un nuevo NFT, llamado NFT inteligente (Smart NFT). [17].

Este proyecto de investigación se centra en la búsqueda, estudio y análisis de distintas posibilidades de combinar los dispositivos de Internet de las cosas (IoT) con las tecnologías blockchain de NFT para hacer que los dispositivos IoT sean seguros, desde el punto de vista del hardware y del software, de modo que los datos que proporcionan sean confiables.

3. RESULTADOS OBTENIDOS/ESPERADOS.

El Objetivo General del proyecto de investigación es analizar las distintas posibilidades de asegurar los datos de los dispositivos de Internet de las cosas (IoT) usando la tecnología blockchain de los Tokens No Fungibles (NFT).

Los Objetivos Específicos son relevar y Estudiar en detalle la tecnología Blockchain de los Tokens No Fungibles; analizar las propuestas de soluciones que garanticen la confiabilidad del hardware y del software de los dispositivos IoT desde su fabricación hasta la operación del usuario final; analizar como los tokens no fungibles inteligentes (Smart NFT), podrían representar dispositivos de Internet de las cosas (IoT) y proponer soluciones usando Smart NFT para asegurar los dispositivos IoT.

4. FORMACIÓN DE RECURSOS HUMANOS.

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas de la Facultad de Ingeniería, específicamente al área de Seguridad Informática, de la Universidad del Salvador.

A este proyecto, se incorporaron tres docentes investigadores con amplia trayectoria académica y un docente investigador con muchos años de desempeño en la industria de TI.

Esto redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes investigadores, como así también conformará las bases para la investigación a futuro desarrollando una línea de trabajo en la Facultad de Ingeniería.

5. BIBLIOGRAFÍA.

- [1] Pal, S.; Rabehaja, T.; Mukhopadhyay, S. Requisitos de seguridad para Internet de las cosas: un enfoque sistemático. *Sensores* 2020, 20, 5897.
- [2] Novo, O. Blockchain se encuentra con IoT: una arquitectura para la gestión de acceso escalable en IoT. *IEEE Internet Things J.* 2018, 5, 1184–1195.
- [3] Dhananjay, S.; Jong-Hoon, K.; Madhusudan, S. *Tecnologías de cadena de bloques*; Springer Nature: Nueva York, NY, EE. UU., 2020.
- [4] Longo, R.; Podda, AS; Saia, R. Análisis de un protocolo de consenso para extender subcadenas consistentes en la cadena de bloques de Bitcoin. *Cómputo* 2020, 8, 67.
- [5] Sol, T.; Yu, W. Un marco de verificación formal para cuestiones de seguridad de los contratos inteligentes de Blockchain. *Electronica* 2020, 9, 255.
- [6] Al-Rakhami, MS; Al-Mashari, M. Un modelo de confianza basado en blockchain para la gestión de la cadena de suministro de Internet de las cosas. *Sensores* 2021, 21, 1759.
- [7] Nawaz, A.; Peña Queralta, J.; Guan, J.; Awais, M.; Gia, TN; Bashir, Ak; Kan, H.; Westerlund, T. Edge Computing para asegurar la propiedad de datos de IoT y el comercio con Ethereum Blockchain. *Sensores* 2020, 20, 3965.
- [8] Omar, AS; Basor, O. Enfoque de tokens no fungibles basado en capacidad para un marco AAA descentralizado en IoT. *Ciberseguridad de cadena de bloques. Confianza. priv.* 2020, 79, 7–31.
- [9] Fotiou, N.; Pittaras, I.; Siris, VA; Siris, VA; Voulgaris, S.; Polyzos, autorización de GC

OAuth 2.0 usando tokens basados en blockchain. arXiv 2020, arXiv:2001.10461.

[10] Hasan, M.; Starly, B. Arquitectura de plataforma descentralizada de fabricación en la nube como servicio (CMaaS) con activos digitales configurables. *J. Manuf. sist.* 2020, 56, 157–174.

[11] Westerkamp, M.; Víctor, F.; Küpper, A. Seguimiento de los procesos de fabricación utilizando composiciones de fichas basadas en blockchain. *Dígito. común Neto.* 2020, 6, 167–176.

[12] Jesús, V. Raíces de confianza mejoradas con Blockchain. En *Actas de la Conferencia internacional IEEE sobre redes y comunicaciones inteligentes (SmartNets)*, Yasmine Hammamet, Túnez, 16 y 17 de noviembre de 2018; pp. 1 a 7.

[13] Propuestas de mejora de Ethereum. Disponible en línea: <https://eips.ethereum.org/erc> (consultado en abril de 2022).

[14] Balagurusamy, VS; Cabral, C.; Coomaraswamy, S.; Coomaraswamy, S.; Delamarche, E.; Dillenberger, D.; Friedman, D.; Gökce, O.; Hinds, N.; Jelito, J.; et al. Anclajes criptográficos. *IBM J.Res. desarrollo* 2019, 63, 4:1–4:12.

[15] Gao, Y.; Al-Sarawi, SF; Abbott, D. Funciones físicas no clonables. *Nat. Electrón.* 2020, 3, 81–91.

[16] Arcenegui, J.; Arjona, R.; Baturone, I. Gestión segura de dispositivos IoT basada en tokens no fungibles de cadena de bloques y funciones físicas no clonables. En *Actas de la Conferencia Internacional sobre Criptografía Aplicada y Seguridad de Redes*; Springer: Berlín/Heidelberg, Alemania, 2020; Volumen 12418, págs. 24–40.

[17] Javier Arcenegui; Rosario Arjona; Roberto Román; Iluminada Baturone. Secure Combination of IoT and Blockchain by Physically Binding IoT Devices to Smart Non-Fungible Tokens Using PUFs. *Sensors* 2021, 21(9), 3119.

Análisis de las Herramientas para Realizar Pruebas Estáticas de Seguridad de las Aplicaciones

**Jorge Eterovic; Valeria Silvestri; Andrea Vera; Martin Zeballos;
Alesio Esteban Sinopoli**

Departamento de Ingeniería e Investigaciones Tecnológicas Universidad
Nacional de La Matanza
Florencio Varela 1903 (B1754JEC), San Justo, (5411) 4480-8900

{eterovic; vsilvestri; avera; mzeballos}@unlam.edu.ar;
asinopoli@alumno.unlam.edu.ar

RESUMEN

Las pruebas estáticas de seguridad de las aplicaciones que se utilizan para proteger el software se denominan SAST (Static Application Security Testing) y consisten en la revisión automática del código fuente para identificar patrones vulnerables.

Las herramientas SAST permiten automatizar la detección de vulnerabilidades y se pueden integrar al sistema de CI/CD (integración continua / distribución continua) para que detecten vulnerabilidades en etapas tempranas del ciclo de vida. Esto ayuda al equipo de Seguridad de las Aplicaciones a implementar un ciclo de vida del desarrollo de software seguro.

CI/CD es un método para distribuir las aplicaciones a los clientes mediante el uso de la automatización en las etapas del desarrollo de las aplicaciones. En este contexto, cualquier equipo de Seguridad de las Aplicaciones se enfrentará al desafío de automatizar los chequeos de seguridad y encontrará la solución en herramientas SAST.

Integrar una herramienta SAST al proceso de CI/CD permite detectar vulnerabilidades en la etapa de desarrollo, en vez de esperar a la etapa de prueba o que se detecten directamente en producción.

Hay disponibles herramientas SAST gratuitas para los repositorios open-source y pagas para los repositorios privados. Algunas son open-source, otras usan un motor privado pero las reglas son open-source y algunas pocas son totalmente privadas.

Este proyecto de investigación propone hacer un análisis de estas herramientas SAST y aportar

los resultados obtenidos a la comunidad open-source para mejorar la seguridad de los repositorios de proyectos de desarrollo de software que las utilizan.

Palabras Clave: *Herramienta SAST; Detección de Vulnerabilidades; Análisis Estático; Seguridad de las Aplicaciones..*

CONTEXTO

Este proyecto de investigación se desarrolla en el marco de un Programa de Incentivos a Docentes Investigadores de la Secretaría de Políticas Universitarias (PROINCE) del Ministerio de Educación, y se ejecuta en el Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza.

El proyecto es financiado por el propio Departamento y es del tipo investigación aplicada. El mismo propone hacer un análisis de las herramientas SAST y aportar los resultados obtenidos a la comunidad open-source. Los trabajos de campo y relevamientos realizados aportaron información valiosa y sirvieron como base para el presente trabajo.

1. INTRODUCCIÓN

Las herramientas SAST permiten automatizar la detección de vulnerabilidades y se pueden integrar al sistema de CI/CD para que detecten vulnerabilidades en etapas tempranas del ciclo de vida del desarrollo del software. Esto ayuda al equipo de Seguridad de las Aplicaciones a implementar un ciclo de vida seguro.

CI/CD es un método para distribuir las aplicaciones a los clientes mediante el uso de la automatización en las etapas del desarrollo de las aplicaciones. Los principales conceptos que se le atribuyen son la integración, la distribución y la implementación continuas. Se trata de una solución para los problemas que la integración de código nuevo puede generar a los equipos de desarrollo y de operaciones.

El proceso de integración y distribución continuas incorpora la automatización y la supervisión permanentes en todo el ciclo de vida de las aplicaciones, desde las etapas de integración y prueba hasta las de distribución e implementación. En este contexto, cualquier equipo de Seguridad de las Aplicaciones se enfrentará al desafío de automatizar los chequeos de seguridad y encontrará la solución en herramientas SAST.

Integrar una herramienta SAST al proceso de CI/CD permite detectar vulnerabilidades en la etapa de desarrollo, en vez de esperar a la etapa de prueba o que se detecten directamente en producción. Una vulnerabilidad en producción implica un riesgo constante, cuesta mucho esfuerzo de los expertos en seguridad detectarla y para los desarrolladores es difícil de corregir. En cambio, si se detecta durante la etapa de desarrollo, nunca generó un riesgo real, no requirió esfuerzo de personas de seguridad para detectarla y es mucho más fácil de corregir.

Hay muchas herramientas SAST y el análisis estático está muy vinculado al lenguaje de programación. Una herramienta puede analizar varios lenguajes, pero en realidad agregar un lenguaje nuevo a la herramienta, es desarrollar un producto nuevo. Es decir, una misma herramienta va a tener distintos grados de madurez, distintas características y distintas limitaciones según el lenguaje a analizar.

Un criterio de comparación entre distintas herramientas SAST es el grado de complejidad. Una herramienta simple se ejecuta rápidamente, soporta código que no compila y es fácil de aprender a usar, pero no es precisa, da muchos falsos positivos y falsos negativos incorregibles porque no maneja la información necesaria para refinar los resultados.

Por otro lado, una herramienta compleja se ejecuta más lentamente, tiene requisitos extras como por ejemplo que el código sea compilable y

completo y lleva tiempo aprender a usarla, pero a su vez como maneja mucha más información y esta se puede usar para evitar falsos positivos y falsos negativos.

Hay bastante variedad entre las herramientas SAST. Algunas son open-source, otras usan un motor privado pero las reglas son open-source y algunas pocas son totalmente privadas. Algunas incluso distinguen su sistema de precios según el código a analizar, siendo gratuitas para los repositorios open-source y pagas para los repositorios privados. Y por último hay sistemas que simplemente se encargan de integrar y ofrecer varias herramientas. Por ejemplo, GitHub tiene una sección de escaneo de código denominada “Code Scanning” con muchas herramientas SAST.

2. LÍNEAS DE INVESTIGACION Y DESARROLLO

Se desarrollan a continuación los aspectos teóricos de este proyecto de investigación:

- DevOps
- DevSecOps
- Pruebas Estáticas de la Seguridad de la Aplicación (SAST)
- Integración Continua y Despliegue Continuo (CI/CD)

DevOps

Existen muchas definiciones diferentes de DevOps disponibles en libros, en artículos de revistas o en Internet. A raíz de esta disparidad en las definiciones, surgen diversos estudios que tratan de darle una descripción académica [1] [2]. Según estos estudios, podemos definir DevOps como una cultura que trata de aunar a los equipos de desarrollo y operaciones, basándose en una serie de principios y prácticas [2] que pretenden acelerar las entregas del producto mejorando el feedback de los clientes y la capacidad de reacción ante los cambios [3].

El término DevOps surge a finales de los 2000, en un contexto en el que las metodologías de desarrollo ágiles cada vez tomaban mayor relevancia en la industria del desarrollo de software. La velocidad a la que se desarrollaban nuevas características o se corregían bugs distaba mucho de la velocidad a la que se realizaban los despliegues de estos cambios, con lo que el ciclo

de desarrollo del software se veía ralentizado. Esta ralentización era resultado de la falta de comunicación existente entre el equipo de desarrollo y el de operaciones.

Fue Patrick Debois quien, en 2007, tras una experiencia frustrante trabajando en la migración de un gran centro de datos, se percató de cómo esta falta de comunicación entre desarrolladores y administradores de sistemas afectaba al flujo de trabajo [4]. En 2009, John Allspaw y Paul Hammond, ingenieros de Flickr, presentaron su charla "10 Deploys a Day: Dev and Ops Cooperation at Flickr" [5] donde propusieron integrar desarrollo y operaciones en un flujo automatizado. Patrick Debois, tras esta conferencia, decidió organizar una similar en Bélgica, a la que llamó DevOpsDays, de donde surge el término DevOps [6].

Un aspecto importante para seguir exitosamente la cultura DevOps es la automatización de procesos, al ser lo que permite mantener la agilidad durante el desarrollo del software. La importancia de la automatización de procesos ha hecho surgir una gran cantidad de herramientas que contemplan la construcción del software, la integración y el despliegue continuos, la gestión de logs y la monitorización [7].

DevSecOps

El crecimiento de las metodologías ágiles de desarrollo del software y la acogida de la cultura DevOps por parte de las organizaciones ha incrementado la velocidad a la que las aplicaciones reciben actualizaciones. Esto tiene grandes ventajas, pues permite tener feedback temprano del cliente para mejorar el producto desde las primeras fases del desarrollo, mejorando la adaptabilidad del producto con el entorno y permitiendo marcar la diferencia con la competencia gracias a la implementación de nuevas funcionalidades y la mejora del funcionamiento de las ya existentes [8]. Sin embargo, a veces esta agilidad se consigue a costa de sacrificar otros aspectos del producto final, como puede ser la seguridad [9]. Los estudios muestran que menos de un 20% de las compañías que siguen la cultura DevOps tienen en cuenta la seguridad como parte del ciclo de desarrollo del software [5].

DevSecOps surge para incluir la seguridad en DevOps, alineando los equipos de desarrollo, de

operaciones y de seguridad durante todo el ciclo de desarrollo. Esto se consigue desplazando la seguridad a la izquierda, es decir, considerándola desde las primeras etapas del desarrollo [9]. De esta manera, al tenerla en cuenta desde el diseño de la aplicación, es posible realizar los controles de seguridad necesarios a lo largo del ciclo de desarrollo del software y automatizarlos para que sean rápidos, escalables y efectivos [10]. De esta manera se mantiene la agilidad en el desarrollo del software y se detectan desde fases tempranas los fallos de seguridad que, de llegar al cliente, conllevarían grandes pérdidas de tiempo y dinero.

Al igual que en DevOps, las herramientas juegan un papel de gran importancia. Existen una gran cantidad de herramientas para llevar a cabo DevSecOps. Se ha tomado como referencia la guía OWASP DevSecOps Guideline [11] para establecer los aspectos más importantes relativos a la seguridad: la detección de secretos, las Pruebas Estáticas de la Seguridad de la Aplicación (SAST), las Pruebas Dinámicas de la Seguridad de la Aplicación (DAST), el escaneo de la infraestructura y la comprobación del cumplimiento normativo.

Pruebas Estáticas de la Seguridad de la Aplicación (SAST)

Las pruebas estáticas de seguridad analizan el código fuente de la aplicación sin ejecutarlo, tratando así de encontrar vulnerabilidades o bugs [12]. Los análisis estáticos se pueden realizar de diferentes formas, desde las más sencillas y rápidas que contemplan sólo un análisis del código fuente en base al árbol sintáctico, hasta las más complejas que combinan diversas representaciones del código como grafos de control y flujo de datos para realizar un análisis semántico en busca de patrones vulnerables [12].

Los factores a tener en cuenta a la hora de decantarse por una herramienta u otra son la velocidad a la que se quiere realizar el análisis y la profundidad del mismo, pues a más profundidad, más se tardará en realizar y viceversa. Además, se ha de considerar el porcentaje de falsos positivos que puede señalar la herramienta y el lenguaje de programación de la aplicación. Las herramientas SAST ayudan a detectar vulnerabilidades como inyecciones SQL, cross-site scripting (XSS) y problemas de gestión de memoria, entre otras.

Algunos ejemplos de estas herramientas son: Semgrep, CodeSonar o CodeQL[13] [14].

Integración Continua y Despliegue Continuo (CI/CD)

La integración continua (CI) surge como una de las prácticas de la metodología ágil Extreme Programming (XP), en la cual se propone que los desarrolladores publiquen sus cambios varias veces al día en el repositorio de código. De esta forma pueden encontrarse problemas de compatibilidad entre estos cambios en etapas más tempranas del desarrollo, y se evitan complejos y largos procesos de integración en los días anteriores de la fecha de entrega del proyecto o del hito [15]. Gracias a estas ventajas, la integración continua se utiliza como práctica independiente de XP, respaldada por referentes en el mundo del desarrollo del software como Martin Fowler [16].

El despliegue continuo (CD) amplía las bases propuestas por la integración continua, proponiendo no solo la integración automatizada del código en el repositorio, sino también el despliegue automatizado del mismo en el entorno de producción [17]. La automatización del despliegue es especialmente beneficiosa cuando existen varios entornos en los que se ha de desplegar el software cuando se genera una nueva versión, y cuando este proceso de despliegue ocupa mucho tiempo [18]. Cabe clarificar las diferencias entre la entrega continua y el despliegue continuo, conceptos que en ocasiones se confunden. Mientras la entrega continua tiene como objetivo mantener siempre el software en un estado que permite su despliegue inmediato [19], el despliegue continuo implica el despliegue de las nuevas versiones del software de forma automática.

Los procesos automatizados de integración y de despliegue pueden tardar varios minutos en ejecutarse, ralentizando el flujo de desarrollo. Además, requieren de una gran coordinación por parte de los desarrolladores para evitar conflictos en los cambios y en el orden en que se realizan las integraciones. Por estos motivos, desde la metodología XP se propone el uso de un servidor dedicado a realizar las integraciones. Este servidor, denominado servidor de integración continua, asegura la realización de los procedimientos necesarios para integrar los nuevos cambios de manera ordenada y evitando

conflictos [20] y su elección es clave para llevar a cabo con éxito estos procesos.

Existen varias alternativas en el mercado, siendo GitLab CI, Jenkins y GitHub Actions los servidores de integración continua más destacables. Tanto GitLab CI como GitHub Actions forman parte del ecosistema de los gestores de repositorios GitLab y GitHub respectivamente, por lo que son los que se tendrán en cuenta en este proyecto, evitando así la dependencia de una herramienta adicional para este propósito.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo general es analizar las herramientas open-source disponibles para realizar las pruebas estáticas de seguridad de las aplicaciones (SAST).

Los objetivos específicos son:

- Comparar las herramientas SAST open-source para determinar sus fortalezas y debilidades.
- Desarrollar casos de ejemplo para analizar qué vulnerabilidad encuentra cada herramienta SAST y establecer sus limitaciones. Esta tarea se realiza para cada lenguaje de programación.
- Analizar las vulnerabilidades para determinar si hay patrones detectables en el código o no.
- Clasificar las vulnerabilidades para asistir a los expertos de seguridad sobre cuándo usar herramientas SAST y cuándo no.
- Desarrollar una regla en una herramienta SAST para un lenguaje específico para detectar una vulnerabilidad determinada.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de este proyecto está formado por dos ingenieras y un licenciado en informática, un especialista en seguridad teleinformática y un alumno avanzado de la carrera. Este trabajo se desarrolló en el marco del proyecto de investigación: “Análisis de las Herramientas SAST”.

Dada la complejidad del desarrollo del proyecto de investigación, fue necesaria la colaboración de

varios expertos con amplia experiencia en la industria y la investigación académica.

5. BIBLIOGRAFIA

- [1] De Franca, B. B. N., Jeronimo, H., & Travassos, G. H. (2016). Characterizing DevOps by Hearing Multiple Voices. Proceedings of the 30th Brazilian Symposium on Software Engineering - SBES '16. Published. <https://doi.org/10.1145/2973839.2973845>
- [2] Jabbari, R., bin Ali, N., Petersen, K., & Tanveer, B. (2016). What is DevOps? Proceedings of the Scienti_c Workshop Proceedings of XP2016. Published. <https://doi.org/10.1145/2962695.2962707>
- [3] Virani, M. (2015). Understanding DevOps & bridging the gap from continuous integration to continuous delivery. Fifth International Conference on the Innovative Computing Technology (INTECH 2015). Published. <https://doi.org/10.1109/intech.2015.7173368>
- [4] Watts, S. (2019, 29 March). A Brief History of DevOps. BMC Blogs. <https://www.bmc.com/blogs/devops-history/>
- [5] Allspaw, J., & Hammond, P. (2009, Jun 22). 10+ Deployes per Day: Dev and Ops Cooperation at Flickr [Talk]. O'Reilly Velocity Conference, San Jose, California. <https://www.youtube.com/watch?v=LdOe18KhtT4>
- [6] Debois, P. (s. f.). About devopsdays. DevOpsDays. Recuperado 18 de enero de 2022, de <https://devopsdays.org/about/>
- [7] Akshaya, H. L., Nisarga Jagadish, S., Bidya, J., & Veena, K. (2015). A Basic Introduction to DevOps Tools. International Journal of Computer Science and Information Technologies, 6(3). <http://ijcsit.com/docs/Volume%206/vol6issue03/ijcsit2015060382.pdf>
- [8] Beck, K., Fowler, M., Martin, R. C., Beedle, M., Cockburn, A., Cunningham, W., Thomas, D., Mellor, S., Schwaber, K., Sutherland, J., Bennekum, A. V., Grenning, J., Highsmith, J., Hunt, A., Je_ries, R., Kern, J., & Marick, B. (2001, 13 February). Principios del Manifiesto Ágil. Agile Manifesto. <http://agilemanifesto.org/iso/es/principles.html>
- [9] Shackleford, D. (2016, 8 March). A DevSecOps Playbook. SANS. <https://www.sans.org/webcasts/devsecops-playbook-101472>
- [10] Amazon Web Services. (2016, 9 November). Introduction to DevSecOps on AWS. <https://www.slideshare.net/AmazonWebServices/introduction-todevsecops-on-aws-68522874>
- [11] The OWASP Foundation. (s. f.). OWASP DevSecOps Guideline. OWASP. Recuperado 24 de enero de 2022, de <https://owasp.org/www-projectdevsecops-guideline/>
- [12] Adkins, H., Beyer, B., Blankinship, P., Lewandowski, P., Oprea, A., & Stubble_eld, A. (2020). Building Secure and Reliable Systems: Best Practices for Designing, Implementing, and Maintaining Systems (Illustrated ed.). O'Reilly Media. <https://sre.google/books/building-secure-reliable-systems/>
- [13] Peterson, J. (2020, 19 November). Software Composition Analysis Explained. WhiteSource. <https://www.whitesourcesoftware.com/resources/blog/softwarecomposition-analysis/>
- [14] Weerasinghe, M. (2019, 24 December). NodeJS Security Tools – Manjula Weerasinghe. Medium. <https://medium.com/@manjula.aw/nodejs-securitytools-de0d0c937ec0>
- [15] Wells, D. (1999). Continuous Integration. Extreme Programming. <http://www.extremeprogramming.org/rules/integrateoften.html>
- [16] Fowler, M. (2000, 10 September). Continuous Integration (original version). martinowler.com. <https://www.martinfowler.com/articles/originalContinuousIntegration.html>
- [17] Rahman, A. A. U., Helms, E., Williams, L., & Parnin, C. (2015). Synthesizing Continuous Deployment Practices Used in Software Development. 2015 Agile Conference. Published. <https://doi.org/10.1109/agile.2015.12>
- [18] Humble, J., Read, C., & North, D. (2006). The Deployment Production Line. AGILE 2006 (AGILE'06). Published. <https://doi.org/10.1109/agile.2006.53>

[19] Fowler, M. (2013, 30 May). Continuous Delivery. [martinfowler.com.
https://martinfowler.com/bliki/
ContinuousDelivery.html](https://martinfowler.com/bliki/ContinuousDelivery.html)

[20] Wells, D. (1999). Dedicated Release Computer. Extreme Programming. <http://www.extremeprogramming.org/rules/dedicated.html>

Tecnología Blockchain Aplicada a la Ciberdefensa

Cipriano, Marcelo^{1,2}; García, Edith¹, Maiorano, Ariel¹
Malvacio, Eduardo¹

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Facultad de Ingeniería del Ejército (FIE), Universidad de la Defensa Nacional - UNDEF

²Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes UNQ.

{marcelocipriano; egarcia; maiorano; emalvacio}@fie.undef.edu.ar

RESUMEN

El objetivo del proyecto es investigar en el contexto de la criptología y dentro del ámbito de la ciberdefensa, las ventajas, aplicaciones y desafíos de la Tecnología Blockchain. Se trata del registro de información distribuida, inmutable, basado en métodos criptográficos. Almacena datos en bloques, que a su vez son enlazados conjuntamente formando una cadena. Si bien inicialmente se consideró para desarrollar criptomonedas y, se hace referencia más comúnmente a ello, hay un número cada vez mayor de aplicaciones fuera de las finanzas y los negocios.

Las características y propiedades de este tipo de cadena de bloques tales como, visibilidad, verificabilidad, confianza, auditabilidad y trazabilidad permiten ser aplicadas en diferentes espacios, donde funciona como una plataforma *ideal* para aplicaciones seguras descentralizadas. En este proyecto, nos enfocaremos en estudiar específicamente los fundamentos criptográficos del paradigma blockchain, incluyendo sus características, propiedades, beneficios, limitaciones y, sus potenciales aplicaciones en la ciberdefensa. Haremos una revisión de proyectos en desarrollo, interconexiones e implementaciones de esta tecnología, con vistas a desarrollar plataformas propias, orientados particularmente a las necesidades inherentes a la defensa nacional.

Con la innovación digital en infraestructuras críticas, tanto civiles como militares, las ciberamenazas y las actividades maliciosas, resultan inevitables.

La tecnología blockchain es una técnica disruptiva emergente, que permite garantizar ciberseguridad, proporcionando una amplia variedad de aplicaciones, criptográficamente robustas, donde no se requiere de un control centralizado, atributos sumamente significativos en el campo de la ciberdefensa.

Palabras Clave

Criptografía, Tecnología Blockchain, Ciberseguridad, Ciberdefensa

CONTEXTO

“CRYPTO BC: *Tecnología Blockchain aplicada a la Ciberdefensa*” es un proyecto perteneciente a la *Facultad de Ingeniería del Ejército (FIE)* “Gral. Div. Manuel N. Savio”, perteneciente a la *Universidad de la Defensa Nacional (UNDEF)*.

Se encuentra enmarcado en el contexto de la carrera de grado de Ingeniería en Informática, la Especialización en Criptografía y Seguridad Teleinformática y la Maestría en Ciberdefensa, que se dictan en la citada unidad académica.

Allí los investigadores conforman el Grupo de Investigación en Criptología y Seguridad Informática (GICSI) que depende del Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática (CriptoLab) lleva adelante tareas de I+D+i. El equipo está conformado por docentes investigadores categorizados en distintos regímenes científicos, profesionales técnicos, becarios y alumnos de la carrera de grado de Ingeniería en Informática, de la Especialización en Criptografía y Seguridad Teleinformática y de la Maestría en Ciberdefensa.

1. INTRODUCCIÓN

Lee-Kim en [1] definen a la Ciberdefensa como *“la práctica de proteger activos nacionales de las amenazas internas o externas propias de las tecnologías de la información y las comunicaciones (TICs)”*, o sea es la actividad que permite garantizar la supervivencia de cualquier Estado o Nación a las ciberamenazas.

En las últimas décadas hemos presenciado una rápida transformación digital. El surgimiento de ciberamenazas se presenta como el resultado, que nos ha cobrado la celeridad de la transición digital. Numerosos cbersistemas presentes en los mundos IT (*Tecnología de la Información*) y OT (*Tecnología de Operación*) son diseñados y usados sin suficiente (o ninguna) consideración de seguridad. Día a día, nuevos tipos de vulnerabilidades son encontrados y deben ser tratados cuidadosamente, razón por la cual el objetivo de identificar, proteger y responder a las ciberamenazas, se ha convertido en una práctica primordial de un Estado y no, en un simple servicio o empresa. La Tecnología Blockchain o Tecnología *Contable* Distribuida (DLT Distributed Ledger Technology en inglés) fue inicialmente considerada para implementar e-cash,

criptomonedas [2], sin embargo, las características y propiedades de este modelo de cadena de bloques, tales como su resiliencia, seguridad criptográfica e inmutabilidad, mayormente asociadas con su naturaleza descentralizada, la hacen sumamente atractiva y, particularmente útil, en el ámbito de la Ciberdefensa. Efectivamente, esta tecnología podría ser utilizada en diferentes escenarios, tanto en roles operacionales como de soporte. Al mismo tiempo, la criptografía, considerada como la ciencia para proteger la información, aporta los mecanismos fundamentales para garantizar la solidez de una Blockchain. Las propiedades fundamentales que deben satisfacer estas cadenas de bloques son alcanzadas a través de diferentes técnicas criptográficas: asegurar un ecosistema Blockchain requiere el uso de distintas primitivas criptográficas, tales como, funciones de hash, criptografía de clave simétrica y criptografía de clave pública o asimétrica [3]. Además de confidencialidad, también se necesita proporcionar otros servicios de seguridad como integridad, autenticación, no-repudio y confiabilidad, requerimientos precisos exigidos en los sistemas de almacenamiento y procesamiento de información.

Numerosos proyectos de Blockchain incluyendo los que se orientan a criptomonedas y contratos inteligentes, han mostrado tener características robustas de seguridad criptográfica. Por esta razón, varios trabajos que involucran dicha tecnología están siendo considerados para ser aplicados en el ámbito de la ciberdefensa. Las aplicaciones pueden ser categorizadas en distintos dominios de acuerdo con su uso, desde las que requieren Integridad de datos hasta aquellas que refieren a Redes de Datos Descentralizadas (civiles o militares), Gerenciamiento de Cadena de Suministros, Internet de las Cosas (Internet-of-Things IoT), Internet de las Cosas Industrial (Industrial-Internet-of-Things

IloT), Comunicaciones, Identificación y Autenticación para escenarios dentro de la ciberdefensa.

Varios trabajos han investigado el rol de blockchain relacionado a ciberdefensa. Taylor et al. [4] hicieron una revisión sistemática de las aplicaciones de blockchain directamente para Ciberseguridad. Lilly and Lilly [5] se focalizaron sobre los proyectos militares en el mundo real de tres países: EE.UU., China y Rusia. Bansal et al. [6] estudiaron aplicaciones específicas de ciberseguridad usando blockchain. Ahmad et al. [7] investigaron blockchain para defensa y espacio aéreo, los cuales son dominios estrechamente relacionados. Su estudio proporciona una arquitectura detallada de aplicaciones blockchain.

Sin embargo, nuestra investigación apunta a indagar acerca de aplicaciones blockchain que refieren a la definición de Ciberdefensa dada en [1], donde esta práctica no está limitada al uso militar, sino que contempla la seguridad nacional.

Cuando consideramos las ciberamenazas actuales y el significado de seguridad nacional [1]: *“es la práctica de proteger una nación o estado y sus infraestructuras críticas, de amenazas internas y externas”*, entendemos que la tecnología blockchain debe ser mejor aprendida y aplicada. Por lo tanto, en este proyecto, nuestro estudio está orientado a indagar acerca del rol de la Tecnología Blockchain en el campo de la Ciberdefensa. En síntesis, surgen las siguientes cuestiones a considerar en el proyecto:

- 1) ¿Cuáles son los principales beneficios de utilizar la Tecnología Blockchain en ciberdefensa?
- 2) ¿Qué tipos de aplicaciones Blockchain son usadas en el ámbito de la ciberdefensa y qué aplicaciones tendrían que ser específicamente impulsadas?

- 3) Comparado con otros dominios de aplicación, ¿cuáles son las limitaciones o desafíos de la Tecnología Blockchain en el campo de la ciberdefensa?

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto está conformado por distintas líneas de acción:

- Seguimiento de bibliografía y publicaciones científicas acerca de las aplicaciones de la tecnología blockchain, proyectos blockchain aplicados a la ciberdefensa.
- Asistencia y/o seguimiento de cursos, jornadas, congresos y workshops de relevancia, tanto nacionales como internacionales acerca de criptografía, tecnología blockchain y sus aplicaciones.
- Elaboración de módulos para pruebas de aplicación de la tecnología blockchain en ciberdefensa.

3. RESULTADOS OBTENIDOS / ESPERADOS

De acuerdo con los tres puntos planteadas en la introducción, nuestras contribuciones serían las siguientes:

Establecer qué roles fundamentales cumple la tecnología blockchain para ser aplicada a la ciberdefensa considerando los aspectos de visibilidad, verificabilidad, resiliencia y auditabilidad.

Estudiar proyectos de desarrollo de sistemas blockchain para ciberdefensa, incluyendo aquellos que respondan a investigación y desarrollo de gobiernos.

Establecer y enumerar los límites y desafíos de la tecnología blockchain aplicada a la ciberdefensa en dominios específicos.

4. FORMACIÓN DE RECURSOS HUMANOS

Los investigadores que llevan adelante el proyecto dictan las asignaturas Crip-

tografía y Seguridad Teleinformática, Matemática Discreta y Paradigmas de Programación. Desde esas cátedras se invita de forma permanente a los alumnos para participar como colaboradores. Asimismo, y por primera vez, varios alumnos de la especialización en Criptografía y Seguridad Teleinformática que están llevando adelante su Trabajo Final Integrador, como así también maestrandos pertenecientes a la Maestría en Ciberdefensa, que se encuentran trabajando en el desarrollo de sus respectivas tesis; y que además abordan temáticas relacionadas a la de este proyecto, han sido invitados a participar del mismo. Se espera que la contribución mutua entre el equipo de investigadores, futuros especialistas y maestrandos permita alcanzar niveles sinérgicos de avance en la investigación, la formación de recursos humanos.

La Formación de Recursos Humanos permite incrementar el Know-How que tendrá el grupo de investigadores a lo largo de la vida del proyecto. Será un importante beneficio de sus integrantes y de la institución en la cual desarrollan sus actividades científico-docentes.

Por último y atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

[1] Lee, Suhyeon & Kim, Seungjoo. (2021). Blockchain as a Cyber Defense: Opportunities, Applications, and Chal-

lenges. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3136328.

[2] Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," [Online]. available: [Online]. available: <https://bitcoin.org/bitcoin.pdf>, 2008, (accessed Oct. 19, 2021).

[3] Bashir, Imran. (2017). Mastering Blockchain.

[4] P. J. Taylor, T. Dargahi, A. Dehghantanha, R. M. Parizi, and K.-K. R. Choo, "A systematic literature review of blockchain cyber security," Digital Communications and Networks, 2019.

[5] B. Lilly and S. Lilly, "Weaponising blockchain: Military applications of blockchain technology in the US, China and Russia," The RUSI Journal, pp. 1–11, 2021

[6] P. Bansal, R. Panchal, S. Bassi, and A. Kumar, "Blockchain for cybersecurity: A comprehensive survey," in Proc. of 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT). IEEE, 2020, pp. 260–265.

[7] R. Wasim Ahmad, H. Hasan, I. Yaqoob, K. Salah, R. Jayaraman, and M. Omar, "Blockchain for aerospace and defense: Opportunities and open research challenges," Computers & Industrial Engineering, vol. 151, p. 106982, 2021.

Aprendizaje por Refuerzo Aplicado al Etiquetado de Tráfico de Red.

Carlos A. Catania, Elina Pacini, Rodrigo Gonzalez, Jorge Guerra, Tatiana Parlanti,
Luciano Robino y Eduardo Pavez

Universidad Nacional de Cuyo, Facultad de Ingeniería, LABSIN
Campus Universitario, Mendoza, Argentina.

{harpo,elina.pacini,rodrigo.gonzalez,jorge.guerra,tatiana.parlanti,luciano.robino,eduardo.pavez}@ingenieria.uncuyo.edu.ar

RESUMEN

Durante los últimos años, se han aplicado varios enfoques basados en el aprendizaje automático (AA) para aligerar el análisis y el reconocimiento de comportamientos maliciosos en las redes informáticas. Estos enfoques se han centrado en facilitar la tarea del personal de seguridad de la red, en mejorar su capacidad de detección y en intentar aumentar el nivel de automatización en el reconocimiento y el análisis de los comportamientos maliciosos. Sin embargo, todos estos enfoques requieren un conjunto de datos anotados con etiquetas. Al utilizar conjuntos de datos etiquetados, estos sistemas pueden generalizar comportamientos no observados previamente. Por desgracia, los conjuntos de datos etiquetados no suelen tener la representatividad y la precisión necesarias. Esta última carencia está asociada, entre otros factores, a la falta de recursos humanos formados y a la dificultad en el proceso de creación de estos conjuntos de datos etiquetados.

En consecuencia, el objetivo general de este proyecto es desarrollar estrategias para reducir el esfuerzo humano durante el proceso de etiquetado y análisis de

conjuntos de datos con tráfico de red. Estas estrategias se centran en la aplicación de diferentes técnicas de Aprendizaje por Refuerzo (RL) para establecer políticas adecuadas que minimicen el grado de interacción del usuario durante el proceso de etiquetado. Al utilizar RL es posible aprender nuevas políticas que tengan en consideración otras recompensas como ser la experiencia o capacidad del usuario junto a aspectos distintivos del proceso de etiquetado como ser la calidad, oportunidad y relevancia, entre otros. De esta manera se evita depender de una única política a la hora de tomar la decisión de consultar al usuario. Esto último resulta fundamental para el desarrollo de sistemas de detección basados en técnicas de AA.

La principal contribución que se espera obtener de la investigación propuesta en el corto plazo es disponer de un conjunto de técnicas que faciliten el proceso de etiquetado en trazas de tráfico de red para su utilización en sistemas de detección de comportamiento malicioso basados en algoritmos de AA, mejorando de manera continua su capacidad de detección.

CONTEXTO

El presente proyecto se desarrolla en el marco de Facultad de Ingeniería dentro Laboratorio de sistemas inteligentes (LABSIN). Este trabajo es parte del proyecto de investigación que dio inicio en septiembre de 2022 en el marco de los proyectos bienales de secretaria de Investigación, Internacionales y Posgrados (SIIP) de la Universidad Nacional de Cuyo.

1 INTRODUCCIÓN

El proceso para la construcción de un modelo basado en aprendizaje automático (AA) asociado a un malware consiste de tres etapas [1,2]. En una primera etapa, se genera un modelo de comportamiento considerando diferentes características de la red (por ej. el tamaño, periodicidad y duración, de las trazas de tráfico). Este modelo permite visualizar de manera sencilla un gran número de conexiones de red a lo largo del tiempo y constituye además una herramienta de análisis. Luego, en una segunda etapa de reconocimiento, se realiza un proceso de asociación entre el modelo de comportamiento y la aplicación maliciosa, tarea normalmente denominada **etiquetado** y que requiere del apoyo de un especialista humano. Finalmente, en una tercera etapa, mediante el uso de algoritmos con base estadística o de aprendizaje automático se generaliza el modelo de comportamiento a fin de mejorar la capacidad del modelo para reconocer de forma automática otras aplicaciones maliciosas que presenten un comportamiento similar.

En los últimos años, se han desarrollado varios **métodos asistidos** para abordar el problema del etiquetado aplicado a las trazas de tráfico de red. En estos métodos, la inclusión de un usuario con cierto nivel de experiencia en el proceso de etiquetado mejora la calidad de los conjuntos de datos resultantes. Dado que los expertos son un recurso escaso, el tiempo de etiquetado debe

utilizarse de forma eficiente. Es por esto que los métodos de etiquetado asistidos se apoyan en diferentes técnicas como el Active Learning (AL), las técnicas de visualización, o una combinación de ambas, para tratar con un tráfico de red más realista y acelerar el proceso de etiquetado [3, 4, 5, 6, 7].

Frente a los enormes desafíos que presenta el tener que lidiar con los volúmenes de tráfico de las redes actuales como ser almacenamiento, capacidad de cómputo y el gran esfuerzo humano requerido para el análisis, el proceso de etiquetado podría verse como un problema que presenta pocas diferencias respecto del proceso de detección de comportamiento malicioso. Sin embargo, en un trabajo reciente presentado por el director del presente proyecto [8], se observó que existen diferencias significativas en el proceso de etiquetado de tráfico frente a un problema de detección de tráfico convencional, como el que lleva a cabo un SDI. Estas diferencias pueden enmarcarse bajo los siguientes aspectos:

Oportunidad: en el proceso de etiquetado no existe la necesidad de realizar la detección en un margen de tiempo relativamente corto. Durante el proceso de etiquetado es posible contar con el tiempo necesario para confirmar la potencial presencia de malware.

Relevancia: un error de etiquetado no presenta la misma importancia para un sistema de detección en tiempo real que para un sistema de creación de conjuntos de datos etiquetados. Como ejemplo, puede mencionarse el hecho de que en el contexto de un SDI, los falsos positivos pueden resultar una molestia para el usuario. Sin embargo, durante un proceso de etiquetado simplemente representa parte del ruido que pudiera ocurrir en el conjunto de datos resultante.

Calidad: durante el proceso de etiquetado se busca fundamentalmente obtener etiquetas de calidad que representen las características de la red. El objetivo es contar con datos de calidad, a fin de alimentar los distintos algoritmos de AA. Como consecuencia, no resulta necesario etiquetar

todas las trazas de tráfico que pertenezcan a un mismo comportamiento malicioso, sino que bastará con etiquetar adecuadamente un subconjunto lo suficientemente representativo.

Economía: el proceso de etiquetado no presenta consecuencias económicas en lo inmediato para el sistema vulnerado. Es decir, frente a un comportamiento malicioso no detectado, en principios no existe una consecuencia más allá de datos espurios para la construcción de un modelo estadístico de predicción. Mientras que, en el caso de un sistema de detección, el no reconocimiento de un comportamiento malicioso puede ocasionar pérdidas importantes a la organización donde se encuentre la red.

Curiosamente, estas diferencias no son tenidas en cuenta en la mayoría de los enfoques orientados al etiquetado de trazas de tráfico, y generalmente se aplica un enfoque que no se diferencia mucho del problema de reconocimiento de comportamiento malicioso en el tráfico de red. La realidad es que hoy en día el reconocimiento de comportamiento malicioso en el tráfico red es una tarea que demanda una gran cantidad de recursos tanto de índole computacionales como humanos. Es por ello que en la actualidad estos sistemas basados en AA solo han sido desplegados en ambientes controlados donde el volumen de tráfico evaluado dista mucho de los volúmenes actuales reales, lo que resta peso práctico a los enfoques. De esta manera, muchas de las tareas de análisis y reconocimiento se realizan fuera de línea y el proceso de etiquetado se realiza de manera manual o asistida por herramientas de visualización.

Se trata de un inconveniente de gran impacto si se tiene en cuenta el hecho de que en muchos casos resulta necesario realizar un reentrenamiento periódico de los modelos de AA a causa de cambios observados en la distribución del tráfico de la red, es decir, que el etiquetado de un conjunto de datos estático en el tiempo resulta inadecuado para hacer frente a las características del tráfico de red actual. Además, es importante destacar que

estos inconvenientes no sólo se aplican al problema de etiquetado en el contexto de detección de intrusiones, sino que también se observan en muchos otros problemas donde se requiere la aplicación de algoritmos de AA. Como consecuencia, la comunidad científica en general y de seguridad en particular, ha comenzado a trabajar en el desarrollo de enfoques orientados a datos (data-centric) en los cuales se pone el foco en el desarrollo de herramientas para la construcción, mantenimiento y evaluación de conjuntos de datos de manera eficiente y fácilmente repetible. Más aún, se ha observado que la comunidad se ha tenido que enfrentar al hecho de que si bien el desarrollo y despliegue de sistemas basados en AA es relativamente simple y rápido, su mantenimiento a lo largo del tiempo puede volverse difícil y costoso. Recientemente, para hacer frente a estos inconvenientes, han surgido distintos roles y procedimientos entre los que se destacan MLE (Machine Learning Engineer) y MLOps (Machine Learning Operations). El rol de MLE combina aspectos de la ingeniería de software tradicional aplicado a los procesos de generación y mantenimiento de aplicaciones basadas en AA. De manera similar, MLOps se focaliza en el desarrollo de procedimientos para simplificar y automatizar el proceso de despliegue y mantenimiento a gran escala de los modelos de AA en ambientes de producción. Entre los aspectos más importantes que involucran el concepto de MLOps se destaca el etiquetado de calidad de manera periódica requeridos para el ajuste y mantenimiento de los modelos de AA. La realidad es que muchos de estos roles todavía no han sido considerados en las propuestas para el desarrollo de Sistemas de detección de intrusos basados en AA.

El presente proyecto se enfoca en la aplicación de técnicas de aprendizaje por refuerzo (RL) dentro de un ciclo de AL de tal manera que permita desarrollar herramientas para facilitar el proceso de etiquetado de tráfico de red. El aprendizaje por refuerzo (RL) es un área del aprendizaje automático

que se focaliza en el desarrollo de agentes inteligentes capaces de realizar acciones en un entorno a fin de maximizar una recompensa acumulada.

Normalmente, en un flujo de trabajo de AL, se aplica una política fija para decidir cuándo preguntar al usuario por el valor para la etiqueta de una instancia en particular. La política más común está basada en el concepto de incertidumbre que consiste en seleccionar aquellas trazas de tráfico que se encuentran cerca de la frontera de decisión del modelo de AA, es decir, para las cuales la política no puede decidir de manera autónoma cómo etiquetar. Sin embargo, esta política puede no resultar adecuada cuando se trabaja con grandes volúmenes de datos. Las técnicas de RL son un enfoque de reciente aplicación en el aprendizaje de una política dinámica de AL a partir de los datos. Al utilizar RL es posible aprender nuevas políticas que tengan en consideración otras recompensas como ser la experiencia o capacidad del usuario junto a aspectos distintivos del proceso de etiquetado como ser la calidad, oportunidad y relevancia, entre otros. De esta manera se evita depender de una única heurística a la hora de tomar la decisión de consultar al usuario. Esto último resulta fundamental para el desarrollo de sistemas de detección basados en técnicas de AA.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO.

El proyecto se enmarca en dos de las áreas de investigación del instituto de investigaciones del LABSIN, en particular la captura y procesamiento de datos a gran escala y el aprendizaje estadístico.

Para el desarrollo del presente proyecto pueden diferenciarse 4 etapas principales:

A) Análisis preliminar del problema. En esta etapa se incluye la recopilación de información bibliográfica sobre el tema poniendo especial énfasis en la aplicación de

técnicas de aprendizaje por refuerzo a problemas relacionados con la temática etiquetado de conjuntos de datos. La intención es evaluar no solo las estrategias de etiquetado de tráfico sino otras que podrían ser fácilmente aplicables al problema formulado en el presente proyecto.

B) Generación de conjuntos de datos.

Esta etapa se focaliza en la construcción o adaptación de uno (o varios) conjuntos de datos adecuados para algún problema de seguridad informática. En principio se considerará la utilización de los conjuntos previamente producidos en el marco del proyecto SIIP B06/036 bajo la dirección del Dr. Carlos Catania.

C) Algoritmos para aprendizaje de políticas. Se analizarán e implementarán algoritmos de etiquetado basados en técnicas clásicas de RL, como políticas del gradiente (Policy Gradient) o Q-learning adaptados para el problema de etiquetado de tráfico de red. Durante esta primera actividad se tendrá en cuenta una única acción posible que consiste en preguntar (o no) al usuario por el valor de una etiqueta. Para las diferentes implementaciones de algoritmos de RL, se considerará en primer lugar un mecanismo de recompensa enfocado principalmente al aspecto de calidad necesario para el etiquetado de trazas de tráfico. Esta actividad tiene como objeto determinar cómo interactúan los elementos básicos de un algoritmo de RL, y sus ventajas y limitaciones para su aplicación dentro de un flujo de trabajo de AL para el etiquetado de trazas de tráfico.

D) Experimentación. Diseño de los experimentos computacionales destinados a medir el rendimiento de los algoritmos desarrollados durante la actividad (C). Para este punto resulta necesario considerar métricas que vayan más allá de la precisión de las etiquetas (i.e. Exactitud, Sensibilidad, Sensitividad), sino otros aspectos necesarios para medir el rendimiento de los enfoques de

AL como por ejemplo la tasa de aprendizaje. Esta última métrica apunta a medir el número mínimo de trazas etiquetadas necesario para que el algoritmo de AA comience a funcionar con niveles aceptables de exactitud. Otra métrica de gran importancia que se considerará durante la experimentación es la tolerancia al ruido, la cual permite conocer el grado de conexiones incorrectamente etiquetadas por el usuario que permite el algoritmo de AA subyacente sin perder exactitud.

3 RESULTADOS ESPERADOS

Al término de los dos años de duración del plan de trabajo se pretende que se haya logrado:

1. Fortalecer la línea de investigación en la aplicación de modelos de aprendizaje automático relacionados con el tráfico de red.
2. Contar con el prototipo de una herramienta computacional que permita facilitar el proceso de etiquetado de tráfico de red dentro de un ciclo de AL. Dicha herramienta deberá ser capaz lidiar no solo con la precisión de las etiquetas sino también con su representatividad.
3. Se espera además contar con un marco de trabajo que facilite la evaluación de variantes aprendizaje por refuerzo a otros de problemas de clasificación más allá de la seguridad informática

4 FORMACIÓN DE RECURSOS HUMANOS

Se espera capacitar en el ámbito de la investigación a profesores y alumnos interesados en participar en un entorno académico y tecnológico innovador y a todos aquellos actores interesados en los resultados del proyecto.

Sobre la temática de este proyecto se está desarrollando el doctorado en ciencias de la computación de la Lic Tatiana Parlanti, en la Universidad Nacional del centro de la provincia de Buenos Aires.

5 BIBLIOGRAFÍA

- [1] J. L. Guerra, C. A. Catania, and E. Veas. Active learning approach to label network traffic datasets. *Journal of Information Security and Applications*, 49:102388, 2019. [Indexada SCI/SCI-E].
- [2] P. Torres, C. Catania, S. Garcia, and C. G. Garino. An analysis of recurrent neural networks for botnet detection behavior. In *2016 IEEE Biennial Congress of Argentina (ARGENCON)*, pages 1–6, 2016.
- [3] N. Görnitz, M. Kloft, K. Rieck, and U. Brefeld. Toward Supervised Anomaly Detection.
- [4] A. Beaunon, P. Chifflier, and F. Bach. Ilab: An interactive labelling strategy for intrusion detection. In M. Dacier, M. Bailey, M. Polychronakis, and M. Antonakakis, editors, *Research in Attacks, Intrusions, and Defenses*, pages 120–140, Cham, 2017. Springer International Publishing
- [5] X. Fan, C. Li, X. Yuan, X. Dong, and J. Liang. An interactive visual analytics approach for network anomaly detection through smart labeling. *Journal of Visualization*, 22(5):955–971, 2019.
- [6] Y. Yang, Z. Ma, F. Nie, X. Chang, and A. G. Hauptmann. Multi-Class Active Learning by Uncertainty Sampling with Diversity Maximization. *International Journal of Computer Vision*, 113(2):113–127, 2015.
- [7] S. McElwee. Active learning intrusion detection using k-means clustering selection. In *SoutheastCon 2017*, pages 1–7, 2017.
- [8] J. Guerra, C. Catania, and E. Veas. Datasets are not enough: Challenges in labeling network traffic, 2021. arxiv.org/abs/2110.05977

Criptografía Liviana y Ciberseguridad aplicados a Sistemas Ciberfísicos.

Cipriano, Marcelo; Eterovic, Jorge; García, Edith; Torres, Luis.

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{marcelo.cipriano; jorge.eterovic; edith.garcia}@usal.edu.ar
Luis.Antonio.Torres@kyndryl.com

RESUMEN

La fusión de las Tecnologías de la Información y Comunicación con la Automatización, Inteligencia Artificial, Robótica, Nanotecnología, Impresión 3D, Computación Cuántica, Biotecnología, entre otras disciplinas, está borrando notablemente las divisiones existentes entre lo digital, físico y biológico. En particular las esferas bien diferenciadas de la *Tecnología de la Información* y la *Tecnología de la Operación* (respectivamente *IT* y *OT* por sus siglas en inglés) ya se encuentran unificadas bajo una misma esfera.

Los llamados *Sistemas Ciberfísicos* (*CPS: cyber-physical system*)[1] son *mecanismos físicos controlados o monitorizado por equipos informáticos, integrados mediante una red de datos y de comunicaciones, siendo Internet la red más ampliamente empleada.*

Este nuevo conjunto de sistemas, agrupa prácticamente la totalidad de los equipos industriales modernos (*Industria 4.0*), como así también los recursos y sistemas de las *Ciudades Inteligentes* (*Smart-Cities*), *Cibermedicina* (*e-Health*), *Internet de las Cosas* (*Internet of Things*)[2–4], solamente por mencionar algunos de los más conocidos.

Esta revolución tecnológica no ha sido acompañada por una revolución equivalente, en la seguridad. Esta aún se encuentra algunos pasos por detrás. Y muy probablemente siempre sea así, tal como la seguridad de los sistemas informáticos ha demostrado en todos estos años. Pues aun habiendo transcurrido varias décadas de la llamada *Revolución Informática* y la *Era de Internet* no se ha logrado vencer la batalla de contra los ciberataques, la existencia de vulnerabilidades, el malware, las intrusiones de

sistemas, robos de datos, ingeniería social y engaños, en su casi innumerable cantidad de situaciones. Factores que atentan contra la seguridad de la información, redes y demás. Sobre todo en una tendencia que se ha visto aumentar en los últimos años. El proyecto persigue el estudio y análisis de las fortalezas y debilidades de los protocolos de comunicaciones de los Sistemas Ciberfísicos en general, como así también los mecanismos y algoritmos de seguridad criptográficos que en ellos se pueden implementar. En particular aquellos mecanismos de autenticación, confidencialidad e integridad de la información[5-7], ofrecidos por la llamada *Criptografía Liviana o Ligera* (*LiCrypt: Lightweight Cryptography*), en la forma de algoritmos de cifrado, intercambio de claves, hash y firma digital.

Asimismo el proyecto también propone la difusión y promoción de estas temáticas en la comunidad científica y tecnológica local, en la que se aprecia un notable desconocimiento. Seguramente provocado por diferentes causas cuyo análisis y comprensión exceden los alcances del proyecto.

La persistente labor investigativa y de difusión que la Universidad del Salvador mantiene acerca de estas temáticas, contribuirá a sensibilizar sobre esta problemática, como así también acercar posibles soluciones. Tanto sea al interior del ámbito académico nacional, como del sector productivo e industrial del país.

Además permitirá nutrir a la futura Diplomatura de Ciberseguridad en Entornos IT/OT (que se encuentra actualmente en evaluación) con conocimientos actualizados y recursos humanos.

Palabras Clave:

Sistemas Ciberfísicos, Criptografía Liviana, Internet de las Cosas, CPS, LiCrypt, IoT.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad Nacional del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndolas como un servicio a la comunidad, entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias Unidades Académicas se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándose a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e internacionales, como así también, apoyo y orientación de recursos para la investigación.

A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto con una duración de 2 años (2023-2025).

El mismo se encuentra aprobado por Disposición Decanal No 58/22 con el Nro. Trámite SIGEVA 80020220200024US.

1. INTRODUCCIÓN

Thomas Kuhn ha expuesto la mecánica de ciertos cambios científicos cuyas características los convierten en verdaderas “revoluciones”. que mueven a la humanidad en su conjunto. Estas pueden afectar a su vez, otros campos del conocimiento, como por ejemplo, la Tecnología. Y es aquí donde se observa un nuevo y vertiginoso cambio de paradigma, llamado “*Revolución Industrial*”.

Jeremy Rifkin declaró ante el Parlamento Europeo en 2006, las características de la llamada *Tercera Revolución Industrial* o *Industria 3.0*[8]. Se distingue por la integración de la Informática y las *Tecnologías de la Información* con la *Automatización de Procesos Industriales*.

En 2016, Klaus Schwab expone las características de la que dio en llamar la *Cuarta Revolución Industrial* o *Industria 4.0*[9], ante el Foro Económico Mundial de ese año. Esta vez la revolución sería mucho más impactante que todas las ante-

rioras. A las Tecnologías de la Información y la *Automatización*, se le suman las diferentes tecnologías: *Comunicación, Inteligencia Artificial, Robótica, Nanotecnología, Impresión 3D, Computación Cuántica y Biotecnología*,[10] entre otras.

Esta fusión comienza a borrar las divisiones existentes entre lo digital, físico y biológico. Se difuminan las fronteras, otrora bien diferenciadas de la *Tecnología de la Información* y la *Tecnología de la Operación* (respectivamente *IT* y *OT* por sus siglas en inglés) las que prácticamente ya se encuentran unificadas.

Sin embargo, la Seguridad de la Información y de las Comunicaciones no han podido acompañar a la par, este desarrollo vertiginoso. Y probablemente jamás lo puedan hacer. Tal vez estén condicionadas a avanzar lo mejor que puedan, pero uno o varios pasos por detrás. Y justamente aquí, en esta persistente brecha, que se vislumbra infranqueable, es donde se fundamenta el problema.

Durante las revoluciones industriales anteriores, la seguridad se acotaba a la dimensión física; la máquina en cuestión, la planta o el producto. Pero la situación con los *Sistemas Ciberfísicos* de la *Industria 4.0*, es diferente. Se pueden mencionar, al menos 4 aspectos relevantes al momento de considerar los problemas de seguridad en estos sistemas[11]:

- a) pueden ser víctima de ataques a distancia, llevados adelante a través de la red de datos y comunicaciones que los conecta.
- b) algunas vulnerabilidades podrían ser explotadas sin que se requiera un conocimiento profundo o avanzado, al alcance de una cantidad mayor de atacantes.
- c) la enorme superficie de ataque disponible, aumentando las posibilidades de éxito de los atacantes.
- d) el alto impacto sufrido por las víctimas de los ataques, sean estos usuarios individuales y/u organismos empresariales, gubernamentales, etc.

En cuanto a este último punto, se pueden destacar, a su vez:

- a) comprometer la información que se procesa y transmite, afectando su privacidad y seguridad[12–13]
- b) secuestrar el/los dispositivos del sistema a cambio de un rescate llamado “*Ransom of Things*” (RoT).
- c) conformar redes zombis que tomen el control parcial o total de los dispositivos y con ellos llevar

adelante ataques masivos a organismos, empresas y gobiernos.

d) afectar el funcionamiento de una o varias *Infraestructuras Criticas* de una nación, como por ejemplo la interrupción del suministro eléctrico, introducir defectos en las plantas potabilizadoras de agua o de tratamientos de desechos, etc.

Este último escenario, no solamente es posible, sino que algunas naciones ya han considerado medidas: a partir del 14 de Junio de 2016 la *Organización del Tratado del Norte (OTAN)* considera a los ciberataques como “*ataques armados convencionales*”. Es decir que si cualquiera de sus miembros fue víctima de un ciberataque adjudicado a alguna potencia extranjera, solicitará ejecutar el *Artículo 5* del tratado. Este artículo indica la *respuesta armada de la parcialidad o totalidad de sus países miembros*[14].

Una de las principales causas de los problemas de seguridad radica en que por su diseño, tamaño y demás características propias, los Sistemas Ciberfísicos (incluidos los dispositivos IoT, entre otros) no suelen disponer de robustos mecanismos de defensa y protección. Las conexiones, datos, información, comandos que procesan y transmiten carecen mayoritariamente de confidencialidad, autenticación, integridad, no repudio y hasta disponibilidad.

En la actualidad existen muchas vulnerabilidades explotadas por Malware y Redes Botnets[15], entre otros.

Es por ello que el gobierno de Estados Unidos vislumbró los peligros de permitir que la Internet de las Cosas quede sin seguridad. pues la mayoría de tales dispositivos no se ve alcanzado por ningún estándar, nacional o internacional y cada fabricante ofrece los servicios a su criterio.

Fue así que en abril de 2017 a través de *National Institute of Standards and Technology (NIST: Instituto Nacional de Estándares y Tecnología en inglés)* Estados Unidos llama a concurso internacional en busca del mejor algoritmo de cifrado-autenticado de criptografía liviana para ambos perfiles: para ser usado indistintamente en hardware y software (profile 1) y sólo en hardware (profile 2). Los ganadores del certamen se convertirán en los primeros estándares para aplicar en *Internet de las Cosas*. Al menos para aquellos productos que sean comercializados en ese país.

Es probable que los estándares que NIST emita, se conviertan también en estándares de facto para el resto del mundo, como ocurrió oportunamente con DES y luego con AES, dos algoritmos de cifrado por bloque.

Es por todo lo expuesto que se manifiesta la relevancia del estudio de la seguridad de los *Sistemas Ciberfísicos* en general como así también los protocolos y algoritmos criptográficos que se adopten para proteger las comunicaciones e información en el contexto de la *Industria 4.0*.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se persigue el estudio y análisis de:

a) algoritmos criptográficos livianos y protocolos de seguridad, observando fortalezas y debilidades matemáticas y criptográficas[16-17].

b) mecanismos de intercambio de claves, autenticación, resumen (hash) y seguridad[18-19].

c) test y pruebas de seguridad para ser aplicados a algoritmos criptográficos.

d) el o los nuevos estándares criptográficos en el campo de la *Criptografía Liviana o Ligera (LiCrypt)*.

e) los ataques criptoanalíticos convencionales/recientes y su incidencia sobre los algoritmos y protocolos[20-22].

Asimismo y por la relevancia, alcance del tema y escaso conocimiento del mismo por gran parte de la comunidad industrial de nuestro país, también se persiguen las siguientes líneas de difusión y concientización:

1) Explicar y difundir la existencia de nuevos algoritmos y estándares criptográficos, como así también sus características de seguridad y su ámbito de aplicación.

2) Transferir a la comunidad académica, científica nacional o internacional, docentes e ingenieros del ámbito IT y OT, la información y resultados obtenidos. En procura de lograr un nexo entre la investigación científico/académica y el mundo de la producción en el marco de Industria 4.0.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Se persigue, tal como se ha indicado, no solamente el alcance de conocimientos teóricos y analíticos en el área de la Criptología. Además se persigue la concientización de la problemática de la ciberseguridad en el entorno industrial, la difusión de estándares de seguridad, protocolos y algoritmos criptográficos.

También este proyecto permitirá nutrir a la *Diplomatura en Ciberseguridad en Entornos IT y OT* que fue diseñada y presentada a las autoridades de la Universidad del Salvador. Actualmente la misma se encuentra en el proceso de evaluación a la espera de su aprobación. Cuyos docentes conforman, en su gran mayoría, el equipo de investigadores de este proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente de *Tecnologías Aplicadas* en la *Facultad de Ingeniería*, el área de la *Seguridad Informática*, de la *Universidad del Salvador*.

Se espera que en el presente año el equipo pueda crecer con la incorporación de más docentes investigadores y alumnos. Ya que redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes participantes, como así también sembrando las bases para la investigación del futuro, a través de la participación de alumnos de la Facultad de Ingeniería.

Se espera, además, que cuando la diplomatura en Ciberseguridad IT/OT sea aprobada y ofrecida a la comunidad, posibilite el ingreso de docentes y alumnos al proyecto.

5. REFERENCIAS

[1] Lee, E. Cyber Physical Systems: Design Challenges. EECS Department University of California, Berkeley Technical Report No. UCB/EECS-2008-8 January 23, 2008. <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-8.html>

[2]Bono, S. Green, M. Stubblefield, A. Juels, A. Rubin, A. Szydlo, M. Security analysis of a cryptographically-enabled RFID device. In Proceedings of the 14th Conference on USENIX

Security Symposium - Volume 14, SSYM'05, pages 1–1, USA, 2005.

[3]Courtois, N. Nohl, K. O'Neil, S. Algebraic attacks on the crypto-1 stream cipher in mifare classic and oyster cards. Cryptology ePrint Archive, Report 2008/166, 2008. <http://eprint.iacr.org/2008/166>.

[4]Dubrova, E. Hell, M. Espresso: A stream cipher for 5g wireless communication systems. Cryptography and Communications, 9(2):273–289, 2017.

[5]Wang, M. Lin, D. Related Key chosen IV Attack on Stream Cipher Espresso Variant. IEEE International Conference on Computational Science and Engineering (CSE) 2017.

[6]Golic, J. Cryptanalytic attacks on MIFARE classic protocol. In Ed Dawson, editor, Topics in Cryptology – CT-RSA 2013, volume 7779 of Lecture Notes in Computer Science, pages 239–258. Springer, Heidelberg, February / March 2013.

[7]Jovanovic, P., Luykx, A., and Mennink, B. (2014). Beyond 2 c/2 security in sponge-based authenticated encryption modes. In Sarkar, P. and Iwata, T., editors, Advances in Cryptology – ASIACRYPT. 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014.

[8]Rifkin, J. The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World. Palgrave Macmillan New York, 2011.

[9]Schwab, K. The Fourth Industrial Revolution. What It Means and How to Respond. Foreign. Affairs. World Economic Forum. 2016.

[10]Hermann, M. Pentek, T. Otto, B. "Design Principles for Industrie 4.0 Scenarios," 49th Hawaii International Conference on System Sciences (HICSS), pp. 3928-3937. Hawaii, 2016.

[11]Garcia, F. van Rossum, P. Verdult, R. Schreur, R. Dismantling SecureMemory, CryptoMemory and CryptoRF. In Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS '10, pages 250–259, New York, NY, USA, 2010. ACM.

[12]Li, R. Li, H. Li, C. Sun, B. A low data complexity attack on the GMR-2 cipher used in the satellite phones. Pages 485–501.

- [13]B. Driessen, R. Hund, C. Willems, C. Paar, and T. Holz. Don't trust satellite phones: A security analysis of two smatphone standards. In 2012 IEEE Symposium on Security and Privacy, pages 128–142, May 2012.
- [14]Organización del Tratado del Atlántico Norte (OTAN). Texto completo del tratado actualizado. https://www.nato.int/cps/en/natohq/official_texts_17120.htm?selectedLocale=es
- [15]Graham, Robert. Mirai and IoT botnet Analysis. RSA Conference 2017. San Francisco. 2017.
- [16]Lu, Y. Vaudenay, S. Faster correlation attack on Bluetooth keystream generator E0. In Matthew Franklin, editor, *Advances in Cryptology – CRYPTO 2004*, volume 3152 of *Lecture Notes in Computer Science*, pages 407–425. Springer, Heidelberg, August 2004.
- [17]Borghoff, J. Knudsen, L. Leander, G. Matusiewicz, K. Cryptanalysis of C2. In Halevi [Hal09], pages 250–266.
- [18]Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: On the security of the keyed sponge construction. In: *Symmetric Key Encryption Workshop (SKEW)*. February 2011.
- [19]Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Duplexing the sponge: single-pass authenticated encryption and other applications. *Cryptology ePrint Archive*, Report 2011/499. 2011.
- [20]Garcia, F. de Koning Gans, G. Verdult, R. Wirelessly lockpicking a smart card reader. *International Journal of Information Security*, 13(5):403–420, 2014.
- [21]Nohl, K. Evans, D. Starbug, S. Plötz, H. Reverseengineering a cryptographic RFID tag. In *USENIX security symposium*, volume 28, 2008.
- [22]Verdult, R. Garcia, F. Ege, B. Dismantling Megamos crypto: Wirelessly lockpicking a vehicle immobilizer. In *Supplement to the 22nd USENIX Security Symposium (USENIX Security 13)*, pages 703–718. USENIX Association, August 2013.

SOFTWARE PARA RECOLECCIÓN DE EVIDENCIAS DIGITALES RÁPIDAS EN SISTEMAS WINDOWS

Autores: Lic. Marcos Adrián Monti, Dr. David Luis La Red Martínez / Facultad de Ciencias Exactas y Naturales y Agrimensura / Universidad Nacional del Nordeste

RESUMEN

La necesidad de análisis de evidencia digital forense comprende ciertos desafíos en cuanto a escasez de herramientas de software disponibles, los altos costos y sobre todo los tiempos requeridos para llevar a cabo estos análisis, los cuales dependen en gran manera de la capacidad de los equipos involucrados. Resulta conveniente disponer de la capacidad de capturar información en el lugar del hecho utilizando las herramientas de software disponibles de manera inteligente de modo que sea posible discernir, cuando exista un gran volumen de equipos, cuáles podrían ser de interés a la causa judicial y cuáles no. En el presente trabajo se lleva a cabo el estudio de necesidad y factibilidad que fundamentan el proceso de desarrollo de un software de aplicación de campo para obtención de evidencia digital forense durante procedimientos de allanamiento; de acuerdo al protocolo de actuación vigente, este software será utilizado, además, como un “indicador” del contenido de los equipos bajo análisis, a fin de permitir al perito actuante tomar decisiones (como proceder al secuestro o no de un determinado equipo) en el lugar donde se está llevando a cabo dicho procedimiento.

Palabras clave:

Actividad de Usuario. Sistema Operativo. Navegador Web. Evidencia Digital. Informática Forense. Virtualización. Imagen Forense.

CONTEXTO

La Problemática se desarrolla en el contexto del análisis forense de evidencias digitales en el Poder Judicial, uno de los organismos públicos de la Provincia de Formosa.

La informática forense o el cómputo forense es el uso de métodos y técnicas científicas probadas, con el fin de identificar, preservar, validar, analizar, interpretar, documentar y presentar evidencia digital obtenida a partir de fuentes de información digital, con el propósito de facilitar la reconstrucción de hechos en una investigación legal, o ayudar a anticipar o prevenir acciones en contra de la ley [1][2]. De esta manera la Informática Forense actúa como una rama de la Informática que provee un complementario a la Criminalística clásica, enfocándose en el análisis de las evidencias digitales, que pudiesen existir, en cualquier escena del hecho [3][4].

En el campo de la criminalística, la tecnología y los sistemas de información resultan fundamentales para las ciencias forenses en el desarrollo de su objetivo que es descubrir y aportar pruebas físicas de los crímenes, a fin de esclarecer los hechos, la identidad de los sujetos que participaron en él, y la forma en que éstos llevaron a cabo la concreción de los mismos.

1. INTRODUCCIÓN

La evolución constante de las tecnologías ha venido de la mano o acompañado el desarrollo en todas las áreas de conocimiento a través de los años; el campo de los sistemas de información, específicamente, ha sido fundamental a la hora de asistir y acompañar a todas las ciencias, en el desarrollo de sus correspondientes actividades.

En la actualidad, los discos de equipos tipo PC que ejecutan alguna versión de Microsoft Windows tienen capacidad de 1 TeraByte o más, por lo cual la envergadura de información que albergan los mismos puede resultar muy significativa, resultando una tarea muy ardua y demandante la recopilación y análisis de datos.

En el laboratorio forense, es común realizar varias imágenes o copias forenses de los discos de los equipos aportados, con su correspondiente cálculo de valor hash, a fin de proceder a analizar los mismos [5], utilizando virtualización del sistema operativo [2] o algún software específico que permita “montar” los discos y sus particiones, conjuntamente con su sistema de archivos, sin tener que utilizar ni modificar el disco físico original [4][6].

Esta práctica, sin embargo, suele demandar cantidades enormes de espacio de almacenamiento en cualquier Laboratorio Forense, del orden de los cientos de Terabytes de datos; proporcional a la cantidad de equipos involucrados que han sido aportados en la causa en particular.

Asimismo los equipos necesarios para llevar a cabo la virtualización de los sistemas operativos y el análisis de datos, normalmente requieren tener la mayor cantidad posible de recursos disponibles en cuanto a procesador, memoria RAM y capacidad de almacenamiento en disco, siendo necesario en muchas ocasiones contar además con un servidor de almacenamiento en red o unidad NAS, instalado preferiblemente con redes cableadas y conexión de alta velocidad para cada uno de los equipos de trabajo.

Actualmente, existen numerosas herramientas de software capaces de realizar la imagen forense de los discos rígidos [8], para luego proceder al análisis de los mismos, según los criterios establecidos por el perito a cargo; sin embargo las mismas en su mayoría son muy exigentes en cuanto a costo, tiempo demandado y hardware requerido.

Por todos estos motivos, cualquier investigación que estuviera en curso puede prolongarse en el tiempo mucho más de lo deseado; es oportuno entonces contar con alternativas viables que sean de bajo costo, con pocas exigencias de hardware y demanden el menor tiempo posible.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para llevar a cabo un proceso de investigación y desarrollo primeramente se definieron cuáles serían los **objetivos Generales abordados**:

Los objetivos del trabajo consistieron en el estudio, recopilación y análisis de datos que son factibles de recuperar en sistemas que ejecutan el sistema operativo Microsoft Windows sobre la actividad de el/los usuario/s del mismo, con el objetivo de desarrollar herramientas de software que permitan automatizar las tareas de captura de información.

Los temas que se abordaron han sido todos aquellos concernientes a los detalles específicos de los sistemas Windows, su funcionamiento general, el almacenamiento de datos de usuario y cómo los mismos podrían aportar soluciones en la búsqueda de información útil para investigaciones forenses y el desarrollo de herramientas que puedan servir para tal fin. La investigación resultó en la concepción y desarrollo de un sistema automatizado para recolección de datos forenses en sistemas Windows, de acuerdo a los principios fundamentales de dicha ciencia [7][8].

Datos y aplicaciones comúnmente disponibles en Windows

De acuerdo a lo que se visualiza en la mayoría de sistemas Windows así como en lo sugerido desde sitios web comunes como “www.makeuseof.com” [9], “es.digitaltrends.com” [10], “www.microsoft.com” [11] y “www.computerworld.com” [12], es posible constatar la presencia de cierto software de base en todas las instalaciones de Windows.

Estos poseen, por ejemplo, Microsoft Internet Explorer o Microsoft Edge, por lo general, el cual ya está incluido desde Windows 10, asimismo existen navegadores alternativos que son de uso popular como ser Google Chrome, Opera o Firefox.

Este conjunto de navegadores, por lo general, constituyen las vías principales por las cuales el usuario realiza su interacción con redes

sociales además de permitir la visualización de imágenes, videos, noticias, etc.

Asimismo a través del análisis de datos resguardados en caché y sus correspondientes archivos de historial, es posible recuperar el registro de búsquedas o consultas realizadas en los mismos, a los diferentes motores de búsqueda online, obteniendo datos extra como el tipo de búsqueda realizada (texto, imágenes, videos), y la fecha y hora en que se realizaron. En cuanto a otros programas incluidos en instalaciones Windows podemos encontrar el editor de texto y explorador de archivos, reproductor de audio y videos a través de Windows Media Player y archivos de imágenes a través de Visualizador de Fotos o simplemente Fotos de Microsoft, los cuales se incluyen por defecto en cualquier instalación de este Sistema Operativo.

Es necesario destacar que, si bien existen aplicaciones instalables a través de la Microsoft store, en las versiones más modernas de Windows, como ser Instagram, Tik Tok, Telegram y WhatsApp Desktop, etc; el análisis de las mismas excede los alcances de los objetivos que fueran planteados durante el desarrollo del presente trabajo, por lo cual en mayor medida, se ha priorizado capturar la actividad del usuario a través de los navegadores de Internet.

Las herramientas mencionadas anteriormente poseen, adicionalmente, información privada sobre el/los perfiles de usuario guardados por cada red social, sitio, cuenta bancaria, etc., a la que el mismo ha accedido. Debido a lo mencionado, además de datos de navegación y marcadores, tendríamos datos sobre las cuentas resguardadas en dicho navegador, siempre y cuando el usuario haya decidido recordar las mismas en dicho navegador y dispositivo.

Software y aplicaciones para extracción de datos en Windows

De acuerdo a lo planteado anteriormente, es posible concluir que, en forma general, es posible obtener bastante información sobre el uso del sistema, a partir de los registros que guarda el propio Sistema Operativo.

Asimismo, la información básica sobre cuentas de usuario, Interacción en redes sociales y otros sitios de internet, se obtendrá únicamente a través de los mismos Navegadores Web, por lo tanto es conveniente centrar el estudio en la obtención de dichos datos; existen numerosos recursos y métodos para obtener estos datos, sin embargo uno de los recursos más mencionado por los sitios web especializados [13] el de las numerosas herramientas gratuitas disponibles en el sitio de herramientas de Nirsoft [14].

Es posible observar la gran colección de software especializado en herramientas para Windows.

Entre ellas se cuentan con herramientas tanto para Obtención de datos en Navegadores Web como para obtención de Información del sistema.

Diseño de la aplicación “Forensic Portable”

De acuerdo a lo mencionado oportunamente, se procedió al diseño y desarrollo de una aplicación de tipo portable, capaz de ejecutarse sin requerir ningún tipo de instalación; liviana, tanto en términos de requerimientos de hardware como en performance y tamaño en disco y con una interfaz simple, buscando un uso intuitivo y entendible, incluso para personal no técnico.

De acuerdo al uso para el cual fue concebida, esta aplicación no incluye un módulo para realizar análisis de superficie y búsqueda de archivos borrados, al ser una operación que, por lo general, demora mucho tiempo y para lo cual se requerirá el secuestro de la unidad a fin de realizar un análisis más extenso, en el Laboratorio Forense.

Se utilizó para esto el entorno de programación de Visual Studio 2019 y el lenguaje C# conjuntamente con las herramientas mencionadas de Nirsoft y la herramienta de Copia incorporada Robocopy.

RESULTADOS OBTENIDOS

Se ha obtenido como resultado una aplicación de tipo portable capaz de ejecutarse en

sistemas Windows desde cualquier dispositivo o unidad de disco portátil.

La aplicación muestra dos secciones donde podemos elegir el destino (carpeta o directorio) donde van a resguardarse todos los datos, y por otro la opción de copia forense de archivos o recuperación de datos del sistema (ver Fig. 1).

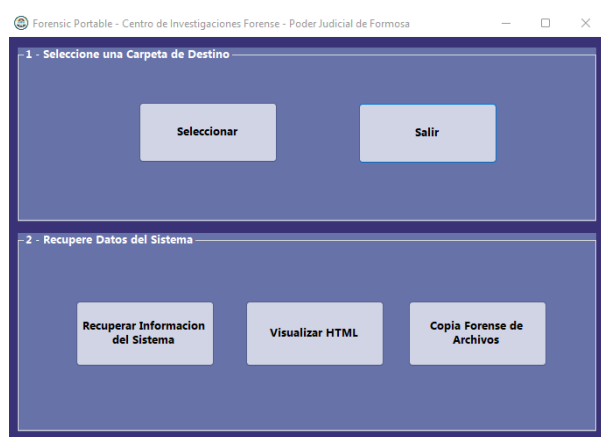


FIG. 1: Ventana Principal de la aplicación “Forensic Portable” [fuente propia].

El primer formulario que se visualiza en “1- Seleccione Carpeta de Destino” se encarga de establecer las configuraciones básicas para iniciar la recuperación de datos, este paso resulta necesario para establecer el destino de los datos, para lo cual se requiere un directorio válido.

A fin de mantener lo más inalterado posible el sistema huésped, es altamente recomendable elegir un dispositivo de almacenamiento externo como destino, en lugar de carpetas ubicadas dentro del sistema de archivos del equipo host.

Módulo de Recuperación de datos del Sistema

A la derecha es posible visualizar información básica del sistema y a la izquierda seleccionar el tipo de información que se busca recuperar. Desde aquí es posible obtener información de uso del sistema mediante los datos proporcionados tanto por los navegadores Web como de los registros del sistema operativo (Fig. 2).

Dicha información es obtenida a través del lanzamiento o ejecución de las aplicaciones mencionadas en capítulos anteriores en “modo silencioso”, es decir, a través de la ventana de comandos, sin utilizar ningún tipo de interfaz de usuario; esto permite recabar los datos de forma eficiente y rápida, al no requerir ningún otro tipo de interacción adicional, además es posible seleccionar, en el mismo módulo, qué procesos específicos lanzar, de acuerdo a la información que se quiere obtener.

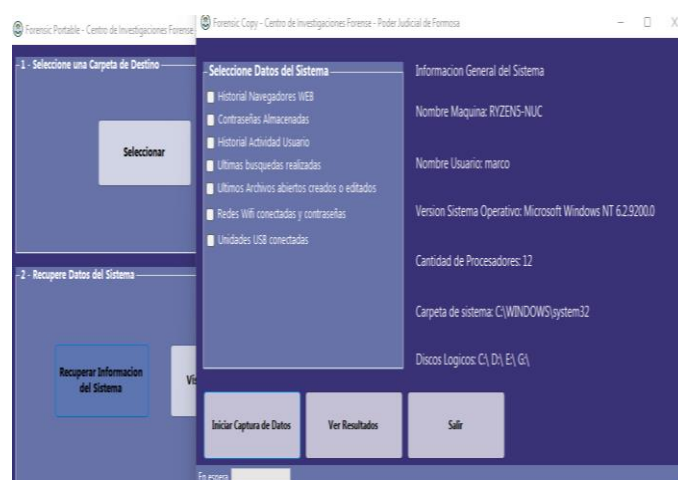


FIG. 2: Ventana módulo de recuperación de datos del sistema [fuente propia].

Módulo de Copia Forense

Finalmente en el último botón del menú principal tendremos el módulo de copia forense; en el mismo tendremos la posibilidad de realizar una copia forense [9][10] de todos los archivos que resulten de interés, junto con su información básica correspondiente (tipo de archivo, tamaño, fecha de creación, etc.), además de permitir obtener una pre visualización de los mismos, utilizando controles especiales incorporados al formulario.

Es posible seleccionar el tipo de archivo que se buscará en la unidad o carpeta de origen indicada.

En este formulario es posible efectuar búsquedas y visualización de archivos de tipo imágenes soportando los formatos - jpg, jpeg, png, gif, raw, cr2, nef, tif, tiff, hdr-, videos de tipo - mp4, mov, wmv, avi, mkv, flv, f4v, swf

-, documentos de tipo – doc, txt, pdf, htm, ppt, xls - y de audio como ser - m4a, ogg, mp3, flac, wav, wma, acc, alac o aif -.

La pre visualización se realiza en el mismo módulo, mediante un control especial por cada tipo de archivo, estos se abren en modo de solo lectura con la sola finalidad de permitir examinar o reproducir el contenido del mismo a fin de poder discriminar si estos serán de interés o no para la investigación en curso, para posteriormente proceder a la selección y, seguidamente copia de los mismos.

La copia de archivos se realiza de modo “forense” bit a bit, conservando todos los atributos originales del mismo, de modo que los metadatos que fueran incluidos en cada archivo permanecerán sin alteración alguna en el dispositivo destino de los mismos.

FORMACION DE RECURSOS HUMANOS

El desarrollo se lleva a cabo en el Departamento de Informática del Laboratorio de Criminalística del Centro de Investigaciones Forense del Poder Judicial de Formosa.

En el marco de este proyecto se desarrolla la Tesis para la Maestría en Tecnologías de la Información de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional del Nordeste contando como director de la misma al Dr. David Luis la Red Martínez.

BIBLIOGRAFÍA

- [1] Da Rocha, J. - de Luca. J. A. (2014) Informática y Delito. Reunión preparatoria del XIX Congreso Internacional de la Asociación Internacional de Derecho Penal (AIDP).
- [2] Piccirilli, D. (2013). La forensia como herramienta en la pericia informática. Revista Latinoamericana de Ingeniería de Software, I(6):237-240.
- [3] Presman, G. D. (2011) Investigación forense en redes sociales. Presentado en el XV

Congreso Iberoamericano de Derecho e Informática, Buenos Aires.

- [4] Presman, G. D. (2014) ISO/IEC 27037. Normalizando la práctica forense informática. Power point. Recuperado a partir de http://www.copitec.org.ar/comunicados/CAIF_2014/CAIF-Presman.pdf.

[5] Salas Ordinola, E., Ramírez García, A.; Núñez Mori, O. (2011) Propuesta de Protocolo para la Recolección de Evidencias Digitales Relacionado con la Legislación Peruana. Pontificia Universidad Católica del Perú. Publicado en alfa-redi, portal de Derecho y Nuevas Tecnologías.

- [6] Acurio del Pino, S. (2010) Manual de manejo de evidencias digitales y entornos informáticos, versión 2.0. AR: Revista de derecho informático.

[7] Bonilla, J. E. (2009) Principios de computación forense. Power point.

- [8] J. L. Rivas López, Análisis forense de sistemas informáticos, Cataluña: Eureka Media, SL, 2009.

[9] Ben Stegner, Digital Article (July 2021), MUO Online Technology Publications, <https://www.makeuseof.com/new-windows-pc-must-have-applications-to-install-first/>

- [10] Revista online Digital Trends, <https://es.digitaltrends.com/computadoras/mejores-apps-para-windows/>

[11] Artículo online, sitio web oficial de Microsoft, <https://www.microsoft.com/en-us/store/collections/essentialapps>

- [12] Revista Digital Computerworld, <https://www.computerworld.com/article/3602030/top-30-free-cheap-apps-for-windows-10.html>

[13] Artículo online, sitio Web de Geekflare, <https://geekflare.com/es/nirsoft-utilities/>

- [14] Sitio Web de Nirsoft, <https://www.nirsoft.net>

PROTECCIÓN DE ACTIVOS DE SOFTWARE

Julieta Gatica⁽¹⁾, Mario Berón⁽¹⁾, Daniel Riesco⁽¹⁾, Maria Joao Pereira⁽²⁾, Pedro Henriques⁽³⁾, Paulo Novais⁽³⁾

⁽¹⁾Universidad Nacional de San Luis, San Luis, Argentina

⁽²⁾Instituto Politécnico de Braganza, Braganza, Portugal

⁽³⁾Universidade do Minho, Braga, Portugal

jag81295@gmail.com, mberon@unsl.edu.ar, driesco@unsl.edu.ar, mjoao@ipb.pt, pjon@di.uminho.pt,
pedrorangelhenriques@gmail.com

RESUMEN

Una organización posee mucha información importante, conocida usualmente como activo. Estos activos pueden estar de forma física, escrita o impresa en papeles; ó de forma digital, a través de medios electrónicos. Considerando este segundo punto, se puede decir que en el software se implementan varios procesos de negocio, se utilizan algoritmos para mejorar el rendimiento, se manejan estrategias para solucionar problemas críticos, entre otras cosas. Por estas razones, el software se considera una fuente de activos muy importante.

Debido a que el software está sujeto a ataques constantes de usuarios maliciosos, se debe tener una constante actualización en materia de seguridad informática y, por ello, es necesario protegerlo utilizando distintas técnicas de protección de software.

En este artículo se presenta una línea de investigación que tiene como principal objetivo la definición de técnicas de protección y el diseño e implementación de lenguajes específicos del dominio orientados a resguardar los activos del software.

Palabras clave: Lenguaje de protección - Activos - Organización - Ofuscación - Marca de agua - Marca de nacimiento

CONTEXTO

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de Software (LaCIS) de la Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto: “Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”, perteneciente a la misma. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional.

1. INTRODUCCIÓN

El avance tecnológico ha llevado a que, a través de dispositivos como computadoras y celulares, se hagan una gran cantidad de actividades que antes eran inimaginables. Hoy en día, imaginar una vida sin estos dispositivos es casi imposible, ya que se utilizan como medio de comunicación y diversión, como medio de trabajo, ayuda a que se puedan hacer actividades de manera sencilla y sin mucho esfuerzo, se pagan impuestos a través de las aplicaciones web, los estudiantes tienen acceso a sus contenidos académicos mediante el uso de aulas virtuales, se realizan transferencias bancarias, las empresas permiten a sus empleados hacer home office, etc.

Aunque no se vea a simple vista, todos los beneficios antes mencionados se dan gracias a la implementación del software, el cual se ha popularizado convirtiéndose en una herramienta muy demandada por empresas, personas, instituciones educativas, etc. De este modo, el software ha pasado de ser una herramienta sólo utilizada por algunos a estar al alcance de cualquier persona.

Sin embargo, todas estas actividades deben realizarse a través de un software que sea seguro, ya que cualquier fallo en las aplicaciones puede causar desastre.

Las empresas de desarrollo de software, en la actualidad, realizan mucho esfuerzo en producir soluciones de calidad y cada vez más seguras. A pesar de ello, los crackers que poseen alta capacidad intelectual hacen uso de sus habilidades para vulnerar las estrategias de protección y, como consecuencia, las empresas deben seguir adaptando sus procesos de desarrollo para incorporar prácticas de seguridad acordes para dificultar la violación de sus sistemas.

En este sentido, Microsoft elaboró SDL (Security Development Lifecycle), que introduce aspectos de privacidad y seguridad en todas las fases de desarrollo, permitiendo a los desarrolladores enfocarse en la creación de un software altamente seguro, capaz de cumplir con los requisitos de seguridad y reducir costos de desarrollo.

Oracle, por su lado, desarrolló OSSA (Oracle Software Security Assurance), que es una metodología para incorporar seguridad en todo el proceso de desarrollo de sus productos: diseño, construcción, pruebas y mantenimiento, cumpliendo con los requisitos de seguridad y brindando una experiencia de propiedad más rentable.

OWASP Foundation describe SAMM (Software Assurance Maturity Model), un modelo de madurez útil para proporcionar una forma efectiva y medible de analizar y mejorar

el ciclo de vida de desarrollo seguro, ya que admite de manera completa el ciclo de vida del software y es independiente de la tecnología y los procesos que se utilizan.

SAMM es de naturaleza evolutiva, está impulsada por el riesgo y es un marco abierto para ayudar a las organizaciones a formular e implementar una estrategia para la seguridad del software. Ayuda a i) evaluar las prácticas de seguridad de software existentes de una organización; ii) construir un programa de garantía de seguridad de software equilibrado en iteraciones bien definidas; iii) demostrar mejoras concretas a un programa de aseguramiento de la seguridad; y iv) definir y medir las actividades relacionadas con la seguridad en toda la organización.

Con las mejoras mencionadas con anterioridad, el software ha aumentado su calidad y seguridad. Sin embargo, todavía sigue siendo vulnerable, es decir, toda estrategia de seguridad tarde o temprano puede ser violada.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La Protección del Software asume que no existen estrategias totalmente seguras, y que siempre el cracker puede encontrar la forma de superarlas. Por lo tanto, es una disciplina destinada a proteger el software de los ataques constantes de los crackers. Se entiende por protección a la estrategia de retrasar el éxito del ataque el mayor tiempo posible. Para lograr este objetivo, la Protección de Software utiliza varias estrategias como: *ofuscación de código*, *marca de agua de software*, *huellas dactilares*, *marca de nacimiento*, etc.

Cada estrategia tiene su campo de aplicación, por ejemplo, la *ofuscación de código* es útil para evitar el análisis del programa. Es bien sabido que los especialistas en informática

pueden comprender rápidamente el software con una alta calidad y solo necesitan crear herramientas de análisis de programas para acelerar el proceso. La ofuscación oculta las propiedades del software al agregar ruido al código fuente, lo que dificulta la extracción de información a través de herramientas de análisis de programas y por lo tanto retrasa el objetivo del cracker.

Las *marcas de agua*, las *huellas dactilares* y las *marcas de nacimiento* de software tienen como objetivo detectar y rastrear la distribución ilegal de software. Por ejemplo, es posible detectar la manipulación del software agregando una marca de agua dentro del código fuente.

En la presente investigación, se realizó un estado del arte acerca de las técnicas de protección de software y la utilización de las mismas.

Muchos autores presentan trabajos acerca de cuáles son las principales técnicas de protección utilizadas y cómo las mismas se utilizan dentro del ciclo de vida del software[1] para elaborar una alta tolerancia a fallos en el mismo[5] y mejorar su seguridad. Los autores mencionan que las técnicas más utilizadas son: serial number, empaquetado, protección de máquina virtual, antidepuración[1], strong name, watermarking, software dog, code obfuscation[8], entre otros.

Por otro lado, existen estudios que hacen uso de las técnicas antes mencionadas, y las combinan para proteger diversas partes del software. Algunos las utilizan a nivel de clase[2]; otros utilizan una combinación de algoritmos, ofuscación y marca de agua en una serie de pasos establecidos[7,10]; otros proponen métodos para la construcción o mejora de las estrategias de protección ya existentes[9,13,14,17]; otros establecen un método de cifrado para proteger el derecho intelectual[6], o utilizan criptografía para generar un mecanismo de protección más

novedoso[4]; y otros especifican modelos para medir la resistencia de un software dedicado a proteger sistemas[16].

Además de las estrategias ya mencionadas hay otras que utilizan una combinación de software y hardware, que también se construyen con el objetivo de evitar la manipulación. Hay autores que proponen un nuevo esquema de resistencia a la manipulación basado en encriptación y haciendo uso de un dispositivo USB[3].

Por otro lado, hay autores que crean su propia técnica o modelo de protección[11,12,15].

Dado el estado del arte realizado hasta el momento podemos concluir en que, si bien se han estudiado diversas formas de proteger el software, no existe aún un lenguaje que, en base a diversas primitivas de protección y funciones de alto nivel, genere estrategias de protección.

En la actualidad, los integrantes de la presente línea de investigación, han implementado algunas primitivas de protección con la intención de utilizarlas en un futuro. Estas primitivas serán utilizadas en la definición de un lenguaje específico del dominio que, haciendo uso de funciones de alto nivel como map, filter, reduce, etc., se puedan combinar las mismas para crear estrategias de protección.

3. RESULTADOS ESPERADOS

A través del trabajo realizado por los integrantes de la presente línea de investigación, se esperan obtener los siguientes resultados:

- Obtener un estado del arte aún más amplio acerca del uso de técnicas y estrategias de protección de software.
- Obtener conocimientos más profundos acerca del diseño y la construcción de lenguajes específicos del dominio.

- La generación de un lenguaje específico del dominio compuesto por diversas primitivas de protección que, haciendo uso de funciones de alto nivel, se combinen para crear estrategias de protección.

4. FORMACIÓN DE RECURSOS HUMANOS

Los progresos obtenidos en esta línea de investigación sirven como base para el desarrollo de tesis de posgrado, ya sea de doctorado o maestría en Ingeniería de Software de la Universidad Nacional de San Luis, en el marco de los proyectos de investigación mencionados en el contexto del presente documento.

5. BIBLIOGRAFÍA

- [1]. L. Guoyuan, T. Jiutao, and G. Bing. "A survey on shareware protection schemes". In: 2010 2nd International Conference on Signal Processing Systems. Vol. 3. 2010, pp. 697–700. doi: 10.1109/ICSPS.2010.5555515.
- [2]. X. Guangli and C. Zheng. "The Code Obfuscation Technology Based on Class Combination". In: 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science. 2010, pp. 479–483. doi: 10.1109/DCABES.2010.104.
- [3]. S. Zheng and J. Liu. "An USB-Key based approach for software tamper resistance". In: 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). Vol. 6. 2010, pp. 506–509. doi: 10.1109/ICACTE.2010.5579199.
- [4]. H. Lin, X. Mo, and Y. Gao. "Based on RSA and Self-Modifying Mechanism of Software Protection". In: 2010 International Symposium on Parallel and Distributed Processing with Applications. 2010, pp. 474–477. doi: 10.1109/ISPA.2010.12.
- [5]. L. Xiong and Q. Tan. "A Configurable Approach to Tolerate Soft Errors via Partial Software Protection". In: 2011 IEEE Ninth International Symposium on Parallel and Distributed Processing with Applications Workshops. 2011, pp. 260–265. doi: 10.1109/ISPAW.2011.45.
- [6]. J. Hu et al. "A key hiding based software encryption protection scheme". In: 2011 IEEE 13th International Conference on Communication Technology. 2011, pp. 719–722. doi: 10.1109/ICCT.2011.6157970.
- [7]. L. Chen and C. Zhang. "A novel algorithm for .NET programs watermarking based on obfuscation". In: 2012 International Symposium on Instrumentation Measurement, Sensor Network and Automation (IMSNA). Vol. 2. 2012, pp. 583–586. doi: 10.1109/MSNA.2012.6324652.
- [8]. B. Chen et al. "Cloud licensing model for .NET software protection". In: 2012 7th International Conference on Computer Science Education (ICCSE). 2012, pp. 1069–1074. doi: 10.1109/ICCSE.2012.6295248.
- [9]. Y. Wang et al. "CHI Based Instruction-Words Based Software Birthmark Selection". In: 2012 Fourth International Conference on Multimedia Information Networking and Security. 2012, pp. 892–895. doi:10.1109/MINES.2012.84
- [10]. C. Collberg, A. Gibson, S. Martin, N. Shinde, A. Herzberg and H. Shulman, "Provenance of exposure: Identifying sources of leaked documents". In: 2013

- IEEE Conference on Communications and Network Security (CNS), National Harbor, MD, USA, 2013, pp. 367-368, doi: 10.1109/CNS.2013.6682731.
- [11]. C. Basile and M. Ceccato, "Towards a unified software attack model to assess software protections". In: 2013 21st International Conference on Program Comprehension (ICPC), San Francisco, CA, USA, 2013, pp. 219-222, doi: 10.1109/ICPC.2013.6613852.
- [12]. A. Averbuch, M. Kiperberg and N. J. Zaidenberg, "Truly-Protect: An Efficient VM-Based Software Protection". In: 2013 IEEE Systems Journal, vol. 7, no. 3, pp. 455-466, Sept. 2013, doi: 10.1109/JSYST.2013.2260617.
- [13]. P. Chan, L. Hui, S. Yiu, "Dynamic Software Birthmark for Java Based on Heap Memory Analysis". In: 2011 Communications and Multimedia Security, pp. 94-107
- [14]. A. Kulkarni and R. Metta, "A New Code Obfuscation Scheme for Software Protection". In: 2014 IEEE 8th International Symposium on Service Oriented System Engineering, Oxford, UK, 2014, pp. 409-414, doi: 10.1109/SOSE.2014.57.
- [15]. A. A. Osti, B. A. P. Botelho and N. Uto, "Implementation aspects of software guards: A case study". In: 2014 The 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014), London, UK, 2014, pp. 262-267, doi: 10.1109/ICITST.2014.7038818.
- [16]. S. Banescu, C. Collberg, V. Ganesh, Z. Newsham, A. Pretschner. "Code obfuscation against symbolic execution attacks". In: 2016 Proceedings of the 32nd Annual Conference on Computer Security Applications, pp. 189-200. doi: 10.1145/2991079.2991114.
- [17]. B. F. Demissie, M. Ceccato and R. Tiella, "Assessment of Data Obfuscation with Residue Number Coding". In: 2015 IEEE/ACM 1st International Workshop on Software Protection, Florence, Italy, 2015, pp. 38-44, doi: 10.1109/SPRO.2015.15.

Técnicas para Incrementar la Seguridad en Web Services basados en WSDL

Edgardo Bernardis⁽¹⁾, Hernán Bernardis⁽¹⁾, Mario M. Berón⁽¹⁾, Daniel E. Riesco⁽¹⁾,
Pedro Rangel Henriques⁽²⁾, Maria Joao V. Pereira⁽³⁾

⁽¹⁾ Departamento de Informática / Facultad Ciencias Físico Matemáticas y Naturales/
Universidad Nacional de San Luis
Ejército de los Andes 950 – San Luis – Argentina
{hbernardis, ebernardis, mberon, driesco}@unsl.edu.ar

⁽²⁾ Departamento de Informática/Universidade do Minho
Braga – Portugal
pedrorangelhenriques@gmail.com

⁽³⁾ Departamento de Informática e Comunicações/ Instituto Politécnico de Bragança
Bragança - Portugal
mjoao@ipb.pt

RESUMEN

En los tiempos que corren ha cambiado de manera considerable la forma en que las personas y empresas interactúan e intercambian información. Este escenario lo vuelve un blanco importante para todo tipo de actores que desean obtener información útil y valiosa para sí mismos o de terceros. Ante esta situación, es sumamente imprescindible implementar todo tipo de medidas y acciones tendientes a evitar estos ataques. Toda acción, herramienta o metodología enfocada a evitar, contrarrestar o retrasar ataques contra activos sensibles hace referencia a lo que se denomina Seguridad Informática [1].

Actualmente muchos sistemas de software son en realidad una agrupación de servicios web en la nube que se invocan cuando es necesario obtener la información que los mismos proveen. Es por ello, que este artículo, se enfoca en comprender y mejorar la seguridad de los servicios web que utilizan un WSDL (Web Services Description Language) [2] como medio para proveer servicios a través de la red.

Palabras Clave: Web Services, Métricas, Comprensión, WSDL, Ofuscación, Seguridad.

CONTEXTO

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de Software (LaCIS) de la Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto: “Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”, perteneciente a la misma. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional.

1. INTRODUCCIÓN

El incremento en el uso de los servicios web dentro del desarrollo de sistemas es cada vez mayor. Muchas organizaciones construyen sus sistemas basándose en una arquitectura orientada a servicios web, algunos de ellos se publican al resto del mundo de manera

pública y libre, mientras que otros son utilizados internamente de manera privada por sus equipos de desarrollo y/o funcionalidades propias de la entidad.

Todo Web Service (WS) posee una especificación que provee la información necesaria para invocarlo y utilizarlo adecuadamente. Uno de los estándares de descripción más conocidos es WSDL (Web Service Definition Language) [3]. Las especificaciones WSDL son un dialecto XML, con reglas bien definidas para especificar cada componente del WS. Así como el archivo WSDL sirve para que un agente de software o persona pueda interpretar para luego usar el servicio web que describe, también puede dar información a personas no deseadas o incluso exponer vulnerabilidades de la organización que lo creó y que lo utiliza. Esto se torna sumamente importante para aquellos casos en donde los servicios web pertenecen a bancos, tarjetas de créditos, servicios de compra/venta online, entre otros. Incluso también para los servicios web que no se publican, son privados y necesitan mayor control y seguridad como los que pertenecen a empresas privadas y redes militares. Esto no solo se trata de competencia, los ataques de seguridad como espionaje de información, suplantación de clientes, inyección de comandos y denegación de servicio también son posibles ya que los atacantes pueden aprender sobre los datos intercambiados y los patrones de invocación de los documentos WSDL. Si bien la legibilidad de las descripciones de los servicios hace que los servicios web sean reconocibles, también contribuye a la vulnerabilidad del servicio [4]. Todos contienen información formal (código fuente) e informal (identificadores, comentarios, documentación, etc.) y es en este tipo de información en donde los atacantes hacen foco para obtener información beneficiosa a sus propósitos. Suena lógico entonces incrementar la seguridad que posee un determinado WSDL para evitar e impedir los ataques.

Si se supone la situación particular de un banco que posee múltiples servicios web en su sistema con arquitectura orientada a servicios y desea determinar qué tan vulnerables son a posibles ataques informáticos. Será importante responder a las siguientes preguntas ¿Cómo puede analizar las vulnerabilidades del mismo? ¿Cómo determinar qué tan comprensibles son a personas externas que deseen encontrar vulnerabilidades en sus descripciones para atacar? Ambos escenarios desnudan la necesidad de técnicas de comprensión de servicios web y mediante las mismas, aplicar técnicas para mejorar la seguridad de los WSDLs.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este equipo de investigación trabaja actualmente en dos líneas principales de estudio. Una de ellas dedicada a la Extracción y Análisis de la Información y la otra enfocada en la Seguridad de la Información. Estas dos líneas se combinan en este trabajo de regulación en la seguridad de servicios web con especificaciones WSDL. A continuación, se describen las líneas de investigación antes mencionadas de manera resumida destacando las principales actividades llevadas a cabo en esta investigación.

Línea de Extracción y Análisis

Esta línea se enfoca en obtener información de la especificación o el código fuente de los sistemas, procesarla y analizarla para generar resultados a partir de la misma. Dentro del contexto de especificaciones WSDL, el análisis y reducción de vulnerabilidades de las especificaciones WSDL posee múltiples etapas. Cada una con su función particular dentro del proceso global. En las subsecciones siguientes se describen brevemente dichas etapas.

Extracción de Información

Debido a que las especificaciones WSDL son un dialecto XML, se pueden usar técnicas de compilación sobre las mismas basadas en el parser DOM (Domain Object Model) [5]. Un parser DOM construye una representación interna del WSDL basada en árboles. A partir de funciones específicamente diseñadas para recorrer la representación construida (funciones transversales) se extrae la información deseada. Estas funciones transversales logran extraer los identificadores de cada componente (*name*, *type element*, etc.), la documentación y los comentarios presentes en el WSDL. Adicionalmente, también se extrae información complementaria como cantidad de líneas del archivo, cantidad de espacios de nombres, versión del WSDL, entre otras.

Preprocesamiento de la Información

Este proceso prepara los datos extraídos para luego realizar el cálculo de las métricas. Esta etapa implica, por ejemplo, quitar las stop-words, separar las palabras de los nombres de los identificadores y calcular valores de base.

Cálculo de Métricas

De la información extraída del WSDL, se calculan múltiples métricas, que pueden ser lógicamente diferenciadas en los siguientes grupos:

- **Métricas de tamaño:** miden la complejidad del WSDL en base a las cantidades de componentes de cada etiqueta dentro del WSDL, como por ejemplo cantidad de tipos complejos, de parámetros, de operaciones, de mensajes, entre otras. Esto permite tener una idea del tamaño de cada sección particular del WSDL y determinar qué tan complejo es, a primera vista, su comprensión.
- **Métricas de calidad:** permiten medir la calidad semántica de la especificación WSDL. Esto es, que

tanta información semántica brinda la especificación WSDL respecto del WS que representa y que tan entendible y comprensible en sí es dicha especificación. Alguno de los parámetros que se miden son calidad de palabras, cantidad de correcciones por palabras y palabras en diccionario.

- **Métrica de entendimiento global:** usando LSP (Logic Scoring of Preference) se calcula el grado de entendimiento que posee la especificación WSDL de un WS [6, 7]. Esto se realiza usando como base todas las métricas calculadas relacionadas mediante estructuras de agregación específicamente creadas para este fin.

Toda aplicación web está conformada por distintos tipos de información, tanto formal como informal. El análisis detallado de la misma y el cálculo de métricas permite detectar, dado su grado de entendimiento, que partes son más susceptibles a los ataques. En este punto, es posible definir estrategias que permitan subsanar las vulnerabilidades y proteger las partes que sean susceptibles de ataques [8]; en este caso particular, ataques dirigidos a WSDLs.

Línea de Seguridad de la Información

Esta línea se enfoca en trabajar la seguridad de la información diseñando y aplicando diferentes técnicas y herramientas diseñadas para tal fin.

En el caso de este trabajo, se aplica ofuscación, transformación de código y otras modificaciones al WSDL para incrementar su seguridad.

Ofuscación de Código

Según el Diccionario de la Real Academia, ofuscar significa deslumbrar, turbar la vista, oscurecer, trastornar o confundir las ideas [9]. Es decir, se refiere a encubrir deliberadamente el significado de alguna cosa haciéndola más confusa y complicada

de interpretar, evitando la comprensión de la misma. La palabra ofuscación fue elegida para esta actividad porque connota oscuridad, ininteligibilidad y desconcierto, y porque ayuda a distinguir este enfoque de otros métodos [10]. Especificando más el concepto, Ofuscación de Código hace referencia a un conjunto de transformaciones de código que convierten un programa en uno funcionalmente equivalente, pero ininteligible haciendo difícil su entendimiento y realizar ingeniería inversa sobre el mismo. La ofuscación de código aplica una o más transformaciones de código que hacen que el código sea más resistente al análisis y la manipulación, pero preservan su funcionalidad [11].

Transformaciones de Código

La compilación de código se ha convertido en mucho más que solo traducir un programa de computadora en uno ejecutable. Por lo general, los programas se escriben en un lenguaje de alto nivel, dadas las ventajas que estos ofrecen. Sin embargo, tienen ciertas desventajas que hacen que la compilación de código incluya implícitamente numerosas técnicas de optimización. Estas, incluyen eliminar código muerto, asignación de registros óptimas y asignaciones eficientes al objetivo conjunto de instrucciones de la arquitectura [11]. En estas etapas de compilación y/o desarrollo de software se pueden realizar y aplicar distintas transformaciones de código. Una buena ofuscación se compone de una o más transformaciones de código que alteran un programa de tal forma que resulte difícil aplicar técnicas de ingeniería inversa. La única restricción para estas transformaciones, sean manuales o automatizadas, es preservar la funcionalidad original del programa. Las transformaciones de código para ofuscar un programa se pueden dividir en cuatro clases principales: i) Léxicas o de Diseño ii) de Flujo de Control, iii) de Flujo de Datos y iv) Preventivas [12].

Incrementar la Seguridad

Utilizando la información extraída del WSDL se pueden manipular diferentes partes del mismo para mejorar su seguridad disminuyendo su nivel de entendimiento. Esto se puede lograr mediante la utilización de funciones de ofuscación y/o encriptación al realizar las modificaciones y/o transformaciones necesarias que aumentaran el nivel de seguridad. Estas transformaciones pueden ser sobre partes específicas del WSDL (identificadores, operaciones, etc.) o en la totalidad del mismo. Dichas modificaciones dependen del nivel de seguridad deseado, partiendo de un nivel básico en donde se ofuscan y/o encriptan partes específicas del WSDL, como por ejemplo el nombre de los identificadores, hasta llegar a un nivel máximo en donde se realiza una transformación completa del WSDL.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados más destacados obtenidos de esta investigación se mencionan a continuación:

- Se diseñaron estrategias para la extracción de información desde las especificaciones WSDL de los WS.
- Se definieron y crearon múltiples métricas de tamaño y semánticas para obtener información de los WSDLs.
- Se construyó la estructura de agregación LSP que permite vincular todas las métricas las métricas y calcular el grado de entendimiento global del WSDL.
- Se construyó WSDLUDTool, una herramienta que realiza la extracción de información, cálculo de métricas, cálculo del grado de entendimiento del WS usando LSP y la visualización de la información.
- Se modificaron representaciones WSDLs originales y ofuscadas y se le calculó el grado de entendimiento

a ambas para comprobar el cambio de comprensión.

Los objetivos planteados a corto y largo plazo son:

- Crear un módulo que permita recomendar automáticamente cambios en el WSDL para incrementar o decrementar el grado de entendimiento según sea necesario.
- Definir otras estrategias para modificar la comprensión de los WSDLs para hacerlos más seguros.
- Definir, analizar y combinar distintas técnicas y estrategias de ofuscación en WSDLs.

4. FORMACIÓN DE RECURSOS HUMANOS

Las tareas realizadas en el contexto de la presente línea de investigación están siendo desarrolladas como parte de trabajos para optar al grado de Magister en Ingeniería de Software. En el futuro se piensa generar diferentes tesis de maestría y doctorado usando como base parte de esta investigación.

5. BIBLIOGRAFÍA

- [1] ISO/IEC. Iso/iec 27032:2012 information technology - security techniques - guidelines for cybersecurity.
- [2] WSDL Specification for W3C <https://www.w3.org/TR/wsdl>
- [3] WSDL Specification for W3C. <https://www.w3.org/TR/wsdl>.
- [4] Pananya Sripairojthikoon, Twittie Senivongse. Concept-Based Readability Measurement and Adjustment for Web Services Descriptions. ICACT Transactions on Advanced Communications Technology (TACT) Vol. 3, Issue 1, January 2014.
- [5] Parser DOM specification for W3C. <https://www.w3.org/DOM>.
- [6] Mario M. Berón, Hernán Bernardis, Enrique A. Miranda, Daniel E. Riesco, Maria João Pereira, Pedro Rangel Henriques. "WSDLUD: a Metric to Measure the Understanding Degree of WSDL Descriptions". Proceedings of the 2015 Symposium on Languages, Applications and Technologies, SLATE'15. Madrid, España 2015.
- [7] Bernardis, Hernán; Berón Mario; Bernardis, Edgardo; Riesco, Daniel; Henriques, Pedro. "Extracción de información y cálculo de métricas en WSDL 1.1 y 2.0". II Congreso Nacional de Ingeniería Informática / Sistemas de información (CoNaIISI). Argentina. 2014.
- [8] Edgardo Bernardis, Hernán Bernardis, Mario Berón, Germán Montejano. "Seguridad en Servicios Web". XIX Workshop de Informática y Ciencias de la Computación (WICC). Buenos Aires, Argentina. Abril de 2017.
- [9] C. Collberg, C. Thomborson, and D. Low. A taxonomy of obfuscating transformations. Technical Report #148, Department of Computer Science, The University of Auckland, 1997.
- [10] Brunton, Finn and Nissenbaum, Helen. Obfuscation: A user's guide for privacy and protest. MIT Press. 2015.
- [11] Cappaert, Jan. Code obfuscation techniques for software protection. Katholieke Universiteit Leuven. 2012.
- [12] Collberg, Christian and Thomborson, Clark and Low, Douglas. A taxonomy of obfuscating transformations. Computer Science Technical Reports. The University of Auckland. 1997.

Avances sobre la Especificación Integral del Sistema OTP-Vote Orientada a su Implementación

Silvia Bast¹ Germán Montejano² Mario Berón²

¹Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-425166 – Int. 28
silviabast@exactas.unlpam.edu.ar – web: <http://exactas.unlpam.edu.ar/>

²Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
Tel.: +54-2652-424027 – Int. 251
[gmonte, mberon]@unsl.edu.ar – web: <http://www.unsl.edu.ar>

RESUMEN

El sistema electoral que rige en Argentina tiene sus bases en la Ley Sáenz Peña de 1912, que establece que el voto debe ser: universal, secreto y obligatorio. Si bien este sistema manual, que se usa actualmente en la mayoría de los distritos del país, funciona de manera aceptable, es también permeable a algunos vicios. Por otro lado, la incorporación del voto electrónico presenta grandes discusiones debido, especialmente, a falta de confianza del electorado. El principal desafío de este proyecto es aportar propuestas para la construcción de un sistema de votación electrónica robusto y confiable. El trabajo de investigación tiene sus orígenes en el modelo inicial de datos de un sistema de voto electrónico denominado OTP-Vote. Con vistas a lograr su implementación, se deben especificar en profundidad un conjunto de características que se mencionan como supuestos en el modelo inicial. En este artículo se describen los aspectos en los que se ha avanzado con el objetivo de obtener un sistema sólido y seguro.

Palabras clave: *Sistemas de Voto Electrónico, Anonimato, Transparencia, Auditoría, One Time Pad, Verificabilidad End to End.*

CONTEXTO

Este trabajo se enmarca en el Proyecto de Investigación denominado: "Especificación Integral del Sistema OTP-Vote Orientada a su Implementación", que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales (FCEyN) de la Universidad Nacional de La Pampa (UNLPam), Resolución N° 55/22 del Consejo Directivo FCEyN - UNLPam y es dirigido por el Doctor Germán Antonio Montejano (Universidad Nacional de San Luis) y por el Doctor Mario Berón (Universidad Nacional de San Luis) y es la continuación de la línea de investigación "Desarrollo de un Modelo de Voto Electrónico basado en Criptografía One Time Pad" del proyecto "Aspectos de Seguridad en Proyectos de Software, Resolución N° 488/14 del Consejo Directivo de FCEyN - UNLPam.

1. INTRODUCCIÓN

El sistema electoral argentino ha presentado algunas mejoras desde sus inicios, una de las más importantes es la aplicación de la Ley Sáenz Peña o Ley Nacional de Elecciones N° 8.871, de 1912 que estableció las características del voto que perduran hasta hoy: universal, secreto y obligatorio.

El sistema manual que se usa en la actualidad funciona de manera aceptable, aunque es permeable a algunos vicios tales como: robo de boletas, colocación al tope de todas las pilas de una boleta determinada, adulteración de actas y telegramas, embarazo de urnas, voto cadena, soborno a personas que realizan el conteo de los votos, entre otras. Además, en algunos países la cantidad de candidatos para cada cargo y también el gran número de votantes dificultan y lentifican el proceso de conteo para la publicación final de los resultados.

Ante este diagnóstico, la incorporación de los sistemas de voto electrónico, parece ser el paso natural en el proceso de transformaciones sociales que se produce día a día. Sin embargo, existen fuertes cuestionamientos del electorado a este tipo de sistemas, que radican especialmente en la falta de transparencia.

Las personas compran en línea, realizan operaciones bancarias online, se reúnen virtualmente, enseñan y evalúan a los estudiantes de manera virtual, pero no se sienten suficientemente confiados para hacer uso de un sistema de voto electrónico.

Estudiando las razones por las que la ciudadanía actúa de forma diferente ante el uso de los sistemas, surge que, los ciudadanos realizan con confianza operaciones en las que de alguna manera queda un comprobante o registro que vincula la actividad realizada con la identificación de quién la realizó. Por ejemplo, una operación bancaria con el documento de identidad de la persona o la clave única de identificación bancaria, o un legajo de estudiante con un examen. En cuanto a los sistemas de voto electrónico, los ciudadanos esperan que la identidad solo quede registrada para verificar la participación en los comicios, pero que de ninguna manera

pueda ser asociada al voto. Otra característica propia de estos sistemas, es que los electores una vez que emitieron su voto, ya no cuentan con el comprobante para hacer el reclamo porque el voto queda en la sede de votación. Estas situaciones no afectan únicamente a los sistemas de votación electrónica sino a todos los sistemas electorales en general.

Desde el equipo de trabajo que lleva adelante esta investigación, se asume que:

- Los sistemas de voto electrónico se clasifican dentro de la categoría de “críticos”.
- El sistema manual que se usa actualmente, tal como viene funcionando ofrece prestaciones aceptables.
- Resulta un desafío generar un modelo que permita el desarrollo de un sistema robusto y confiable.

Sistemas de Voto Electrónico

Según McGaley and Gibson [1] los sistemas de voto electrónico son “cualquier forma de recolección de votos que implique el uso de dispositivos electrónicos (usualmente computadoras)”. Los requisitos que deben cumplir estos sistemas son expresados por importantes autores como Epstein [2], Kazi, Alam y Tamura [3], Prince [4], van de Graaf, Henrich y Müller-Quade [5], Hao y Ryan [6], Rivest [7], Ryan, Schneider y Teague [8], Rabin y Rivest [9] y Awad y Leiss [10].

Teniendo en cuenta que uno de los principales cuestionamiento hacia estos sistemas es la falta de transparencia y confiabilidad, resulta de suma importancia que cuenten con la característica de verificabilidad extremo a extremo o verificabilidad end to end (VE2E). Esta característica permite que los electores puedan verificar que sus votos se emitieron y se contaron correctamente. Chaum [11], Ryan y Bryans [12], Chaum, Ryan y Schneider [13], Fisher, Carback, y Sherman [14], Adida y Rivest [15] y Benaloh et al [16] presentan propuestas para aportar verificabilidad a los sistemas de voto electrónico.

El Modelo OTP-Vote

En [17] se presentan las bases del modelo teórico OTP-Vote, que consiste de tres etapas, tal como se muestra en la Figura 1.

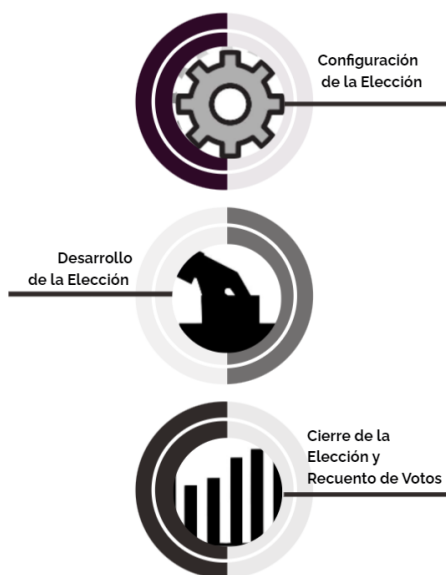


Figura 1. Etapas del Proceso Electoral en OTP- Vote

El modelo asegura:

- Anonimato incondicional
- Seguridad computacional que puede llevarse a cualquier nivel exigible durante el proceso electoral.

Para alcanzar estos objetivos usa:

- Claves One Time Pad (OTP) que cumplen con la característica de Secreto Perfecto de Shannon [18].
- El esquema de almacenamiento denominado Múltiples Canales Datos único (MCDU) y sus fórmulas propuestas para alcanzar dimensiones de los datos con comportamiento óptimo [19], [20], [21] y [22].
- La operación lógica XOR,
- La redundancia apropiada [23] en el almacenamiento de los datos.

El modelo teórico presentado supone, para cada una de las etapas mencionadas, el cumplimiento de condiciones de seguridad que

resultan imprescindibles para alcanzar el normal funcionamiento del sistema.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El modelo inicial debe ser refinado para lograr un sistema robusto e implementable. Este trabajo de refinamiento inicia con el análisis profundo de cada una de las etapas, para detectar cuáles son los aspectos críticos que deben mejorarse o especificarse, para posteriormente avanzar en propuestas superadoras que aporten soluciones a cada uno de ellos.

Será necesario profundizar en:

- La identificación y clasificación de los datos que deben permanecer inalterables durante el proceso y de aquellos que deben modificarse de forma controlada. Además, debe realizarse la búsqueda de mecanismos que permitan controlar los accesos de acuerdo a sus características.
- Propuestas de optimización de la configuración de las tuplas que almacenan los datos de los votos y los atributos de control.
- Especificación y validación de una propuesta de generación automática de tablas relacionales a partir de los datos del sistema.
- Análisis de la información que permita ofrecer transparencia al proceso a la vista de terceros, y especificación, validación y desarrollo de propuestas de auditoría.
- Especificación y validación de una propuesta de verificabilidad End to End.

3. RESULTADOS Y OBJETIVOS

Los avances de la investigación son:

- Se ha continuado y profundizado el trabajo de análisis de mecanismos que permitan realizar el acceso controlado

- a los datos de acuerdo a su nivel de criticidad.
- Se ha desarrollado una propuesta de configuración de las tuplas y los atributos de control haciendo especial énfasis en la seguridad de los datos de los votos.
- Se ha trabajado en la revisión sistemática de literatura en relación a VE2E.
- Se ha desarrollado y presentado una propuesta inicial de VE2E para el sistema OTP – Vote.
- Se comenzó a trabajar en la propuesta de generación automática de tablas de datos relacionales a partir de los datos obtenidos en la etapa de configuración de la elección.

Como trabajos futuros debe avanzarse en:

- Refinamiento de protocolos antifraude que deben usarse durante todo el proceso.
- Análisis y selección de un método criptográfico que asegure la transmisión de datos entre estaciones y servidor.
- Propuestas de auditoría.
- Evaluación y mejora de los avances ya realizados sobre el modelo original.

4. FORMACIÓN DE RECURSOS HUMANOS

En cuanto a la formación de recursos humanos, Silvia Bast está avanzando en el desarrollo de la tesis denominada “Especificación Integral del Modelo OTP-Vote orientada a su implementación” para alcanzar el grado de Doctora en Ingeniería Informática en la Universidad Nacional de San Luis. Resolución de Inscripción y Aprobación de Plan de Tesis 408/21 Decanato. FCFMyN – Universidad Nacional de San Luis.

5. BIBLIOGRAFÍA

- [1] M. McGale, J. Gibson, “A Critical Analysis of the Council of Europe Recommendations on E-Voting”, Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop 2006 on Electronic Voting Technology Workshop, USENIX Association.
- [2] J. Epstein, “Electronic Voting” in Computer, vol. 40, no. 8, pp. 92-95, Aug 2007. doi: 10.1109/MC.2007.271.
- [3] K. M. Rokibul Alam and S. Tamura, “Electronic voting - Scopes and limitations”, International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, Bangladesh, pp. 525-529, . 2012. doi: 10.1109/ICIEV.2012.6317324.
- [4] A. Prince, “Consideraciones, aportes y experiencias para el Voto electrónico en Argentina”, Editorial Dunken, 2006.
- [5] J.van de Graaf, C. Henrich, J. Müller-Quade, “Requirements for secure voting”, Work Notes 2011.
- [6] F. Hao, P. Ryan, “Real -World Electronic Voting. Design, Analysis and Deployment”. CRC Press. ISBN-13: 978- 1498714693. ISBN-10: 1498714692. 2017.
- [7] R. Rivest, “On the notion of ‘software independence’ in voting systems”. Philosophical Transactions of the Royal Society A, 366(1881):3759–3767. 2008.
- [8] P. Ryan, S. Schneider, V. Teague, “End-to-End Verifiability in Voting Systems, from Theory to Practice”. Voting Systems, from Theory to Practice. IEEE Security & Privacy, 13(3):59–62, 2015.
- [9] M. Rabin, R. Rivest, “Efficient End to End Verifiable Electronic Voting Employing Split Value Representations” Bregenz, Austria.

Proceedings of EVOTE 2014. ISBN 978-9949-23-688-6. 2014.

[10] M. Awad, E. Leiss, "End-to-End Cryptography: Spreading Democracy". International Journal of Applied Engineering Research. Volume 11, Issue 11. Ps. 7391-7394. 2016.

[11] D. Chaum, "Secret-ballot receipts: True voter-verifiable elections" IEEE security & privacy, IEEE, 2004, 2, 38-47

[12] P. Ryan, J. Bryans. "A simplified version of the Chaum voting scheme." School of Computing Science Technical Report Series. 2004.

[13] D. Chaum, P. Ryan, and S. Schneider. "A practical voter-verifiable election scheme." European Symposium on Research in Computer Security. Springer, Berlin, Heidelberg, 2005

[14] K. Fisher, R. Carback, and A. Sherman. "Punchscan: Introduction and system definition of a high-integrity election system." Proceedings of Workshop on Trustworthy Elections. 2006.

[15] B. Adida, and R. Rivest. "Scratch & vote: self-contained paper-based cryptographic voting." Proceedings of the 5th ACM workshop on Privacy in electronic society. 2006.

[16] J. Benaloh, R. Rivest, P. Ryan, P. Stark, V. Teague, & P. Vora, (2015). End-to-end verifiability. arXiv preprint arXiv:1504.03778.

[17] S. Bast, "Confidencialidad e Integridad de Datos en Sistemas de E-Voting – Un Modelo para la Implementación Segura de un sistema de Voto Presencial", Editorial Académica Española. ISBN 978-3-639-53793-2. 2017.

[18] C. E. Shannon, "Communication theory of secrecy systems," in The Bell System Technical Journal, vol. 28, no. 4, pp. 656-715,

Oct. 1949, doi: 10.1002/j.1538-7305.1949.tb00928.x..

[19] P. García, "Una Optimización para el Protocolo Non Interactive Dining Cryptographers" - Editorial Académica Española (<https://www.eae-publishing.com/> - ISBN-13: 978-3-639-85270-7. ISBN-10: 3639852702. EAN: 9783639852707 – 2017.

[20] J. van de Graaf, G. Montejano, P. García, "Manejo de Colisiones en un Protocolo Non Interactive Dining Cryptographers". Anales de las 42° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). Páginas 29 a 43. 2013 Disponible en: <http://42jaiio.sadio.org.ar/proceedings/simposios/Trabajos/WSegI/03.pdf>.

[21] García P., van de Graaf J., Montejano G., Bast S., Testa O.: "Implementación de Canales Paralelos en un Protocolo Non Interactive Dining Cryptographers". 43° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO 2014), Workshop de Seguridad Informática (WSegI 2014). <http://sedici.unlp.edu.ar/handle/10915/42066>. 2014.

[22] P. García, J. van de Graaf, A. Hevia, A. Viola, "Beating the Birthday Paradox in Dining Cryptographers Networks". En "Progress in Cryptology – Latincrypt 2014". Springer International Publishing. ISSN: 0302-9743. ISSN (electronic): 1611-3349. ISBN: 978-3-319-16294-2. ISBN (eBook): 978-3-319-16295-9. Ps. 179 – 198. Octubre, 2014.

[23] P. García, G. Montejano, S. Bast, E. Fritz, "Codificación de Sufragios con Detección de Colisiones en NIDC con Canales Paralelos de Slots" Congreso Nacional de Ingeniería en Informática / Sistemas de Información. CoNaIISI 2016.

Criptología Maliciosa para la Ciberdefensa

Cipriano, Marcelo^{1,2}; García, Edith¹, Maiorano, Ariel¹
Malvacio, Eduardo¹,

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Facultad de Ingeniería del Ejército (FIE), Universidad de la Defensa Nacional - UNDEF

²Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes UNQ.

{marcelocipriano; egarcia; maiorano; emalvacio}@fie.undef.edu.ar

RESUMEN

En línea con el trabajo realizado al momento, plasmado en las publicaciones del grupo de investigación en el XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021) y en el XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022), el proyecto tiene por finalidad la continuidad del estudio, análisis y aplicación de paradigmas y herramientas criptológicas modernas para la creación de software malicioso y puertas traseras, como así también indagar técnicas de prevención, detección y protección para ser consideradas en el ámbito de la Ciberdefensa Nacional. Se planea la continuación del desarrollo de diferentes esquemas anti-kleptográficos para la experimentación y evaluación de su factibilidad y eficacia. En particular, se intentará abordar el estudio de otras funcionalidades criptográficas como por ejemplo: generación de números aleatorios, funciones de hashing, y algoritmos de cifrado y firma asimétrico. Aunque comúnmente se entiende a la criptografía y a sus aplicaciones como herramientas de carácter defensivo, también pueden emplearse para usos ofensivos y maliciosos; a saber, el secuestro, extorsión y pérdida de información producidos mediante software malicioso denominado *ransomware*, en sus distintas variantes. Por otra parte, la literatura también da cuenta de ataques en las etapas de diseño e implementación de algo-

ritmos criptográficos, comúnmente llamados *backdoors* o puertas traseras, que pueden vulnerar la confidencialidad, integridad y disponibilidad.

Es fácil observar las consecuencias directas sobre la ciberseguridad de usuarios particulares, empresas y organismos no gubernamentales de tales ataques. Pero un aspecto de este tipo de ataques pasa desapercibido y sin embargo tiene un impacto mayor. Pues amenaza directamente a la población de una ciudad, provincia o llanamente, un país al ser capaz de afectar a sus organismos públicos, fuerzas de seguridad, estructura militar, política y diplomática, como así también sus activos de información. Puede contribuir a la realización exitosa de ataques a las Infraestructuras Críticas, es decir, aquellas organizaciones relacionadas con la generación y distribución de energía, sistema financiero y bancario, organismos de salud como hospitales, servicio de potabilización y distribución de agua, saneamiento de desechos, entre otras. Es decir, los ataques basados en puertas traseras o *backdoors* podrían menoscabar la ciberdefensa de una nación.

Palabras Clave

Criptología, Criptografía Maliciosa, Puertas Traseras Criptográficas, Ciberdefensa.

CONTEXTO

“*Criptología Maliciosa para la Ciberdefensa*” (CRIPTO-MC) es un Proyecto de Desarrollo Tecnológico y Social (PDTS) aprobado por la Disposición Decanal Nro 667/2022, perteneciente a la Facultad de Ingeniería del Ejército (FIE) “Gral. Div. Manuel N. Savio”, perteneciente a la Universidad de la Defensa Nacional (UNDEF).

Se encuentra enmarcado en el contexto de la carrera de grado de Ingeniería en Informática, la Especialización en Criptografía y Seguridad Teleinformática y la Maestría en Ciberdefensa, que se dictan en la citada unidad académica.

Allí los investigadores conforman el Grupo de Investigación en Criptología y Seguridad Informática (GICSI), que conforma el Laboratorio de Informática, |Software Seguro y Criptografía (LISSyC), que lleva adelante tareas de I+D+i. El equipo está conformado por docentes investigadores categorizados en distintos regímenes científicos, profesionales técnicos, becarios y alumnos de la carrera de grado de Ingeniería en Informática, de la Especialización en Criptografía y Seguridad Teleinformática y de la Maestría en Ciberdefensa.

1. INTRODUCCIÓN

Es común apreciar a la criptografía y a sus aplicaciones como instrumentos de carácter defensivo que proporcionan confidencialidad en sistemas de comunicaciones, redes y bases de datos, entre otros. En el año 1996 en un trabajo fundacional, Adam Young y Moti Yung [1] presentan lo que han dado en llamar Crip-tovirología. Los autores previeron y mostraron la posibilidad de llevar adelante ataques mediante virus informáticos, que cifran la información de sus víctimas a través de criptografía de clave pública, pidiendo luego rescate para su recuperación. En la actualidad este tipo de malwa-

re se conoce como *ransomware*. Al año siguiente, los mismos autores presentan la llamada Kleptografía: esto es el diseño e implementación de *backdoors* o puertas traseras en algoritmos criptográficos [2-4]. En particular los autores presentan el mecanismo criptográfico *Secretly Embedded Trapdoor with Universal Protection*, conocido por sus siglas en inglés por el acrónimo de SETUP. Este kleptograma es una modificación a nivel matemático del algoritmo de intercambio de llaves Diffie-Hellman. Con las modificaciones pertinentes, estas técnicas kleptográficas se podrían implementar en el corazón de otros algoritmos, podrían ser embebidos en otros mecanismos criptográficos como son los esquemas de cifrado y de firma digital ElGamal, DSA, el algoritmo de firma de Schnorr, y el PKCS de Menezes-Vanstone y finalmente el reconocido algoritmo RSA [5-6, 8,13].

Es importante destacar que estos diseños no se limitan a la criptografía de llave pública. Se cuenta en la literatura con publicaciones que describen ejemplos aplicados a funciones de hash, en donde se presentan colisiones para una versión de SHA-1 modificando sus correspondientes parámetros [7], como también se presentan alternativas para protegerse de funciones de hash comprometidas en algoritmos de nivel superior, como HMAC y HKDF [14]. Por otro lado, tampoco los generadores de números de pseudo-aleatorios conocidos en la bibliografía como Pseudo Random Numbers Generators o PRNG por sus siglas en inglés. Estos algoritmos no serían inmunes a este tipo de ataques [10-13]. Básicamente la vulnerabilidad insertada afecta las propiedades estadísticas de un generador haciéndolo muy sensible a la entropía de la entrada. Por ejemplo, cuando los inputs tienen una distribución correcta, este mecanismo no tiene efecto, pero cuando están afectados por algún sesgo, el generador malicioso empeora considerablemente. En síntesis, la seguridad de los esquemas criptográficos se mide tradicionalmente como la incapacidad

de un adversario, que cuenta con recursos limitados, de violar un objetivo de seguridad deseado [14]. Sin embargo, este argumento de seguridad generalmente se basa en un diseño sólido de los componentes subyacentes. Podría decirse que uno de los fracasos más devastadores de este enfoque se puede observar al considerar adversarios con la capacidad de influir en el diseño, implementación y estandarización de primitivas criptográficas. Es entonces que considerando el impacto y la relevancia actual [16,17] de las técnicas y mecanismos mencionados se justificaría esta línea de investigación.

En la publicación realizada por el grupo en el *XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022)* [19], se presentó una muestra. Esta implementación se muestran a manera de prueba de concepto y aplicación experimental. De manera que permita la validación y análisis de los paradigmas y herramientas criptológicas modernas para la creación de software malicioso, y la elaboración de técnicas de prevención, detección temprana y protección, para ser consideradas como un aspecto más de la Ciberdefensa. El código fuente de referencia fue publicado en *Github.com*.

Específicamente, a través de un ejemplo de implementación, se presentó una técnica kleptográfica. Aunque específico, este primer ejemplo representa un caso posible de aplicación a un algoritmo en particular, presente en todas las ciphersuites aceptadas para TLS 1.2. Incluso, es la única definida como obligatoria. Aunque el escenario allí presentado, requiere e implica el compromiso inicial del sistema, el ataque es una puerta trasera o *backdoor* no detectable “*over the wire*”, ni siquiera monitoreando la red. Además persiste, aún frente a actualizaciones del software del Sistema Operativo, incluyendo aún, a aquellas que actualicen la librería OpenSSL, Es por estas y otras capacidades, que se puede apreciar la potencia y alcance de este tipo de malware.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto está conformado por distintas líneas de acción:

- Seguimiento de bibliografía y publicaciones científicas acerca de las aplicaciones de la tecnología kleptográfica y criptología maliciosa.
- Asistencia y/o seguimiento de cursos, jornadas, congresos y workshops de relevancia, tanto nacionales como internacionales.
- Análisis de primitivas matemáticas, el estudio de las puertas traseras kleptográficas y ataques conocidos.
- Estudio de las propiedades criptológicas específicas de la Keptografía aplicadas a los diferentes algoritmos criptográficos asimétricos y generación de números pseudo aleatorios (PRNG) entre otros.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se persiguen los siguientes objetivos y resultados:

- Estudiar y analizar la aplicación de los paradigmas y herramientas criptológicas modernas en la creación de software malicioso, como así también las técnicas de prevención, detección y protección para ser considerados en el ámbito de la Ciberdefensa Nacional.
- Estudiar y analizar las diferentes variantes de criptovirología y ataques kleptográficos existentes en la literatura, que son aplicados a diferentes algoritmos o primitivas criptográficas, con el objetivo de su detección y/o prevención

4. FORMACIÓN DE RECURSOS HUMANOS

Se espera que los resultados previstos puedan enriquecer el capital de conoci-

miento y recursos humanos del espacio académico civil y militar de la Facultad de Ingeniería del Ejército, de la Universidad de la Defensa Nacional y de las Fuerzas Armadas en general. Mediante la difusión y publicación de los resultados de la investigación, como así también ampliar la posibilidad de realizar jornadas de capacitación acerca de los mismos. Asimismo, se procura la difusión de estas temáticas en el ámbito de la Ciberdefensa Nacional, permitiendo dar a conocer nuestra maestría y logrando promocionar la misma. Evitar un ciberataque que vulnere las capacidades de defensa, afecte la confidencialidad de la información y hasta evite la afectación de una infraestructura crítica, en defensa de las vidas humanas involucradas.

Los investigadores que llevan adelante el proyecto dictan las asignaturas tanto en la carrera de grado como en la especialización antes mencionadas. Por ejemplo las asignaturas Matemática Discreta, Criptografía y Seguridad Teleinformática, Criptografía y Criptografía Avanzada. Desde esas cátedras se invita de forma permanente a los alumnos, de grado y posgrado, a para participar en los proyectos que se llevan adelante desde LISSyC.

En particular, los alumnos de posgrado, están llevando adelante su Trabajo Final Integrador, como así también su Tesis de Maestría. Se espera que la contribución mutua entre el equipo de investigadores, futuros especialistas y maestrandos permita alcanzar niveles sinérgicos de avance en la investigación, la formación de recursos humanos.

La Formación de Recursos Humanos permite incrementar el Know-How que tendrá el grupo de investigadores a lo largo de la vida del proyecto. Será un importante beneficio de sus integrantes y de la institución en la cual desarrollan sus actividades científico-docentes.

Por último y atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investiga-

ción, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

- [1] Young, Adam L. and Moti Yung. "Cryptovirology: extortion-based security threats and countermeasures." *Proceedings 1996 IEEE Symposium on Security and Privacy* (1996): 129-140.
- [2] Young, Adam L. and Moti Yung. "The Prevalence of Kleptographic Attacks on DiscreteLog Based Cryptosystems." *CRYPTO* (1997).
- [3] Young, Adam L. and Moti Yung. "Kleptography: Using Cryptography Against Cryptography." *EUROCRYPT* (1997).
- [4] Young, Adam L. and Moti Yung. "Malicious cryptography - exposing cryptovirology." (2004).
- [5] Young, Adam L. and Moti Yung. "A Space Efficient Backdoor in RSA and Its Applications." *Selected Areas in Cryptography* (2005).
- [6] Young, Adam L. and Moti Yung. "An Elliptic Curve Backdoor Algorithm for RSASSA." *Information Hiding* (2006).
- [7] Albertini, Ange, Jean-Philippe Aumasson, Maria Eichlseder, Florian Mendel and Martin Schl  ffer. "Malicious Hashing: Eve's Variant of SHA-1." *Selected Areas in Cryptography* (2014).
- [8] Young, Adam L. and Moti Yung. "Cryptography as an Attack Technology: Proving the RSA/Factoring Kleptographic Attack." *The New Codebreakers* (2015).
- [9] Russell, Alexander, Qiang Tang, Moti Yung and Hong-Sheng Zhou. "Cliptography: Clipping the Power of Kleptographic Attacks." *ASIACRYPT* (2015).
- [10] Indarjani, Santi. Sugeng, Kiki. Widjaja, Belawati. "Modification Attack Effects on PRNGs: Empirical Studies and Theoretical Proofs." (2015).
- [11] Young, Adam L. and Moti Yung. "Cryptovirology: the birth, neglect, and explosion of ransom-

ware" Commun. ACM 60 (2017): 24-26.

[12] Teseleanu, George. "Random Number Generators Can Be Fooled to Behave Badly." IACR Cryptology ePrint Archive (2018). | Página 16 de 16 [13] Markelova, A. V. "Vulnerability of RSA Algorithm." (2018). [14] Fischlin, Marc. Janson, Christian. Mazaheri, Sogol. "Backdoored Hash Functions: Immunizing HMAC and HKDF." (2018): 105-118. [15] Xiao, Dianyan and Yang Yu. "Klepto for Ring-LWE Encryption." Comput. J. 61 (2018): 1228-1239. [16] Yogi, Manas. Aparna, S.. "Novel insights into Cryptovirology A Comprehensive Study." International Journal of Computer Sciences and Engineering. 6. (2018): 1252-1255. [17] Zimba, Aaron. Chishimba, Mumbi. "On the Economic Impact of Crypto-ransomware Attacks: The State of the Art on Enterprise Systems." European Journal for Security Research. (2019) [18] Cipriano, Marcelo. García, Edith. Maiorano, Ariel. Malvacio, Eduardo. Pazo Robles, María Eugenia. "Criptografía maliciosa y ciberdefensa". XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021). Chilecito, La Rioja (2021). [19] Cipriano, Marcelo. García, Edith. Maiorano, Ariel. Malvacio, Eduardo. Pazo Robles, María Eugenia. "Ataque por Sustitución de Algoritmo Criptográfico: Implementación de prueba para OpenSSL". XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022). Ciudad de La Rioja, La Rioja (2022).

Métodos, Técnicas y Herramientas para la Protección de Sistemas de Software

José Pedro Montejano Massa, Mario Marcelo Berón, German Montejano, Daniel Riesco

Universidad Nacional de San Luis

Facultad de Ciencias Físico Matemáticas y Naturales

Área de Programación y Metodologías de Desarrollo de Software

mberon@unsl.edu.ar, {jose.p.montejano, german.a.montejano, riesco.daniel}@gmail.com

Resumen

Se presenta una línea de investigación y desarrollo que se basa en la seguridad informática y en la protección de los sistemas informáticos. La seguridad informática se ha vuelto cada vez más importante en los últimos años debido a la creciente cantidad de ataques cibernéticos y vulnerabilidades de seguridad en las aplicaciones de software.

Una de las principales amenazas a la seguridad de una aplicación de software es la explotación de vulnerabilidades de seguridad en el código fuente de la aplicación. Estas vulnerabilidades pueden ser explotadas por los atacantes para obtener acceso no autorizado al sistema, robar información confidencial, o incluso causar daños irreparables al sistema.

Es fundamental que los desarrolladores de software presten especial atención a la seguridad de sus aplicaciones y tomen medidas para proteger el código fuente de las mismas. Esto da lugar a una necesidad y es la que se pretende cubrir en esta línea de investigación. Se pretende encontrar soluciones informáticas que deriven en la implementación de técnicas y herramientas de protección de código para ser brindadas al desarrollador como un instrumento que puedan sumar en el desarrollo de sus proyectos de software para aumentar la seguridad de los mismos.

Palabras clave

Seguridad informática – Vulnerabilidad – Protección de código fuente.

Contexto

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de Software (LaCIS) de la Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto: “Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”, perteneciente a la misma. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional.

1. Introducción

El apogeo del software es, sin lugar a dudas, cada día más tangible y evidente. Es posible encontrar aplicaciones de software en: una casa, en la calle, en los campos, en plataformas marítimas y hasta en el espacio exterior. Es decir, el software no solo se encuentra inmerso en cada aspecto del planeta tierra, sino que es absolutamente necesario para el día a día.

El uso del software soluciona gran parte de los problemas de la vida cotidiana, partiendo de la base de que el usuario lleva siempre un celular, el cual se usa para cosas tan simples como chequear el clima, la hora, o también para cosas tan complejas como hacer transferencias bancarias. El uso del software soluciona necesidades a gran escala que son absolutamente necesarios para el correcto funcionamiento de la población, como sistemas de gestión de hospitales, sistemas bancarios, sistemas de aeronáuticas, sistemas embebidos en maquinaria industrial, etc.

Como es posible observar, el software es fundamental para el bienestar social con lo cual es altamente requerido por la comunidad. Por esta razón, en la actualidad, se ha producido una alta demanda de profesionales de la informática y asociado con ello aumentó masivamente la producción de software. Estos nuevos profesionales deben ser capaces de desarrollar software de calidad. Esta relevante característica, conduce a que haya un conocimiento absoluto del uso de la información que pasa por dicho software; entiéndase por información tanto a la interacción de un usuario con un juego de celular como a el almacenamiento y tratado de cuentas bancarias, información de salud personal, etc. Por ende, los sistemas desarrollados deben ser seguros.

Lograr desarrollar software seguro no es tan simple, dado el crisol de metodologías de desarrollo de software y estilos de programación diferentes que existen [1]. Muchas veces, un uso indebido de una metodología, o el poco soporte que proporciona para el desarrollo de software seguro, puede llevar a no garantizar la integridad y confidencialidad tanto de la información como de los datos que se utilizan en un sistema [2].

Por supuesto que ningún software es inviolable y no se puede garantizar la protección total de

un sistema, pero si es fundamental ofrecer una base de protección del código fuente para prevenir un ataque indebido o un acceso no autorizado. Ya no es una tarea aislada desarrollar software y pensar en la seguridad del mismo, es decir estos dos conceptos van intrínsecamente conectados. Es tan relevante la aseveración antes mencionada que muchas empresas de software han adaptado sus metodologías de desarrollo de software para producir software seguro, como es el caso de Microsoft con SDL (Security Development Lifecycle) [3], Oracle con OSSA (Oracle Software Security Assurance) [4], entre otras. A partir de lo antes mencionado, es importante recalcar que no debe lanzarse el despliegue de un sistema o módulo de software sin poder garantizar su seguridad.

En esta línea de investigación y desarrollo se está abordando la temática de la seguridad de los sistemas informáticos a través del desarrollo de herramientas y técnicas que permitan la protección del código fuente de un programa.

Se considera que la investigación en el campo de la seguridad informática es de vital importancia en la actualidad, y se espera que este trabajo pueda contribuir a mejorar la protección de los sistemas informáticos y prevenir los ataques cibernéticos y las vulnerabilidades de seguridad en las aplicaciones de software.

2. Líneas de investigación y desarrollo

Las vulnerabilidades en informática [5] son, hoy por hoy, un eje central a tener en cuenta en el desarrollo de cualquier proyecto de software. La norma ISO/IEC 27005:2008 [6] define una vulnerabilidad como “Una debilidad de un activo o grupo de activos que pueden ser explotadas por una o más

amenazas, donde un activo es cualquier cosa que tiene valor para la organización, sus operaciones comerciales y su continuidad, incluidos los recursos informáticos que respaldan la misión de la organización”. En cualquier proyecto uno de los activos [7] principales es el código fuente del software en cuestión, esto da paso al eje central de esta línea de investigación que se basa en la protección de dicho código fuente.

Siguiendo con esto, se llega a la deducción de que sería de gran valor, para cualquiera de las partes interesadas en un proyecto de software, lograr proteger el código fuente [8]. Pero es necesario ahondar un poco más y ver cuáles serían posibles beneficios de lograr esto. La protección del código de un programa puede ofrecer varios beneficios, entre ellos:

- **Protección de la propiedad intelectual:** Al proteger el código fuente, se protege la propiedad intelectual del creador o propietario del programa, lo que puede impedir la copia o uso no autorizado del software [9]. Esto puede ser especialmente importante para los desarrolladores que buscan monetizar su trabajo a través de la venta de licencias de software [10].

- **Seguridad:** La protección del código fuente puede ayudar a proteger contra ataques malintencionados [11]. Si los detalles internos del código fuente son de conocimiento público, los atacantes pueden explorar el código en busca de vulnerabilidades y usar esta información para comprometer el sistema o el software en sí. Al proteger el código fuente, se hace más difícil para los atacantes identificar las debilidades y crear ataques [12].

- **Ventaja competitiva:** La protección del código fuente puede otorgar una ventaja competitiva a una empresa o desarrollador de software. Si el código fuente es inaccesible, otras empresas o desarrolladores no podrán estudiarlo para replicar la funcionalidad del

software, lo que puede ayudar a mantener una ventaja en el mercado [13].

Ahora bien, se sabe que es imposible garantizar la seguridad de un programa al cien por ciento, entonces el objetivo es dificultar la tarea de un usuario que intenta robar o manipular código fuente lo máximo posible para que fracase en su intento. Y para esto se pueden utilizar de base innovadoras estrategias que ya han sido desarrolladas como Ofuscación de Código, Marca de Agua, Marca de Nacimiento, Protección de Software por Hardware, entre otras tantas aproximaciones orientadas a proteger los activos contenidos en los sistemas. Cada estrategia tiene su finalidad, por ejemplo, la ofuscación intenta confundir al atacante de forma tal que no pueda identificar parte del sistema que desea vulnerar [14]. La marca de agua es una técnica usada para prevenir los ataques a través de la inserción identificadores confiables que representan al propietario del sistema. Si un atacante modifica el software, el propietario puede extraer el identificador para verificar si es el software original o el mismo ha sido modificado por el atacante [15]. Otras técnicas como la marca de nacimiento se basan en encontrar propiedades innatas del software el cual cuando es modificado cambia la marca indefectiblemente [16]. Finalmente, las estrategias de protección basadas en hardware se basan en la elaboración de dispositivos electrónicos que son necesarios para que el sistema funcione [17].

Según esta línea de investigación, las herramientas que se han encontrado en la literatura implementan sofisticadas técnicas de protección, y la finalidad principal de ellas es la de probar el nivel de fortaleza de la estrategia en sí [18]. Mientras que el enfoque de esta línea de investigación se centra en encontrar una solución más general la cual permita aplicar primitivas de protección a

diferentes lenguajes y paradigmas de programación.

3. Resultados obtenidos/esperados

A través de los trabajos realizados por los integrantes de esta línea de investigación se han podido obtener diferentes resultados. Se desarrolló una aplicación web que permite proteger sistemas escritos en Java a través del uso de técnicas de protección de software que pueden ser desarrolladas por el mismo usuario de la herramienta según su necesidad particular.

Esta aproximación que permite la aplicación automatizada de técnicas de protección al código fuente de un sistema está dirigida a profesionales de la informática y cuenta con dos funcionalidades principales: la primera de ellas permite que las técnicas de protección sean subidas a la herramienta por los usuarios, mientras que la segunda posibilita la aplicación de estas técnicas al sistema que se desea proteger.

La aplicación es fácil de usar y permite proteger el código fuente de las aplicaciones Java de una manera simple y efectiva. También cuenta con algunas técnicas de protección ya implementadas que están listas para usar, las mismas fueron analizadas y escogidas para que sirvan de la mejor manera como ejemplo de caso práctico. Entre estas técnicas de protección implementadas se encuentra una técnica de marca de agua y una de señuelo.

Se espera desarrollar nuevas estrategias de protección que involucren la elaboración de métodos para detectar las partes vulnerables de un sistema, elaboración de estrategias de reingeniería que dado un sistema detecte la parte vulnerable y lo proteja. Estas soluciones pueden ser utilizadas para detectar patrones en los datos y comportamientos sospechosos en

los sistemas, lo que permitiría una respuesta más rápida y eficiente a las amenazas de seguridad.

4. Formación de recursos humanos

Las tareas realizadas en el contexto de la presente línea de investigación están siendo desarrolladas como parte de un Trabajo Final Integrador para optar al grado de Ingeniería en Informática. En el futuro se piensa generar diferentes tesis de maestría y doctorado usando como base parte de esta investigación. Se busca, entre otras cosas, escalar la implementación de la herramienta para que no solo proteja sistemas escritos en Java, sino que también proteja sistemas escritos en varios lenguajes.

5. Bibliografía

- [1] Grembi, J., & Chandra, P. (2015). Secure Software Development: A Security Programmer's Guide. Syngress.
- [2] Waters, K. (2012). The Dark Side of Agile Software Development. Routledge.
- [3] Blokdyk, G. (2021b). Microsoft Security Development Lifecycle a Complete Guide - 2019 Edition. 5STARCook.
- [4] Oracle Software Security Assurance. Oracle Corporation. (2023, 7 marzo). <https://www.oracle.com/security/assurance/>
- [5] Thompson, H. H., & Chase, S. G. (2007). The Software Vulnerability Guide. Laxmi Publications Pvt Limited.
- [6] ISO/IEC, "Information technology -- Security techniques-Information security risk management" ISO/IEC FIDIS 27005:2008.
- [7] Holsing, N. F., & Yen, D. C. (1999). Software Asset Management. Information Resources Management Journal, 12(3), 14-26. <https://doi.org/10.4018/irmj.1999070102>

- [8] J. Yang, C. Barrientes, J. Sanchez and Y. R. Kim, "Source Code Analysis for Secure Programming Practices," 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2018, pp. 819-824, doi: 10.1109/CSCI46756.2018.00164.
- [9] Publications, LandMark. (2019). Software Copyright (English Edition). LandMark Publications.
- [10] Información legislativa. (2023, 7 marzo). servicios.infoleg.gob.ar.
<http://servicios.infoleg.gob.ar/infolegInternet/anexos/40000-44999/42755/texact.htm>
- [11] Ransome, J., & Misra, A. (2019). Cybersecurity for Software Engineers. Boca Raton, FL: CRC Press.
- [12] Nagra, J. & Collberg, C. (2009b). Surreptitious Software: Obfuscation, Watermarking, and Tamperproofing for Software Protection (English Edition) (1.a ed.). Addison-Wesley Professional.
- [13] Michael Porter. (1985). Competitive Advantage: Creating and Sustaining Superior Performance. Free Press.
- [14] Brunton, F. & Nissenbaum, H. (2016). Obfuscation: A User's Guide for Privacy and Protest. The MIT Press.
- [15] Arnold, M., Wolthusen, S. D. & Schmucker, M. (2003). Techniques and Applications of Digital Watermarking and Content Protection (Ilustrado.). Artech House Publishers.
- [16] Park, H., Lim, H. I., Choi, S. & Han, T. (2008). A Static Java Birthmark Based on Operand Stack Behaviors. 2008 International Conference on Information Security and Assurance (isa 2008).
<https://doi.org/10.1109/isa.2008.15>
- [17] Ju, H., Jeon, Y. & Kim, J. (2015). A Study on the Hardware-Based Security Solutions for Smart Devices. 2015 International Conference on Computational Science and Computational Intelligence (CSCI).
<https://doi.org/10.1109/csci.2015.105>
- [18] Wagner, D., & Jones, A. K. (2010). Software Protection and Simulation on Code Virtualization. Springer.

La blockchain y la certificación de documentos en Argentina

Carolina Canessa^{1,2,3}, Elisa Galdame¹, Marcelo Palma¹, y Luis Canessa^{1,3}

¹ Universidad Champagnat, Facultad de Informática y Diseño,
Mendoza, Argentina.

{canessacarolina, elisagaldame, mpalma, canessaluis}@uch.edu.ar

² Universidad Tecnológica Nacional, Facultad Regional Mendoza, Mendoza,
Argentina.

³ Universidad de Mendoza, Facultad de Ingeniería, Mendoza, Argentina.

RESUMEN

La tecnología blockchain es adecuada para escenarios en los que se requiera almacenar de forma creciente datos ordenados en el tiempo, sin posibilidad de modificación ni revisión y cuya confianza pretenda ser distribuida en lugar de residir en una entidad certificadora. Los potenciales beneficios de la tecnología blockchain son suficientes para alcanzar una adopción generalizada en los próximos años en la certificación académica de documentos privados en cuanto que ha demostrado cumplir con mayor seguridad, registros inmutables, reducción de costos y tiempo en transacciones.

Palabras Claves:

Blockchain, Bitcoin, Criptomonedas.
Certificación. Crypto

CONTEXTO

El presente proyecto se desarrolla en el marco de la Facultad de Informática y Diseño. Este trabajo es parte de las líneas de investigación y de trabajo llevadas adelante por Carolina Canessa y Luis Canessa desde el año 2019.

1. INTRODUCCIÓN

El nacimiento de Bitcoin, en el año 2009, puso en evidencia la existencia de una nueva tecnología denominada Blockchain, que posibilitaba pasar del actual internet de la información al internet del valor. En menos de 10 años de existencia, esta nueva tecnología disruptiva está creando a su alrededor todo un nuevo ecosistema que va mucho más allá de Bitcoin y su uso original como mera criptomoneda. En torno a la Blockchain se está construyendo un nuevo modelo económico que

se conoce como cripto economía caracterizado por la descentralización y porque puede transformar radicalmente muchas de las estructuras económicas y sociales actuales. El Blockchain es una tecnología basada en una cadena de bloques de operaciones descentralizada y pública. Esta tecnología genera una base de datos compartida a la que tienen acceso sus participantes, los cuáles pueden rastrear cada transacción que hayan realizado. Es como un gran libro de contabilidad inmodificable y compartido que van escribiendo una gran cantidad de ordenadores de forma simultánea

El cambio tecnológico de las últimas décadas permite observar la íntima relación que existe entre los archivos digitales, su disponibilidad, la protección de los patrimonios / memorias y el desarrollo de inteligencias público-comunitarias. Por ello, cada día es más evidente que los archivos valen por lo que preservan y, sobre todo, por los usos futuros que de ellos se construyan.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo general de este proyecto es: explorar, reconocer e interpretar las implementaciones jurídicas existentes al presente que operan técnicamente en la blockchain con el objeto de transferir el conocimiento generado a las unidades académicas correspondientes, tendientes a formar futuros profesionales que sean capaces de resolver problemáticas en torno a su utilización y a generar en las mismas, nuevos conocimiento científico y tecnológico y

transferirlo a la comunidad.

Para el desarrollo del presente proyecto pueden diferenciarse 4 etapas principales.

1. Conocer el mundo de las criptomonedas, su historia, su funcionamiento y su proyección a futuro.
2. Reseñar brevemente los antecedentes históricos a nivel nacional e internacional.
3. Describir, a nivel macro, el funcionamiento tecnológico del sistema de criptomoneda más conocido: Bitcoin,
4. Exponer en detalle los conceptos de la tecnología blockchain y hacer un estudio de su impacto en la sociedad y sus perspectivas de futuro como posibles aplicaciones diseñadas para la certificación legal de documentos privados.

3. RESULTADOS ESPERADOS

El alcance del estudio contempla investigar el entorno de la tecnología Blockchain, su valor estratégico y potenciales aplicaciones de la tecnología. Entender cuáles son las industrias en las que se presentará un mayor impacto y cómo afectará a los modelos de negocio, de esta manera entender sus beneficios y áreas en las que se puede innovar utilizando esta tecnología como herramienta. Analizar los desafíos y oportunidades jurídicas de la tecnología Blockchain para brindar seguridad en la certificación documentos privados.

4. FORMACIÓN DE RECURSOS HUMANOS

Para la realización de esta investigación, contamos con un equipo interdisciplinario de profesionales nacionales e internacionales como así de alumnos becarios apasionados con este tema de tanta actualidad por lo cual se espera formar recursos humanos con conocimientos fundamentales para afrontar los distintos escenarios que esta tecnología nos propone.

El equipo de trabajo de este proyecto está integrado por:

1. Esp. Ing. Carolina Canessa directora de este proyecto, docente e investigadora de la Universidad Champagnat, Universidad de Mendoza y Universidad Tecnológica Nacional. Se especializa en técnicas de Investigación Operativa y Modelado de Sistemas Estocásticos. Además, cuenta con experiencia práctica en análisis técnico de criptomonedas.

2. Esp. Lic. Elisa, docente coordinador en la UMAZA y Secretraria Técnica de Evaluación y Acreditación de Carreras, de la misma Universidad. Docente Investigador en la UCH. Y profesora de Posgrados en la UMAZA.

3. Lic. Marcelo Palma, docente e investigador de Universidad Champagnat. Especialista en Ingeniería de Software Cuenta con experiencia práctica en el ciclo de vida completo del desarrollo de software, en particular en Arquitectura y programación.

4. Luis Canessa, fundador y primer Director del Nodo de Informática y Comunicaciones de la Universidad Nacional de Cuyo, logrando el primer enlace de Internet entre dos países Latino Americanos. Formó parte del grupo de gerentes que implementó la Ley Bitcoin en El Salvador fundando la Asociación “Chivo Maker”. Actualmente, es Managing Partner en Sophia Projects y en la Asociación Chivo Maker dedicado a la implementación de sistemas cripto para gobiernos y grandes compañías en el mercado de LATAM.

En el campo de la formación de recursos humanos, se espera:

1. Iniciar la investigación a alumnos de la UCH que deseen sumarse al proyecto como becarios.
2. Iniciar en investigación a alumnos de los últimos años de carrera que deseen desarrollar su proyecto final de carrera en el marco de este proyecto.

5. BIBLIOGRAFÍA

Andreas Antonopoulos (2019). Mastering Bitcoin: Traducción al español de la guía

completa del mundo de bitcoin y blockchain editorial: PROFIT I.S.B.N: 9788417209728.

Benito Arruñada (2018). Limitaciones de blockchain en contratos y propiedad. Traducción publicada en la Revista Crítica de Derecho Inmobiliario, vol. 94, núm. 769, 2018, pp. 2465-93. Disponible en <https://www.arrunada.org/files/research/ARRU%C3%91ADA%202019%20Blockchain%20RCDI.pdf>

Cristhian Jesús Mejía Fernández (2021). Modelo Blockchain para el aseguramiento de la Información en Certificados de Notas SSA-UMSA (Sistema de Seguimiento Académico - UMSA) Postgrado en Informática Universidad Mayor de San Andrés La Paz - Bolivia. Disponible en https://ojs.umsa.bo/ojs/index.php/inf_fcpn_pg/article/view/47

Franz Gustavo Laime Rojas (2021). La criptoconomía y tecnología blockchain (la cuarta revolución industrial) retos y oportunidades de futuro. Máster en Comercio y Finanzas Internacionales. Universitat de Barcelona.

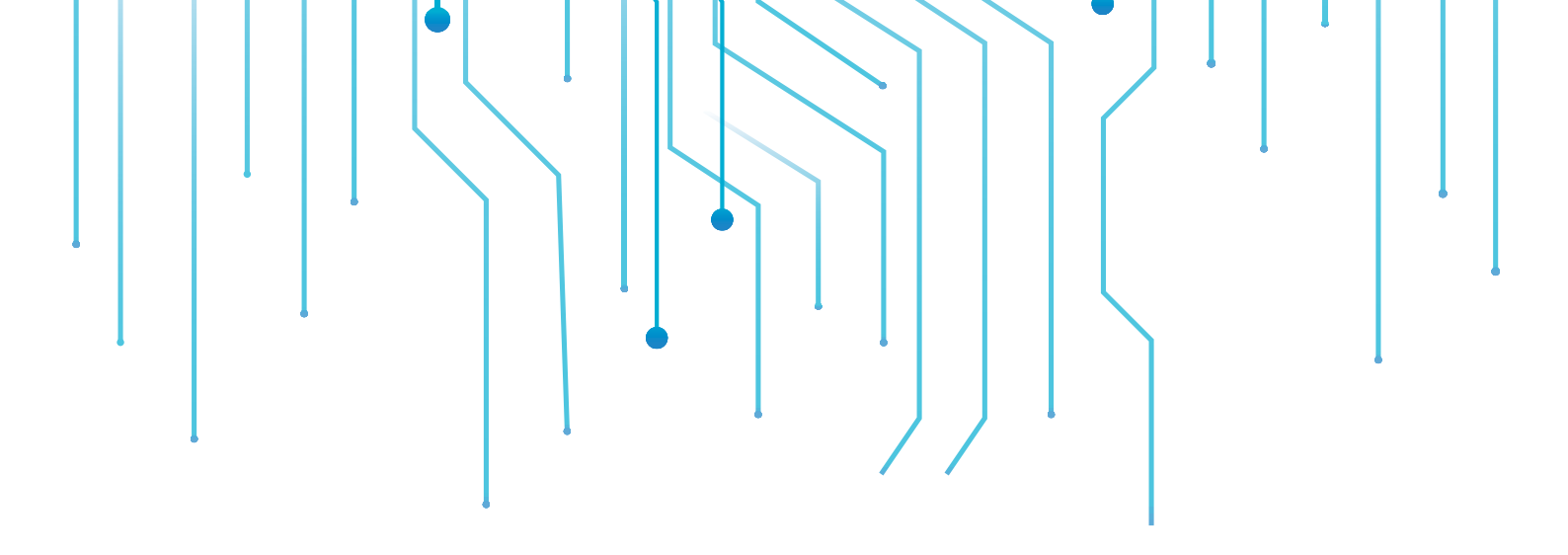
Kevin Leonel Young (2021). Incidencias de la blockchain sobre la autoría e integridad de los documentos electrónicos desde una doble perspectiva española-argentina. Trabajo final de máster. Tarragona 2021. Disponible en https://repositori.urv.cat/estatic/TFM0011/en_TFM810.html

Loiacono, S. (2021). Blockchain, sus aplicaciones más allá de las criptomonedas. Revista Abierta De Informática Aplicada, 2, 49–52. Recuperado a partir de <http://portalrevisien.uai.edu.ar/ojs/index.php/RAIA/article/view/69>

Mora. Santiago J. (2019). La tecnología blockchain Contratos inteligentes, ofertas iniciales de monedas y demás casos de uso. Tomo La Ley 2019-B, AÑO 1, XXXIII N°62. Disponible en <https://www.udes.edu.ar/sites/default/files/1693.pdf>

Moreno, Santiago (2021). Introducción al Blockchain y Criptomonedas. Editorial: Nowtilus. Formato:Rústica. Edición: 2021. ISBN:9788413051741

Vercelli, A. (2020). Los archivos de la informática argentina y el uso de blockchain: aspectos jurídicos y tecnológicos a considerar. Simposio de Informática y Derecho, 49JAIIO, pps.: 131 – 141. Disponible en <https://49jaiio.sadio.org.ar/pdfs/sid/SID-11.pdf>



TD

Tesis Doctorales

Generalización del modelado de cadencias de tecleo con contextos finitos para su utilización en ataques de presentación y canal lateral

**Doctorado en Ciencias Informáticas
Universidad Nacional de La Plata (UNLP)**

Autor: Nahuel González¹

Fecha de Defensa: 07/06/2022

Director: Dr. Jorge S. Ierache¹

Codirector: Dr. Waldo Hasperué²

Asesor científico: Dr. Enrique P. Calot¹

¹ Laboratorio de Sistemas de Información Avanzados (LSIA)

Facultad de Ingeniería, Universidad de Buenos Aires

Ciudad Autónoma de Buenos Aires, Argentina

² Instituto de Investigación en Informática (III-LIDI)

Facultad de Informática, Universidad Nacional de La Plata

La Plata, Provincia de Buenos Aires, Argentina

Resumen

Los patrones individuales de escritura en un teclado, denominados cadencias de tecleo, configuran un atributo biométrico comportamental que puede utilizarse como segundo factor de autenticación (2FA) para verificar en forma transparente la identidad de los usuarios. Los sistemas de autenticación basados en cadencias de tecleo son vulnerables a ataques de presentación, que utilizan muestras sintetizadas para imitar el comportamiento del usuario legítimo; adicionalmente, cuando otros sistemas filtran los tiempos de escritura, estos pueden ser recolectados en un ataque por canal lateral para luego identificar el texto ingresado o potenciar un ulterior ataque por fuerza bruta, reduciendo los candidatos posibles. A pesar de estos riesgos, los métodos de autenticación basados en cadencias de tecleo que se han propuesto en la literatura han sido generalmente evaluados bajo un modelo de esfuerzo cero, también llamado de impostores no entrenados, que subestima o ignora el riesgo de los ataques mencionados. En esta tesis se propone un sistema de detección de vida para mitigar los riesgos de un ataque de presentación y una familia de estrategias de síntesis de muestras artificiales que son utilizadas como adversarios durante el entrenamiento. También se proponen nuevas distancias basadas en los histogramas empíricos del perfil del usuario legítimo, que presentan mejor capacidad de discriminación que las distancias clásicas. El modelo de detección de vida logra tasas de falsos positivos menores al 2% contra muestras sintetizadas con perfiles interusuario y menores al 15% incluso en el caso extremo en que el atacante cuenta con el perfil biométrico completo del usuario legítimo; estos valores deben contrastarse con el 98%+ de falsos positivos alcanzado sin detección de vida. Adicionalmente, una modificación del modelo propuesto permite abordar el problema de identificación del texto ingresado utilizando sólo los tiempos de escritura, potenciando los ataques de canal lateral y de fuerza bruta ya existentes.

Palabras clave: seguridad informática, biometría comportamental, cadencias de tecleo, detección de vida, ataques de presentación, ataques por canal lateral.

1. Introducción

Las sutiles variaciones en la forma en que distintas personas teclean son suficientes para revelar su identidad. Hace cuarenta años, Gaines y colaboradores (Gaines 1980), pioneros del análisis de cadencias de tecleo, reconocieron la utilidad de este fenómeno para la autenticación de usuarios. Utilizando sólo los tiempos entre eventos de presión y liberación de teclas, es posible construir un segundo factor de autenticación (2FA) tanto para endurecer claves de usuario (Monrose 1990) como para verificar la identidad en forma continua (Bours 2009). Más recientemente, el análisis de cadencias de tecleo también ha encontrado usos fuera del dominio de la seguridad informática, como descubrir ciertas características fisiológicas o impedimentos clínicos del usuario (Milne 2018), e incluso determinar en forma aproximada las variaciones de su estado emocional mientras escribe (Epp 2011).

El desafío que hoy enfrenta el análisis de cadencias de tecleo consiste en prevenir modelos de ataque sofisticados (Rahman 2011) en donde el impostor malintencionado cuenta con acceso parcial o total al perfil biométrico del usuario y realiza un esfuerzo considerable para imitar al usuario legítimo (Gonzalez2021b). O, también, en explorar el uso de las cadencias de tecleo para potenciar otro tipo de ataques como los de canal lateral (Gonzalez 2021d) o fuerza bruta (Song 2001). Dentro de esta línea, Monaco y colaboradores (Monaco 2019) han demostrado que es posible reconstruir la consulta escrita en un motor de búsqueda empleando los intervalos entre teclas sucesivas, inferidos en base a los tiempos de arribo de paquetes de red.

1.1 Contexto

El análisis de cadencias de tecleo pertenece al ámbito de la *seguridad informática* ya que permite, como único factor o en combinación con otros, la autenticación y verificación de identidad. Los métodos de autenticación pueden dividirse en aquellos basados en el conocimiento, basados en la posesión, y *biométricos*; el tema que nos compete pertenece a estos últimos, que se dividen en convencionales, cuando estudian

las características fisiológicas del usuario, y *comportamentales* cuando analizan sus patrones de comportamiento. El análisis de cadencias de tecleo también pertenece al ámbito de la *interacción hombre-máquina* ya que la fuente de información es un dispositivo de entrada: el teclado, o a veces también un dispositivo móvil.

Dentro de la *seguridad informática* se estudian los tipos de *ataque* a los que puede ser sometido un sistema de información. Estos pueden ser *directos* o *indirectos*. Los primeros son aquellos que demandan una interacción con el sistema; los segundos, aquellos que se realizan con información filtrada sin interactuar con el sistema.

1.2 Definiciones

Un *ataque de presentación* es un tipo de ataque directo, en el cuál un actor malintencionado muestra o presenta al sistema de autenticación las credenciales biométricas del usuario legítimo, ya sea copiándolas o imitándolas, con la intención de hacerse pasar por él y lograr obtener acceso privilegiado. La prevención de ataques de presentación requiere robustecer al sistema considerado con un modelo de *detección de vida*, capaz de distinguir entre el comportamiento del usuario legítimo y el de un atacante que intenta imitarlo. En particular, los sistemas de autenticación basados en cadencias de tecleo son vulnerables a ataques de presentación con *muestras sintéticas*, es decir aquellas construidas artificialmente por el atacante (en contraste con las muestras legítimas, escritas naturalmente por el usuario) con la ayuda de una herramienta de síntesis y, posiblemente, conocimiento parcial o total del comportamiento del usuario legítimo o de su perfil biométrico.

Un *ataque por canal lateral* es todo aquel que explota la información extrínseca filtrada inadvertidamente por una cierta implementación particular de un sistema. Es un tipo de ataque indirecto, y se caracteriza por tener como objetivo dicha implementación particular en lugar del algoritmo, método, o sistema implementado. Por ejemplo, en (Song 2001) se detalla cómo una implementación del protocolo SSH que transfiera por red y sin demora los caracteres escritos filtra

inadvertidamente los tiempos de escritura y cómo estos pueden emplearse para potenciar un ataque por fuerza bruta, reduciendo en dos órdenes de magnitud el tiempo y la cantidad de intentos requeridos para adivinar la clave del usuario.

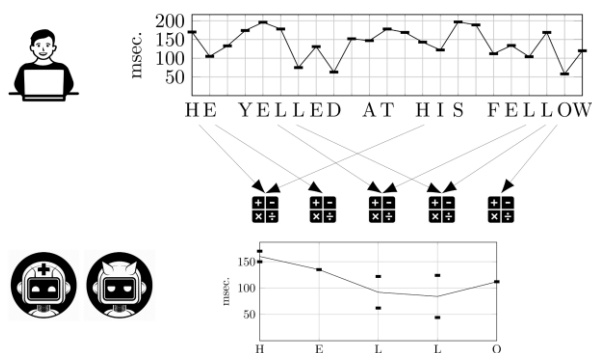


Figura 1 - Modelado por contextos finitos

El *modelado por contextos finitos*, que se ilustra en la Figura 1, es un método general de síntesis de cadencias de tecleo que emplea un conjunto de muestras de escritura en texto libre del usuario legítimo para reconstruir con fidelidad los atributos temporales, y quizás otros, que corresponden a cualquier secuencia de caracteres, así no existan observaciones previas de esta última entre las muestras de entrenamiento (González 2015b). Los métodos de síntesis de muestras propuestos en esta tesis (ver Sección 4) se basan en el modelado por contextos finitos.

Un *perfil intrausuario* es aquel que sólo incluye muestras del usuario a imitar. En contraste, un *perfil interusuario* es aquel que se compone de muestras de varios usuarios de una población genérica.

1.3 Motivación

Las técnicas de análisis y modelado de cadencias de tecleo han sido y continúan siendo ampliamente estudiadas por su interés intrínseco y para su aplicación como un segundo factor en la autenticación de usuarios. En particular, estos métodos tienen la ventaja de ser transparentes, en el sentido de que no demandan ninguna acción ulterior al usuario legítimo más que escribir naturalmente, y no le imponen demoras, confirmaciones, u otras

tareas como los biométricos convencionales y demás esquemas de autenticación multifactor.

Sin embargo, las vulnerabilidades de los sistemas de autenticación basados en cadencias de tecleo ante ataques de presentación y falsificaciones sintéticas han recibido mucha menos atención que en el caso de los biométricos tradicionales, a pesar de que las debilidades a las que los expone no es inferior. Hoy en día y en vista de la ubicuidad de las filtraciones de datos, no es descabellado imaginar que un atacante cuenta con acceso, al menos parcial, a las muestras biométricas del usuario. Por este motivo es necesario investigar métodos sofisticados de ataque y las contramedidas de defensa necesarias para mitigarlos. Asimismo, el análisis y modelado de cadencias de tecleo es promisorio como vector de amplificación de otro tipo de ataques, de canal lateral, al permitir la filtración de información sobre el texto escrito en la forma de valores temporales correlacionados con este.

1.4 Marco

La presente tesis fue radicada en el Laboratorio de Sistemas de Información Avanzados (LSIA) de la Facultad de Ingeniería de la Universidad de Buenos Aires, y se enmarca en el proyecto de desarrollo estratégico PDE-44-2019, *Reconocimiento de Patrones de Tecleo en Ambientes Web*,

2. Definición del problema

El problema que ha guiado los experimentos de esta tesis es la creación de contramedidas de defensa eficaces ante ataques de presentación que utilicen muestras sintetizadas artificialmente, en base a un perfil intrausuario total o parcial. En el camino hacia la solución se estudiaron las distribuciones temporales subyacentes y se propusieron estrategias de síntesis de muestras capaces de engañar a los actuales sistemas de detección de vida con mayor frecuencia que los empleados usualmente para la evaluación. Por último, se descubrió que una modificación de los mismos métodos permite potenciar los ataques por canal lateral para identificación del texto ingresado.

2.1 Objetivos

Identificar las distribuciones subyacentes y los patrones de comportamiento que generan la cadencia de tecleo en texto libre.

Proponer estrategias de síntesis de muestras artificiales para su uso en ataques de presentación, que engañen a los actuales sistemas de autenticación con frecuencia suficiente como para constituir una amenaza.

Construir un sistema de detección de vida que sirva como contramedida de defensa ante las anteriores y otras estrategias de síntesis de muestras artificiales del estado del arte.

Explorar las técnicas anteriores para su uso en ataques de canal lateral, incrementando la capacidad de identificar un texto en base a los tiempos de escritura.

3 - Contribuciones

La presente tesis contribuye a incrementar la seguridad de los métodos de autenticación por medio de cadencias de tecleo, robusteciéndolos frente a ataques de presentación con muestras sintetizadas, a la vez que potencia los ataques por canal lateral a través la reconstrucción del texto ingresado en base a los tiempos de escritura.

3.1 - Métodos

Los métodos propuestos se describen en detalle en la Sección 4.

3.2 - Herramienta

Se creó una herramienta integrada que implementa los métodos propuestos en esta tesis. La herramienta y el código fuente fueron puestas a disposición del público en el repositorio del Laboratorio de Sistemas de Información Avanzados y en la publicación con revisión de pares (Gonzalez 2023a), en *Software Impacts*, Elsevier, que acompaña al artículo (Gonzalez 2021b).

3.3 - Producción científica

La producción científica derivada de esta tesis incluye dos publicaciones en revistas internacionales con revisión de pares: *On the Shape of Timing Distributions in Free-Text Keystroke Dynamics Profiles* (Gonzalez 2021a) en *Heliyon*, Elsevier, revista de primer

cuartil (Q1) según Scimago Journal Ranking; y *Towards Liveness Detection in Keystroke Dynamics: Revealing Synthetic Forgeries* (Gonzalez 2021b) en *Systems and Soft Computing*, que acompaña a *Applied Soft Computing*, revista de primer cuartil (Q1) según Scimago Journal Ranking. Dentro del mismo tópico y como resultado de la investigación, dos artículos adicionales en revistas internacionales con revisión de pares fueron publicados posteriormente a la fecha de aprobación de la tesis: *KSDSLD—A tool for keystroke dynamics synthesis & liveness detection* (Gonzalez 2023a) en *Software Impacts*, Elsevier, y *Dataset of Human-written and Synthesized Samples of Keystroke Dynamics Features for Free-text Input* (Gonzalez 2023b) en *Data in Brief*, Elsevier. La publicación *Exploring Internal Correlations in Timing Features of Keystroke Dynamics at Word Boundaries and Their Usage for Authentication and Identification* (Gonzalez 2020b) en *Computer Science—CACIC 2020: Revised & Selected Papers* pertenece a series y capítulos de libro en idioma inglés.

Se realizaron tres presentaciones en congresos internacionales: (Gonzalez 2015b), (Gonzalez 2016), y (Gonzalez 2021d); mientras que las presentaciones en congresos nacionales y regionales fueron tres: (Gonzalez 2020a), (Gonzalez 2021c), y (Gonzalez 2022). Adicionalmente, se realizaron en colaboración seis presentaciones, dos como autor principal y cuatro como coautor, de avances de líneas de investigación en el Workshop de Investigadores de Ciencias de la Computación (WICC): (Gonzalez 2015a), (Calot 2015), (Calot 2016), (Ierache 2019), (Concilio 2020), y (Gonzalez 2021e).

3.4 - Transferencia Tecnológica

En el ámbito nacional, se ha contribuido con entidades del Ministerio de Seguridad, entre ellos la Policía de Seguridad Aeroportuaria (PSA), facilitando el uso experimental de la herramienta solamente con fines exploratorios, para su empleo en la identificación de personal para auditoría forense. Adicionalmente, se dictó un seminario de la temática de la tesis al personal de la fuerza.

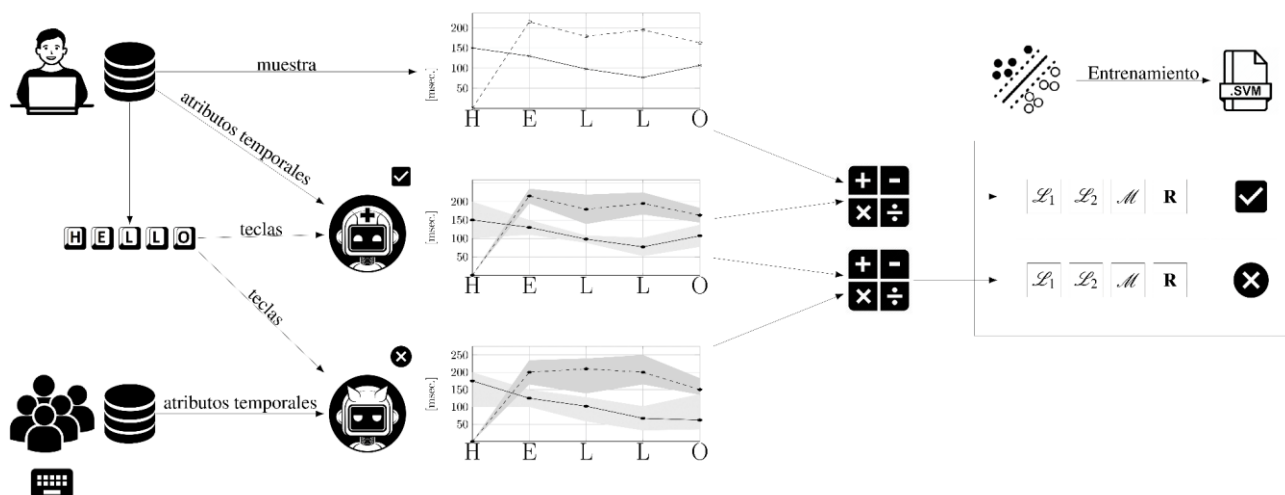


Figura 2 - Entrenamiento del modelo de detección de vida con muestras legítimas y sintetizadas.

En el ámbito internacional se llevó a cabo un proceso de transferencia tecnológica en contacto con empresas privadas del ámbito de la seguridad informática, con el objetivo de robustecer productos existentes para la autenticación con cadencias de tecleo y aplicarlos a la seguridad bancaria.

3.5 - Conjuntos de datos

Los conjuntos de datos de entrenamiento, evaluación, y resultados para los tres experimentos principales de esta tesis fueron puestos a disposición de la comunidad de investigadores en forma libre y gratuita, en los repositorios IEEE DataPort y Mendeley Data. Los conjuntos de datos del experimento sobre distribuciones subyacentes pueden encontrarse en (Gonzalez 2021l) y (Gonzalez 2021m), los del experimento sobre síntesis de muestras y detección de vida en (Gonzalez 2021h) y (Gonzalez 2021i), y los del experimento sobre identificación del texto ingresado en (Gonzalez 2021j) y (Gonzalez 2021k). Un conjunto de datos extendido que subsume los dos anteriores ha sido descrito en la publicación con revisión de pares (Gonzalez 2023b), que acompaña al artículo (Gonzalez 2021b).

Adicionalmente, se han hecho públicos en los repositorios IEEE DataPort y Mendeley Data los conjuntos de datos de entrenamiento, evaluación, y resultados de los siguientes artículos que forman parte de esta tesis: para (Gonzalez 2021c) en (Gonzalez 2021f) y

(Gonzalez 2021g), para (Gonzalez 2020b) en (Gonzalez 2021n).

3.6 - Honores y menciones

En el marco de esta tesis, el artículo *Exploración de correlaciones internas de los parámetros temporales generados en dinámicas de tecleo* (Gonzalez 2020a), presentado en el XXVI Congreso Argentino de Ciencias de la Computación (CACIC), año 2020, ha obtenido el premio al mejor trabajo en el IX Workshop de Seguridad Informática (WSI). Asimismo, el artículo *Un método de ensamble basado en subsecuencias a nivel de palabras para la autenticación de usuarios con cadencias de tecleo en textos libres* (Gonzalez 2021c), presentado en el XXVII Congreso Argentino de Ciencias de la Computación (CACIC), año 2021, ha obtenido el premio a la mejor presentación en el X Workshop de Seguridad Informática.

4 - Métodos propuestos

4.1 - Distancias basadas en histogramas empíricos

El experimento sobre distribuciones empíricas de los tiempos de retención y latencia entre teclas en la escritura en texto libre (Gonzalez 2021a) demostró que los modelos estadísticos paramétricos no son suficientes para capturar las variaciones en la forma de estas distribuciones. En consecuencia, se propuso una familia de

distancias basadas en los histogramas empíricos del perfil biométrico del usuario, utilizando la función inversa de la distribución acumulada de probabilidad como contraparte de las distancias clásicas de Manhattan, euclídea, Canberra, y de Minkowski. Las distancias propuestas demostraron ser más sensibles que las clásicas para la verificación de identidad y la detección de muestras sintéticas (Gonzalez 2021b), a la vez que proporcionaron mayor ganancia de información a los clasificadores empleados para la detección de vida e identificación del texto ingresado.

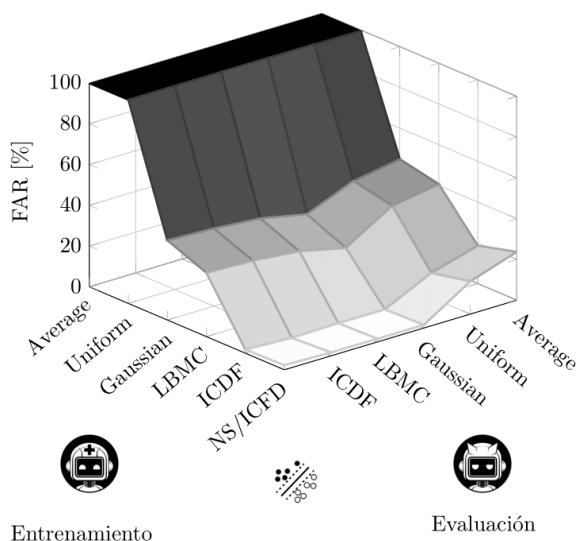


Figura 3 - Tasas de falsos positivos de métodos de síntesis propuestos, al ser utilizados para entrenar y para atacar al modelo de detección de vida.

4.2 - Métodos de síntesis basados en histogramas empíricos

La efectividad de las distancias basadas en histogramas empíricos para discriminar entre usuarios legítimos e impostores motivó la introducción de una familia de métodos de síntesis de muestras de cadencias de tecleo también basados en histogramas empíricos. El experimento sobre síntesis de muestras y detección de vida (Gonzalez 2021b) reveló que estos métodos, siempre que se empleen

contextos de orden elevado que consideren las correlaciones internas a nivel de palabras (Gonzalez 2020a, Gonzalez 2020b), son capaces de engañar a un sistema de verificación de identidad con mayor frecuencia que otros métodos del estado del arte.

4.3 - Modelo de detección de vida con adversarios sintéticos

Se propuso un modelo de detección de vida, capaz de discriminar entre muestras auténticas y sintetizadas, que emplea los métodos de síntesis basados en histogramas empíricos y otros del estado de arte como adversarios para el entrenamiento. El experimento sobre síntesis de muestras y detección de vida (Gonzalez 2021b) demostró que un modelo tal es capaz de neutralizar un ataque de presentación con muestras sintéticas con muy bajas tasas de falsos positivos ($< 2\%$) si el atacante no cuenta con demasiada información sobre el usuario legítimo, y con tasas aceptables ($< 15\%$) incluso si el atacante cuenta con el perfil biométrico completo del usuario legítimo. La tasa de falsos negativos en todos los casos es menor al 5%. Al utilizar atributos derivados basados en el grado de desorden local (Gonzalez 2022) se logró mejorar el rendimiento gradualmente.

4.4 - Identificación de texto ingresado en base a tiempos de escritura

Dada una muestra de tiempos de escritura que carezca de la secuencia de teclas correspondiente, como por ejemplo la resultante de una ataque por canal lateral, es posible reconstruir el texto ingresado con alta probabilidad si el texto no es demasiado extenso (Monaco 2019). Un método para tal fin pero capaz de abordar muestras más largas (en el orden de los 100 caracteres), basado en el anterior modelo de detección de vida e ilustrado en la Figura 4, fue propuesto y evaluado como parte de esta tesis (Gonzalez2021d), alcanzando tasas de error muy bajas ($< 2\%$) pero con la limitación de requerir una lista de candidatos parciales.

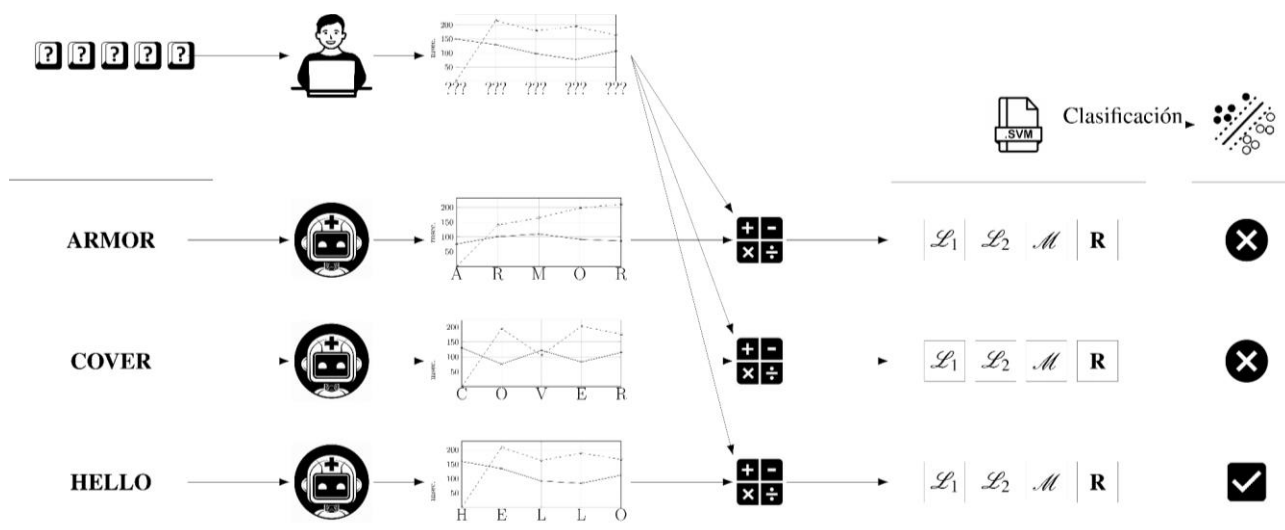


Figura 4 - Identificación del texto ingresado en base a los tiempos de escritura

5. Síntesis de los resultados

La conclusión principal de esta tesis consiste en haber establecido la validez de la hipótesis de trabajo; esto es, que *ninguna distribución suave es adecuada para modelar en general todos los atributos temporales que caracterizan la cadencia de tecleo; sólo las distribuciones empíricas del perfil intra-usuario pueden capturar con precisión su comportamiento característico, y este fenómeno puede ser capitalizado tanto para generar muestras sintéticas que engañan a los actuales sistemas de autenticación como para construir medidas de defensa eficaces contra ataques de presentación, al distinguir la escritura del usuario humano legítimo de una muestra construida artificialmente.*

Adicionalmente, cada experimento arrojó una conclusión particular dentro de su alcance. El experimento sobre distribuciones subyacentes nos enseñó que *la distribución log-logística es una clara ganadora entre todos los candidatos para ajustar los histogramas de tiempo, tanto de retención como de latencia, pero las tasas de rechazo de hipótesis y los méritos relativos muestran que un enfoque que considere los histogramas empíricos en su individualidad es preferible.*

De esta forma se motivaron y justificaron tanto las distancias como los métodos de síntesis basados en histogramas empíricos. Al evaluar su rendimiento en un ataque de presentación, se estableció que *las*

estrategias de síntesis basadas en histogramas empíricos alcanzan un mejor rendimiento que aquellas basadas en distribuciones suaves al intentar engañar a un sistema de autenticación.

Pero también se estableció que, al utilizarlas como adversarios en un esquema de detección de vida, resulta que *las distancias basadas en histogramas empíricos alcanzan un mejor rendimiento que las tradicionales.*

Finalmente, al evaluar una extensión del método anterior se descubrió que *utilizando sólo atributos temporales es posible identificar el texto ingresado dentro de una lista de candidatos de tamaño mediano, alcanzando tasas de error muy bajas y competitivas con el estado del arte aunque tratemos con textos y listas de candidatos más largas.*

5.1 Resultados cuantitativos

Para la evaluación de los métodos propuestos, se utilizaron tres conjuntos de datos (LSIA, KM, y PROSODY, dividido en GAY, GUN, y REVIEW) de gran extensión, públicamente accesibles, y que han sido empleados en estudios previos sobre cadencias de tecleo. Las figuras 5 y 6 muestran las tasas de falsos positivos que cada método de síntesis alcanza contra el sistema de detección de vida, para cada conjunto de datos, con perfiles inter- e intrausuario, y sus intervalos de confianza del 95%.

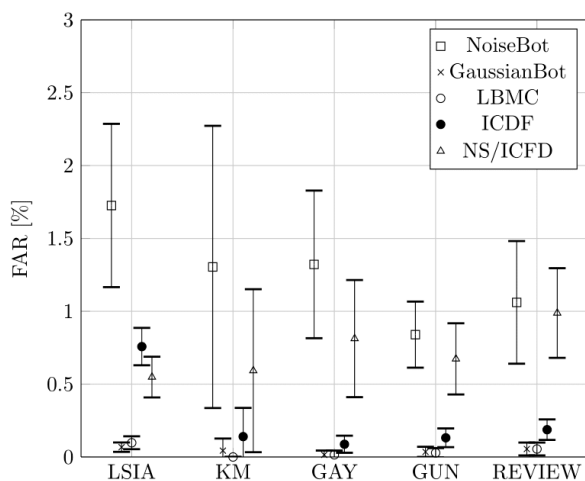


Figura 5 - Tasas de falsos positivos para distintos métodos de síntesis de muestras, CI 95%, perfil interusuario.

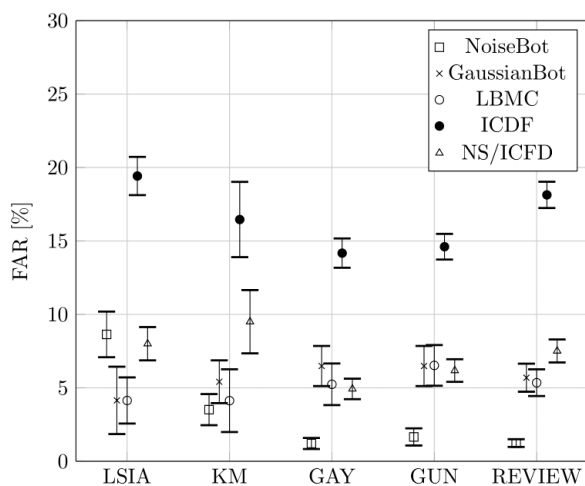


Figura 6 - Tasas de falsos positivos para distintos métodos de síntesis de muestras, CI 95%, perfil intrausuario.

6. Conclusiones

El resultado principal de esta tesis es un modelo de detección de vida que defiende los sistemas de autenticación continua con cadencia de tecleo frente a ataques de presentación. Basado en un modelo de aprendizaje automático entrenado con muestras sintéticas, éste alcanza una tasa de falsos positivos menor al 2% contra muestras sintetizadas con perfiles interusuario y menores al 15% incluso en el caso extremo en que el atacante cuenta con el perfil biométrico completo del usuario legítimo; estos valores deben contrastarse con el 98%+ de falsos positivos alcanzado sin detección de vida. Para

ser utilizados como adversarios, se propusieron diversos métodos de síntesis capaces de inducir tasas elevadas de falsos positivos en sistemas convencionales, y cuyas características principales son el empleo de modelos de orden elevado e histogramas empíricos recabados del perfil biométrico del usuario. Adicionalmente, se descubrió que una modificación del método de detección de vida permite reconstruir el texto ingresado en base a los tiempos de escritura, extendiendo el largo de las muestras tratables con los métodos existentes y potenciando los ataques por canal lateral que obtengan tales tiempos.

Se ha revelado la vulnerabilidad inherente en los sistemas de autenticación continua por cadencias de tecleo cuyo diseño no tiene en consideración los recursos a los que puede acceder un atacante sofisticado, justificando la necesidad de evaluar su rendimiento con modelos de ataque realistas y de incluir modelos de detección de vida como parte del proceso de verificación. Por otra parte, se ha mostrado que estos últimos hacen factible reducir la eficacia de un ataque de presentación con muestras sintéticas en un orden de magnitud o más, restaurando la confianza en la verificación de identidad con cadencias de tecleo en textos libres.

Los resultados de esta tesis, tanto para el robustecimiento de sistemas de autenticación existentes como para la auditoría forense, fueron validados en el transcurso del proceso de transferencia tecnológica con entidades del sector público nacional y privado internacional.

Como producción científica se obtuvieron dos publicaciones internacionales en revistas de primer cuartil (Q1), dos publicaciones suplementarias en revistas internacionales, una en series y capítulos de libro en inglés, tres presentaciones en congresos internacionales, tres en congresos nacionales y regionales, y seis presentaciones en workshops nacionales. Adicionalmente, se publicaron seis conjuntos de datos relacionados con los experimentos de esta tesis.

6.1 Futuras líneas de investigación

Durante el transcurso de esta tesis, los resultados han abierto diversos interrogantes. El hilo conductor de todos ellos puede

resumirse como la búsqueda de modelos, cuantitativos y explicativos, que den cuenta de las diferencias entre una muestra escrita por un usuario humano y otra sintetizada, para que el método de detección de vida explote en su tarea de discriminarlas. Lo que, en vista de los recientes desarrollos con redes neuronales adversarias generativas y *deepfakes*, plantea la

crucial pregunta: *¿Es factible sintetizar muestras de cadencias de tecleo indistinguibles de sus contrapartes reales?* Intentar contestarla abre, sin duda, la más relevante y promisorio futura línea de investigación que puede derivarse de los resultados de esta tesis.

Referencias

- (Bours 2009) Patrick Bours y Hafez Barghouthi. “Continuous authentication using biometric keystroke dynamics”. En: The Norwegian Information Security Conference (NISK). Vol. 2009. 2009.
- (Calot 2015) Enrique P. Calot y col. “Líneas de investigación del Laboratorio de Sistemas de Información Avanzados”. En: Proceedings del XVII Workshop de Investigadores en Ciencias de la Computación (WICC). Salta, Argentina, jun. de 2015, pág. 4. ISBN: 978-987-633-134-0. URL: hdl.handle.net/10915/46107
- (Calot 2016) Enrique P. Calot y col. “Avances en educación de dinámica de tecleo y el contexto emocional de un individuo aplicando interfaz cerebro-computadora”. En: Proceedings del XVIII Workshop de Investigadores en Ciencias de la Computación (WICC). Entre Ríos, Argentina, jun. de 2016, págs. 872-876. ISBN: 978-950-698-377-2.
- (Concilio 2020) Germán Concilio y col. “Avances en reconocimiento de patrones de tecleo para la identificación de personas en ambientes web”. En: XXII Workshop de Investigadores en Ciencias de la Computación (WICC). El Calafate, Santa Cruz, Argentina, 2020, págs. 842-846. ISBN: 978-987-3714-82-5.
- (Epp 2011) Clayton Epp, Michael Lippold y Regan L Mandryk. “Identifying emotional states using keystroke dynamics”. En: Proceedings of the sigchi conference on human factors in computing systems. 2011, págs. 715-724
- (Gaines 1980) R Stockton Gaines, William Lisowski, S James Press y Norman Shapiro. “Authentication by keystroke timing: Some preliminary results”. Inf. téc. Rand Corp Santa Monica CA, 1980.
- (Gonzalez 2015a) Nahuel González y col. “Educación de dinámica de tecleo centrado en el contexto emocional de un individuo”. En: Proceedings del XVII Workshop de Investigadores en Ciencias de la Computación (WICC). Salta, Argentina, jun. de 2015, pág. 5. ISBN: 978-987-633-134-0. URL: hdl.handle.net/10915/4626
- (Gonzalez 2015b) Nahuel González y Enrique P. Calot. “Finite Context Modeling of Keystroke Dynamics in Free Text”. En: Biometrics Special Interest Group (BIOSIG), 2015 International Conference of the, IEEE. Septiembre 2015. ISBN: 978-3-88579-639-8. doi: [10.1109/BIOSIG.2015.7314606](https://doi.org/10.1109/BIOSIG.2015.7314606).
- (Gonzalez 2016) Nahuel González, Enrique P. Calot y Jorge S. Ierache. “A replication of two free-text keystroke dynamics experiments under harsher conditions”. En: 2016 International Conference of the Biometrics Special Interest Group (BIOSIG), IEEE. ISBN: 978-1-50900-780-6. doi: [10.1109/BIOSIG.2016.7736905](https://doi.org/10.1109/BIOSIG.2016.7736905).
- (Gonzalez 2020a) Nahuel González, Germán Concilio, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “Exploración de correlaciones internas de los parámetros temporales generados en dinámicas de tecleo”. En: XXVI Congreso Argentino de Ciencias de la Computación (CACIC), 2020, págs. 726-735. ISBN: 978-987-4417-90-9.
- (Gonzalez 2020b) Nahuel González, Germán Concilio, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “Exploring Internal Correlations in Timing Features of Keystroke Dynamics at Word Boundaries and Their Usage for Authentication and Identification”. En: Computer Science—CACIC 2020: 26th Argentine Congress, CACIC 2020, San Justo, Buenos Aires, Argentina, October 5–9, 2020, Revised Selected Papers. Vol. 1. Communications in Computer and Information Science, Springer Nature, págs. 321-332. ISBN: 978-3-030-75835-6. doi: [10.1007/978-3-030-75836-3_22](https://doi.org/10.1007/978-3-030-75836-3_22).
- (Gonzalez 2021a) Nahuel González, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “On the shape of timing distributions in free-text keystroke dynamics profiles”. En: Heliyon 7.11 (2021), Elsevier. ISSN: 2405-8440. doi: [10.1016/j.heliyon.2021.e08413](https://doi.org/10.1016/j.heliyon.2021.e08413).
- (Gonzalez 2021b) Nahuel González, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “Towards liveness detection in keystroke dynamics: Revealing synthetic forgeries”. En: Systems and Soft Computing (2022), Elsevier. ISSN: 2772-9419. doi: [10.1016/j.sasc.2022.200037](https://doi.org/10.1016/j.sasc.2022.200037).

- (Gonzalez 2021c) Nahuel González, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “*Un método de ensamble basado en subsecuencias a nivel de palabras para la autenticación de usuarios con cadencias de tecleo en textos libres*”. En: XXVII Congreso Argentino de Ciencias de la Computación (CACIC), 2021, págs. 685-694. ISBN: 978-987-633-574-4.
- (Gonzalez 2021d) Nahuel González, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “*The Reverse Problem of Keystroke Dynamics: Guessing Typed Text with Keystroke Timings Only*”. En: 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), 2021, IEEE. ISBN: 978-1-6654-4231-2. doi: [10.1109/ICECET52533.2021.9698782](https://doi.org/10.1109/ICECET52533.2021.9698782).
- (Gonzalez 2021e) Nahuel González y col. “Avances en robustecimiento ante ataques de presentación y falsificación para sistemas basados en el análisis de patrones de tecleo”. En: XXIII Workshop de Investigadores en Ciencias de la Computación (WICC). Chilecito, La Rioja, 2021, pág. 177. ISBN: 978-987-24611-4-0.
- (Gonzalez 2021f) Nahuel González. Dataset for An Ensemble Method for Keystroke Dynamics Authentication in Free-Text Using Word Boundaries. Ver. 1. 2021. doi: 10.17632/np9hhy6gt7.1
- (Gonzalez 2021g) Nahuel González. Dataset for An Ensemble Method for Keystroke Dynamics Authentication in Free-Text Using Word Boundaries. Ver. 1. 2021. doi: 10.21227/jdzh-4m97
- (Gonzalez 2021h) Nahuel González. Dataset for Towards Liveness Detection in Keystroke Dynamics: Revealing Synthetic Forgeries. Ver. 1. doi: 10.17632/xvg5j5z29p.1.
- (Gonzalez 2021i) Nahuel González. Dataset for Towards Liveness Detection in Keystroke Dynamics: Revealing Synthetic Forgeries. Ver. 1. doi: 10.21227/1ka3-er49.
- (Gonzalez 2021j) Nahuel González. Dataset for The Reverse Problem of Keystroke Dynamics: Guessing Typed Text with Keystroke Timings. Ver. 1. doi: 10.17632/94dwkbf2d.1.
- (Gonzalez 2021k) Nahuel González. Dataset for The Reverse Problem of Keystroke Dynamics: Guessing Typed Text with Keystroke Timings. Ver. 1. doi: 10.21227/7616-7964.
- (Gonzalez 2021l) Nahuel González. Dataset of Timing distributions in free text keystroke dynamics profiles. Ver. 1. doi: 10.17632/sjk7kz35nh.1. URL: <https://data.mendeley.com/datasets/sjk7kz35nh/1> (visitado 04-03-2021)
- (Gonzalez 2021m) Nahuel González. Dataset of Timing distributions in free text keystroke dynamics profiles. Ver. 1. doi: 10.21227/ngv9-fa18.
- (Gonzalez 2021n) Nahuel González. Dataset for Exploring internal correlations in timing features of keystroke dynamics at word boundaries and their usage for authentication and identification. Ver. 1. doi: 10.17632/vx83444p8n.1.
- (Gonzalez 2022) Nahuel González, Jorge S. Ierache, Enrique P. Calot, y Waldo Hasperué. “*Atributos derivados para la clasificación de cadencias de tecleo en texto libres basados en el grado de desorden local*”. En: XXVIII Congreso Argentino de Ciencias de la Computación (CACIC), 2022.
- (Gonzalez 2023a) Nahuel González. “*KSDSLD—A tool for keystroke dynamics synthesis & liveness detection*”. En: Software Impacts (2022), Elsevier. doi: [10.1016/j.simpa.2022.100454](https://doi.org/10.1016/j.simpa.2022.100454).
- (Gonzalez 2023b) Nahuel González. “*Dataset of Human-written and Synthesized Samples of Keystroke Dynamics Features for Free-text Inputs*”. En: Data in Brief (2022), Elsevier (en prensa).
- (Ierache 2019) Jorge S. Ierache y col. “*Líneas de investigación del Laboratorio de Sistemas de Información Avanzados: Dinámica de Tecleo, Computación Afectiva, Extracción de Relaciones Semánticas, Blockchain, y Smart Contracts*”. En: XXI Workshop de Investigadores en Ciencias de la Computación (WICC).
- (Milne 2018) A Milne, K Farrahi y MA Nicolaou. “*Less is more: Univariate modelling to detect early Parkinson’s disease from keystroke dynamics*”. En: International Conference on Discovery Science. Springer. 2018, págs. 435-446.
- (Monaco 2019) John V Monaco. “*What are you searching for? a remote keylogging attack on search engine autocomplete*”. En: 28th {USENIX} Security Symposium ({USENIX} Security 19). 2019, págs. 959-976.
- (Monrose 1990) Fabian Monrose y Aviel D Rubin. “*Keystroke dynamics as a biometric for authentication*”. En: Future Generation computer systems 16.4 (2000), págs. 351-359.
- (Rahman 2011) KA Rahman, KS Balagani y VV Phoha. “*Making impostor pass rates meaningless: A case of snoop-forged-replay attack on continuous cyber-behavioral verification with keystrokes*”. En: CVPR 2011 workshops. IEEE. 2011, págs. 31-38.
- (Song 2011) DX Song, DA Wagner y X Tian. “*Timing analysis of keystrokes and timing attacks on ssh.*” En: USENIX Security Symposium. Vol. 2001. 2001.

Cálculo Científico Distribuido sobre Clientes Móviles Indeterminados en redes MANETs

Autor: Pablo José Iuliano

Director/Co Director: Ing. Luis Marrone - Dr. Fernando G. Tinetti

Afiliación: Facultad de Informática. Universidad Nacional de La Plata. La Plata, Argentina

e-mail de contacto del tesista: piuliano@info.unlp.edu.ar

Palabras Claves: MANET; Protocolo; Incertidumbre; Clustering; Cross-Layer;

1. MOTIVACION

En este trabajo de investigación se propone el diseño de un novel protocolo que combina un enfoque para lidiar con la indeterminación de clientes al momento de comenzar un cálculo científico distribuido basado en un diseño cross layer con un mecanismo que agrupa los dispositivos presentes en una red móvil ad hoc o MANET (*Mobile Ad hoc Networks*) en clusters, sacando rédito de la heterogeneidad inherente a este tipo de redes a fin de determinar qué dispositivos serán los líderes de los agrupamientos conformados.

2. INTRODUCCIÓN

La tecnología wireless en los últimos años ha experimentado una expansión sin precedentes, llegando hasta el lugar más recóndito de la tierra. La tecnología inalámbrica permite la creación de redes sin la necesidad de cables, dando al usuario de este tipo de redes la libertad de moverse por su entorno sin la restricción de movilidad asociada a tener una conexión cableada.

De acuerdo al criterio dominante actualmente este tipo de redes tienden a ser centralizadas en torno a un dispositivo inalámbrico estacionario el cual ofrece el servicio de acceso a la red, sumado en la mayoría de los casos el acceso a internet. Sin embargo, los dispositivos de computación móviles continúan reduciendo sus precios e incrementando sus capacidades de cómputo; volviéndose cada vez más interesantes como alternativa barata de computación distribuida para cálculos científicos, sumado al hecho que los bajos

precios han popularizado la computación móvil y en ese sentido la exploración en cuanto a la posibilidad de conformar de una red móvil ad hoc o MANET (*Mobile Ad hoc Networks*) a gran escala destinadas al cómputo científico se está convirtiendo plausible.

Cuando el objetivo de una investigación científica es obtener información válida de algún entorno con el cual se interactúa, la mayoría de las veces es prácticamente imposible establecer vínculos fijos permanentes, en otras palabras enlaces cableados normales o enlaces inalámbricos centralizados en torno a un dispositivo estacionario. Para estos escenarios existe una nueva tecnología, basada en un nuevo paradigma de los sistemas de comunicación, que han venido a denominarse red móvil ad hoc o MANET.

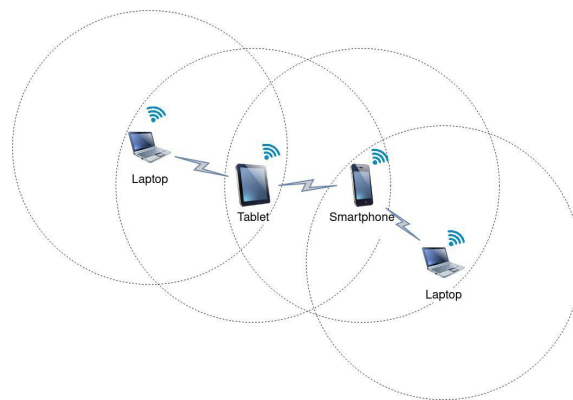


Figura 1: Esquema simplificado de una red móvil ad hoc

3. REDES MÓVILES AD HOC

Formalmente las redes móviles ad hoc se

pueden definir como sistemas autónomos de nodos móviles con capacidades de enrutamiento y conectados a través de enlaces inalámbricos formando un grafo de comunicación arbitrario.

Con respecto al modelo de referencia TCP/IP, las redes móviles ad hoc definen una capa adicional entre la capa de red y las tecnologías inalámbricas subyacentes. Esta capa es necesaria debido a que este tipo de redes deben mantener un solo dominio de difusión en entornos donde la visibilidad del nodo está limitada por el rango de transmisión de la interfaz inalámbrica.

En este tipo de redes la topología varía dinámica y libremente a medida que los nodos se van desplazando por su entorno. Básicamente el modo de operación de este tipo de redes es *peer to peer multihop*, donde los paquetes de información son transmitidos de manera de almacenamiento y reenvío desde un origen a un destino arbitrario pudiendo involucrar varios o ningún intermediario para alcanzar al destino.

Dentro de las MANET el enrutamiento de paquetes que es soportado de manera distribuida y la falla de un solo dispositivo no debería afectar significativamente al mismo.

Las problemáticas asociadas a las cuestiones energéticas deben ser consideradas debido a que, muchos o todos los nodos que componen las redes móviles ad hoc utilizan baterías o alguna otra forma de alimentación agotable. Con lo cual para este tipo de nodo se debe considerar que tengan presentes medias de preservación de la energía. El ancho de banda es un problema que está presente en todos los tipos de redes actuales y particularmente en las MANET en donde se hace más evidente esta problemática. Con lo cual el *Throughput* real es siempre menor que la tasa máxima de transmisión de los canales de radio, esta merma radica por efecto del *multihop*, el ruido ambiental, la atenuación, etc.

4. PROTOCOLOS DE ENRUTAMIENTO PARA MANETS

En las redes móviles ad hoc, tanto la topología y los estados de los enlaces entre los dispositivos que la componen son dos propiedades que son impredecibles y que constantemente se van modificando, conllevando a que la eficiencia de las técnicas antes descritas se vea seriamente degradada. Como consecuencia directa de esta última contrariedad se han desarrollado varios protocolos de ruteos específicos para este tipo de redes y caen dentro de una de estas tres categorías: Protocolos proactivos, reactivos o híbridos [1].

- Los protocolos proactivos mantienen la información de enrutamiento incluso antes de que se necesite. Todos y cada uno de los nodos en la red mantienen información de las rutas hacia todos los demás nodos de la red. La información de las rutas se mantiene en las tablas de enrutamiento y se actualizan periódicamente a medida que cambia la topología de la red. Las actualizaciones generan una excesiva sobrecarga en la red debido a los cambios frecuentes en las rutas por lo dinámico de este tipo de redes y el ancho de banda limitado que esta posee.
- Los protocolos reactivos no mantienen información o generan actividad de ruteo en los nodos de la red si no es necesario establecer comunicación entre ellos. Un nodo al enviar un paquete a un destino, provoca que se busque la ruta y se establezca la conexión. El procedimiento anterior es realizado a demanda. El descubrimiento de ruta por lo general se produce inundando toda la red con paquetes de petición de ruta. Estos protocolos experimentan cierto retardo asociado a la búsqueda de las rutas o porque las rutas pueden no ser válidas durante el tiempo suficiente. Por otro lado, utilizan eficientemente el ancho de banda de la red y el consumo de energía, debido a que buscan las rutas sólo cuando se necesitan.
- Los protocolos híbridos son derivación de los dos primeros, combinan técnicas de

los protocolos proactivos y de los protocolos reactivos.

AODV (*Ad hoc On Demand Distance Vector*) [2] [3], es un protocolo de ruteo reactivo para redes móviles ad hoc basado en vector distancias que funciona bajo demanda. Solo comienza el descubrimiento de rutas cuando es requerido y además las rutas se almacenan en tablas de ruteo. Estas últimas son mantenidas y almacenadas en cada nodo de la red. Debido a que las rutas se crean únicamente cuando son requeridas se minimiza el consumo de ancho de banda y de energía.

Las tablas creadas por AODV guardan información solamente de las rutas hacia los destinos con los cuales se mantiene comunicación, lo cual genera poca sobrecarga en la red y las rutas se van descubriendo conforme se van necesitando.

AODV soporta tres tipos de transmisión: unicast, multicast y broadcast; no obstante, AODV descubre las rutas basándose en peticiones broadcast y las respuestas son enviadas en modo unicast [4]. Inicialmente la tabla de ruteo solamente contiene entradas para los nodos vecinos. Si un origen necesita alcanzar un destino que no es vecino, debe obtener una ruta que le indique como hacerlo. Para ello, el nodo origen envía por broadcast mensajes de descubrimiento de ruta (RREQ *Route Request*). Si un nodo que no es el destino recibe este paquete, verifica los números de secuencia a fin de determinar si lo ha recibido previamente. En caso de que no lo haya recibido, lo reenvía incrementando el número de saltos y crea el camino inverso para la respuesta RREP (*Route Reply*). La respuesta del establecimiento de ruta RREP es generada por los nodos intermedios que cuentan con una ruta hacia el destino o el mismo destino una vez que el mensaje RREQ le haya llegado. Los nodos envían estas respuestas RREP al resto que le habían enviado el mensaje RREQ. La ruta se confirma una vez que el RREP llega al origen que había comenzado con el descubrimiento de ruta.

Es necesario mantener el enlace activo entre el origen-destino después de establecerse una

ruta, ya que ésta se mantiene válida durante un periodo de tiempo. Si los nodos que tienen rutas activas realizan envíos periódicos de mensajes RREP ajustando a 1 el TTL, es decir, mediante la difusión local del mensaje a sólo los dispositivos situados a 1 salto de distancia este objetivo se asegura. Otro mecanismo adicional para garantizar la validez de las rutas consiste en enviar periódicamente mensajes HELLO entre los nodos vecinos a fin de determinar si se encuentran activos en ese instante en la red. Si luego de algún tiempo no se reciben mensajes HELLO desde un vecino o se produce un error de envío, se origina un mensaje RERR (*Route Error*), que se emplea para invalidar una ruta. De modo que, el mensaje RERR llegará hasta el origen de la ruta y este desechará la ruta borrándola de la tabla enrutamiento.

5. CLUSTERING EN MANET

El proceso mediante el cual se divide una red en estructuras interconectadas se denomina agrupamiento o clustering y las subestructuras resultantes se las denomina grupos lógicos o clusters. En estos escenarios los nodos de la red son divididos y ubicados dentro de grupos lógicos. Una estructura típica de clusters es la que se muestra en la Figura 2.

Los nodos normalmente en este tipo de esquemas asumen roles diferentes, como líder o coordinador, gateway o simplemente miembro del agrupamiento lógico. Los coordinadores mantienen y administran toda la actividad dentro del cluster al que pertenecen. Un nodo gateway se encarga de toda la comunicación entre clusters. Hay esquemas donde este último tipo es omitido y por lo tanto no existe la diferenciación entre tráfico intra del inter cluster.

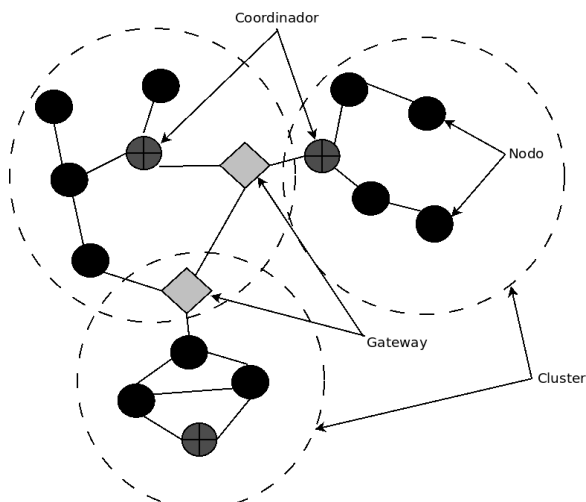


Figura 2: Ejemplo de una red dividida en clusters

Una topología dividida en clusters mejora la performance de una MANET junto con otras ventajas como el mejor aprovechamiento del ancho de banda, el consumo de energía y el rendimiento también se mejoran.

6. DISEÑO CROSS LAYER

Históricamente las arquitecturas de red asumen que las funciones de comunicación están organizadas en capas y la entrega de paquetes son llevadas a cabo por protocolos que operan en distintas capas [5]. De esta manera las funcionalidades de la red en su conjunto son llevadas a cabo a través del modelo en capas. Así cada una de estas define un aspecto particular de la operatoria de la red del cual se encargará de resolver y proveer servicios a la capa inmediatamente superior. Otra característica de las capas dentro de este modelo es que son modulares y cada una de ellas implementa una tarea en específica.

La arquitectura en capas prohíbe la comunicación directa entre capas no adyacentes, mientras que la comunicación entre las adyacentes funciona gracias a la utilización de interfaces pre establecidas y estandarizadas [6].

Alternativamente, existen protocolos que son diseñados explícitamente para ignorar la arquitectura en capas, permitiendo la interacción e intercambio de datos entre capas que no se encuentran dispuestas en forma

adyacentes. Este tipo de diseño se utiliza para obtener mejores prestaciones cuando se construyen nuevas soluciones en particular con los nuevos protocolos de enrutamiento [7] [8].

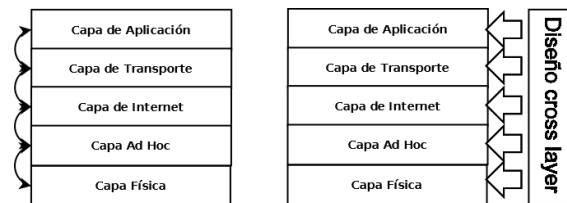


Figura 3: Arquitectura en capas y el diseño cross layer

En la Figura 3 se puede apreciar la arquitectura en capas y el diseño cross layer que permite la interacción e intercambio de datos entre capas no adyacentes, por ejemplo, en el diseño cross layer es posible que parámetros de la capa de aplicación puedan tener comunicación directa con la capa de enlace. Bajo el enfoque cross layer, una capa puede adaptarse a los requerimientos y condiciones particulares de las otras capas.

La motivación para adoptar un diseño cross layer en un protocolo que opere en una MANET se puede fundamentar principalmente en tres tópicos: Soportar la necesidad de que sea adaptable a las condiciones dinámicas de este tipo particular de red [9] [10], satisfacer requisitos específicos del ámbito de aplicación y atender restricciones en cuanto al consumo de energía que debe alcanzar la solución a desarrollar [11] [12].

7. PROTOCOLO PARA EL CÁLCULO CIENTÍFICO DISTRIBUIDO SOBRE CLIENTES MÓVILES INDETERMINADOS EN MANET (PROCCDIMI)

El protocolo propuesto se denomina PROCCDIMI, el cual establece la creación de clústers dinámica y distribuidamente, donde el envío de mensajes relacionados con el cómputo únicamente se circunscribe al clúster. Esta

última cualidad tiene por objetivo de optimizar el consumo de energía, memoria, potencia de cómputo y ancho de banda, gracias al uso eficiente de recursos.

PROCCDIMI implementa de un balanceo de cargas dinámico basándose en multi-agentes y se fundamenta en los protocolos de enrutamiento de vector distancia de MANETs, particularmente en AODV [2]. También se contempla la heterogeneidad existente en este tipo de redes y es por ello que se incorpora la idea en esta propuesta de aprovechar la existencia de algunos dispositivos que poseen capacidades superiores (cómputo, comunicación y energía) con el objetivo de ser elegidos como los líderes de los clústers, o en este caso como coordinadores del cómputo distribuido, a fin de llevar a cabo la conformación del grupo, la asignación de tareas dentro del mismo y realizar el proceso de descubrimiento de nuevos clientes. Por ello es necesario distinguir entre dispositivos con capacidades superiores (DCSs) y dispositivos con capacidades limitadas (DCLs). Una vez hecha la distinción, es posible una clasificación en base a sus capacidades.

Las capacidades de los dispositivos se obtienen a partir del hardware que ellos mismos poseen, es decir, inicialmente un dispositivo conoce sus propias características respecto a la cantidad de memoria, energía, capacidad de cómputo y potencia de señal; por lo tanto, la capacidad de un dispositivo se asigna en función del tipo de dispositivo que se trata.

El criterio para la elección de coordinador del clúster se basa en la selección de los dispositivos con capacidades superiores, en este caso la capacidad de un dispositivo se determina por medio de una estrategia sencilla que asigna un identificador binario a cada dispositivo; es decir, si se trata de un dispositivo con capacidades superiores es un DCS, de lo contrario es un DCL [13] [14].

Para definir la capacidad de un dispositivo se tiene en cuenta las siguientes suposiciones, cada dispositivo es consciente del grado de la capacidad que posee y este puede ser solo de dos tipos: alta (DCS) o limitada (DCL). En

PROCCDIMI los nodos que son líderes de clústers son todos ellos DCS, no implicando que el recíproco sea siempre verdadero.

8. ENTORNO DE SIMULACIÓN

Network simulator NS-2 [15] fue la herramienta seleccionada para evaluar la performance del protocolo. La velocidad de movilidad del nodo varía según la Tabla 1 y de acuerdo con el modelo de movilidad RWP (*Random Waypoint*). El modelo de movilidad describe el patrón de movimiento de los nodos móviles y cada nodo es responsable de calcular su propia posición y velocidad. El tiempo para cada instancia de simulación se establece en 200 segundos. El generador de tráfico CBR (*Constant Bit Rate*) se utiliza para la transferencia de paquetes de datos. Las simulaciones se han realizado utilizando los parámetros antes mencionados y los restantes se pueden observar en la Tabla 1.

Parámetro	Valor
Nro. de dispositivos	10, 20, 50, 100, 150
Nro. de DCSs	6, 12, 30, 60, 100
Nro. de DCLs	4, 8, 20, 40, 50
Nro. de clusters	2, 4, 10, 20, 50
Coords. por cluster	2, 4, 20, 60, 100
Tiempo por ronda	200 seg.
Área de simulación	500 x 500 mts
Mod. de movilidad	RWP
Protocolos de Ruteo	PROCCDIMI, AODV, DSR, DSDV
Protocolo MAC	IEEE 802.11b

Rango de transmisión (Tx)	100 mts.
Velocidad	5, 10, 15 Km/h
Pausas	Ninguna
Energía inicial	100.0 Joules

Tabla 1: Parámetros de simulación

9. MÉTRICAS DE RENDIMIENTO

La evaluación del rendimiento, como así el diseño y el desarrollo de los protocolos de ruteo en MANET requieren que sean examinados meticulosamente bajo parámetros estandarizados dentro de la materia de estudio. Estos últimos son seleccionados de acuerdo a la IETF RFC 2701 [16], la cual recomienda el uso de un conjunto de métricas destinadas al estudio y evaluación de rendimiento de diferentes protocolos pensados para trabajar en redes MANET.

A continuación se detallan las métricas más relevantes utilizadas en las pruebas de rendimientos realizadas en el contexto de la tesis doctoral:

Tasa de Entrega de Paquetes

La Tasa de Entrega de Paquetes se define como el número de paquetes entregados al nodo destino sobre el total enviados por el nodo origen. Con lo cuál esta tasa tomará valores entre 0 y 100 por ciento.

Un paquete que no puede ser ruteado debido a que el nodo origen no conoce ninguna ruta hacia el destino es contado como un paquete enviado pero no entregado, con lo cual se penaliza al protocolo interpretando que la actualización de su conocimiento de la topología de la red es no es óptima. Por lo tanto, una tasa menor a 100 significa que los paquetes se perdieron en el camino o que no se pudo encontrar ninguna ruta al destino. Claramente es deseable tener una alta proporción de entrega de paquetes en cualquier tipo de red y más en las redes inalámbricas

móviles [17]. Esta métrica dimensiona la confiabilidad del protocolo en cuanto a la transmisión de información [18].

Throughput

Throughput es definida como la tasa en que se entregan todos los datos al destino. Es el total de todos los bits exitosamente entregados a destinos individuales y es expresado como bytes por segundo (byte/sec) o bits por segundo (bit/sec). También puede verse como el tiempo que le toma al destino recibir el último mensaje con información [19]. Por consiguiente, es una medida de efectividad de un protocolo.

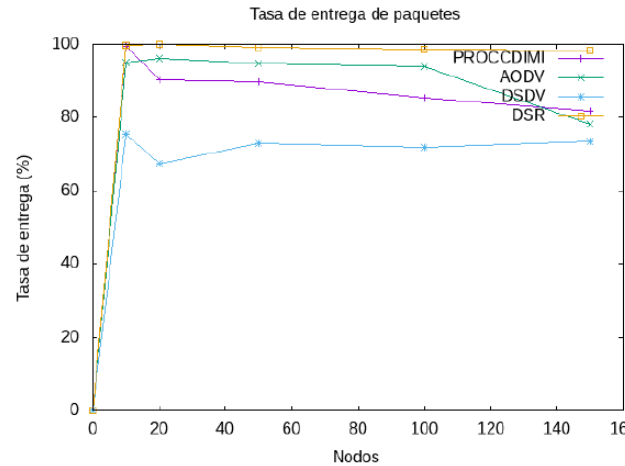
Costo energético promedio

El costo energético promedio dimensiona el consumo de energía promedio de los nodos que conforman la red al utilizar el protocolo que se está evaluando durante la simulación. Este valor se calcula como la suma de las diferencias registradas entre la energía inicial del nodo menos la energía que posee el mismo luego de terminada la simulación, dividida la cantidad de nodos que componen la red. La energía es un recurso escaso y un mal manejo de la misma por parte de un protocolo podría desencadenar el agotamiento de la batería de nodos críticos en términos de conectividad y afectar la disponibilidad de la red, es decir, dejando unidades de red aisladas debido a que las que proveían conectividad de encuentra off line debido a que su batería no posee carga. Por lo tanto, es de suma importancia utilizar la energía de manera eficiente [20].

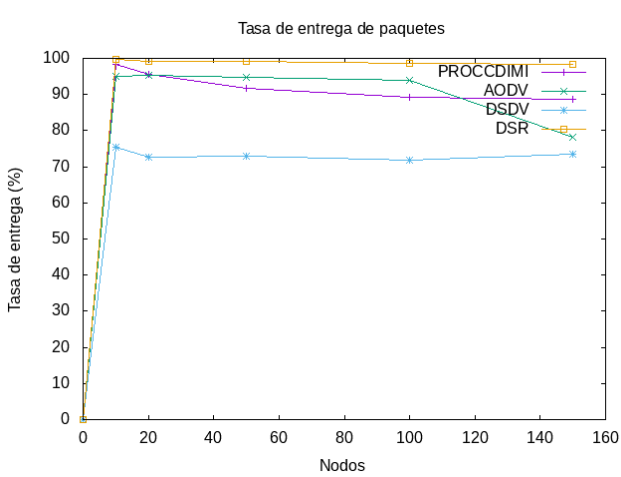
10. RESULTADOS

Los protocolos de enrutamiento AODV, DSR y DSDV solamente se remiten al encaminamiento de paquetes sobre las redes MANET y no realizan ninguna operación adicional a esta última. Mientras que PROCCDIMI resuelve el ruteo junto con todas las tareas involucradas en la gestión del cómputo científico y la administración de los clientes que resuelven las tareas dentro de la computación distribuida. Dada esta

disparidad de responsabilidades y a pesar de la misma, las pruebas de rendimiento se llevaron a cabo utilizando los protocolos antes mencionados, ya que en los escenarios tradicionales lo único común que puede haber son justamente los protocolos de ruteo, debido a que el resto es muy dependiente de cada proceso, de cada forma de construcción, de cada configuración y hasta de cada plataforma, con lo cual es imposible poder realizar una evaluación en los mismos términos. Al no disponerse de un mecanismo, aplicación o herramienta unificadora comparable a nuestra propuesta, que resuelva tanto el ruteo de paquetes como la indeterminación de los clientes de la red MANET más la asignación de tareas y recolección de resultados, entonces se prefirió utilizar los protocolos de enrutamiento tradicionales para realizar la comparaciones con PROCCDIMI donde se tiene presente siempre que este último realiza operaciones adicionales además del enrutamiento de paquetes y evaluándose la performance incluyendo tareas propias del ruteo y lo antes mencionado. Pese a estas diferencias de procesos involucrados los resultados analizados son similares en cuanto a las métricas destinadas a los protocolos de red y en lo concerniente a la métrica del consumo de energía promedio, PROCCDIMI exhibe un comportamiento superior al resto de los protocolos con los cuales se lo contrastó.



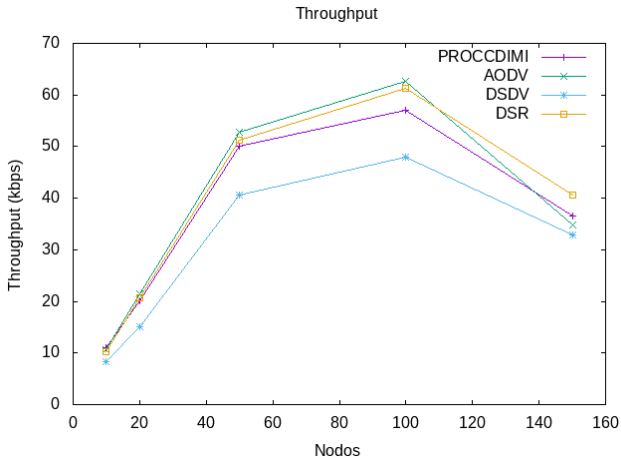
Tasa de entrega de paquetes vs. Número de nodos (enrola nuevos clientes)



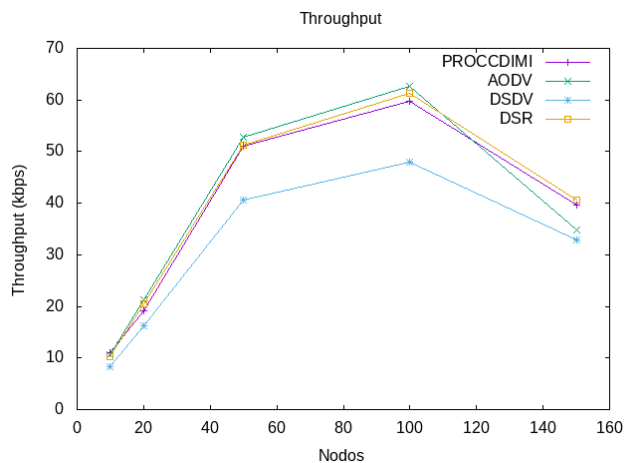
Tasa de entrega de paquetes vs. Número de nodos (enrola nuevos clientes, asigna tareas y colecta los resultados)

Figura 4: Resultados de evaluar la tasa de entrega de paquetes sobre diferentes números de nodos.

En la Figura 4 se observa que PROCCDIMI tiene una tasa de entrega de paquetes notoriamente superior cuando se trata de redes con una cantidad de nodos que oscila entre 10 y 20 nodos y por ende convirtiéndolo en un protocolo más confiable respecto a los demás analizados. Cuando se analiza PROCCDIMI en redes con un número de nodo superior a 20 nodos exhibe una confiabilidad superior al 80% convirtiéndolo en un protocolo que otorga un nivel de confiabilidad que rivaliza con los otros protocolos analizados.



Throughput vs. Número de nodos (enrola nuevos clientes)

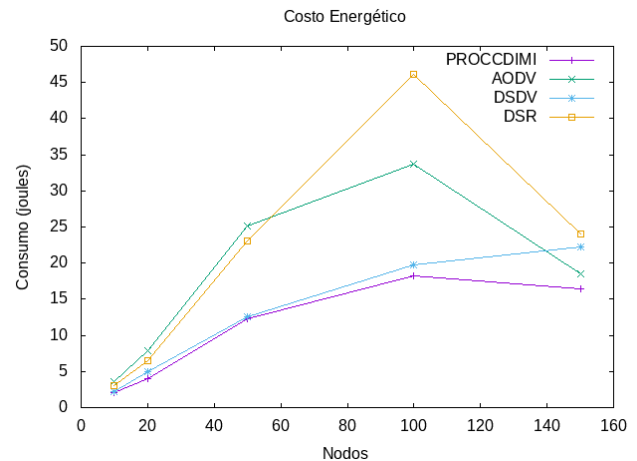


Throughput vs. Número de nodos (enrola nuevos clientes, asigna tareas y colecta los resultados)

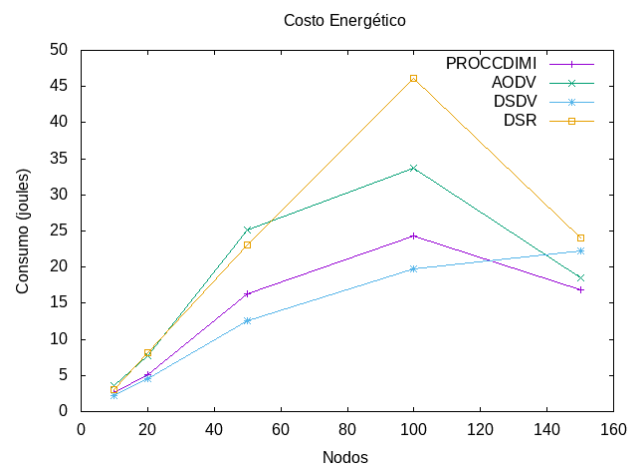
Figura 5: Resultados de evaluar el *Throughput* sobre diferentes números de nodos.

Los resultados que se pueden apreciar en la Figura 5 indican que no hay diferencias significativas entre las curvas de desempeño, sobre todo en las curvas que se corresponden a las simulaciones realizadas con números de nodos relativamente pequeños 10 y 20, puesto que la velocidad de transmisión de todos los protocolos ronda los 10 kbps en el primer caso y 20 kbps en el segundo, con la salvedad de DSDV que está cercano a los 15kbps. Para redes más grandes en términos del número de nodos, los rendimientos se tornan más heterogéneos aunque PROCCDIMI se mantiene en el rango medio de desempeño, no siendo el que registra ni la mejor, ni la peor velocidad de transmisión. Si se tiene en cuenta que el protocolo aquí propuesto además de resolver cuestiones de enrutamiento de paquetes como los demás protocolos contra los que se lo compara, resuelve el descubrimiento de nuevos clientes y asignación de clusters de pertenencia, que se mantenga velocidades de entrega similares a los otros protocolos que solo se encargan de enrutar paquetes es realmente significativo. Este aumento en la tasa de *Throughput* se explica por el diseño de clustering de

PROCCDIMI que otorga una ventaja sobre la capa de acceso al medio ayudando a incrementar la rapidez de la conexión y por tanto, a conseguir mayor tasa de transmisión de datos [21].



Consumo de energía promedio vs. Número de nodos (enrola nuevos clientes)



Consumo de energía promedio vs. Número de nodos (enrola nuevos clientes, asigna tareas y colecta los resultados)

Figura 6: Resultados de evaluar el consumo de energía promedio sobre diferentes números de nodos.

En la Figura 6 se puede apreciar que la curva correspondiente a PROCCDIMI asciende por debajo de las demás curvas, lo que indica que este protocolo consume menos energía que el resto. El hecho de que PROCCDIMI tenga menor gasto energético, es producto de la estrategia de agrupamiento de los nodos, dado que uno de los beneficios que se esperan al formar clusters, es la

reducción del consumo de energía [20].

11. APORTES DE LA TESIS

Los aportes de la investigación realizada son informados a continuación:

- Desarrollo de un nuevo modelo que utiliza un diseño cross layer y de organización de nodos en agrupamientos lógicos que gestiona el cómputo y asigna tareas diferenciales de acuerdo a las prestaciones del nodo para lidiar con la indeterminación de clientes al momento de realizar cálculos científicos sobre MANET.
- Utilizar el protocolo de ruteo reactivo AODV como sustrato para construir una extensión orientada a descubrir los clientes disponibles en una red móvil ad hoc para llevar a cabo un cómputo distribuido.
- Desarrollo de un nuevo protocolo que emerge como solución de los problemas asociados a la indeterminación de clientes al momento de realizar cálculos científicos sobre MANET y la gestión del cómputo distribuido.
- Pruebas de rendimiento de la solución propuesta en la tesis doctoral que proveen un análisis cuantitativo del comportamiento de PROCCDIMI con respecto a los demás protocolos contra los que se lo comparó.

12. LÍNEAS DE I/D FUTURAS

A continuación, se resumen las principales líneas futuras de investigación:

- **Probar el desempeño de PROCCDIMI en entornos más reales:** Modelar la generación de tráfico en la red de forma más realista, para que no sea únicamente la información transmitida entre los nodos de la ruta establecida la que circule por la red, si no que pueda haber interferencias debidas a otras comunicaciones.
- **Implementar un mecanismo de**

seguridad: Otro aspecto a evaluar en el protocolo PROCCDIMI es la incorporación de medidas que provean seguridad-privacidad sobre los datos transmitidos.

- **Implementar la automatización de elección del coordinador de cluster:** En las simulaciones realizadas los escenarios definían estáticamente los tipos de dispositivos y esta definición era realizada ad hoc. Con el propósito de hacer más sólida y estricta la estrategia de elección de coordinador es preciso emplear algoritmos ampliamente probados como por ejemplo: *Highest-Degree heuristic*, *Lowest-ID heuristic*, *Node-weight heuristic*, *Weighted Clustering algorithm* o *Distributed Weighted Clustering Algorithm*.
- **Probar PROCCDIMI con otro protocolo de enrutamiento:** A la luz de los resultados obtenidos en la simulaciones y a la naturaleza intrínseca de los protocolos de enrutamiento que funcionan sobre las MANET, se puede decir que los protocolos reactivos introducen poca carga de paquetes de control, como resultado se optimiza el ancho de banda y se obtienen altas tasas de entrega de paquetes durante el proceso de descubrimiento de nuevos clientes; pero en cambio se registran tiempos de latencias muy elevados. Por tal razón resultaría de interés evaluar a PROCCDIMI con un protocolo de ruteo proactivo como sustrato para realizar toda su operatoria reemplazando a AODV, con el objetivo de dimensionar en qué medida mejoran los tiempos de latencias.

13. REFERENCIAS

- [1] H. Kaur, V. Sahni, and M. Bala, "A survey of reactive, proactive and hybrid routing protocols in manet: A review," *network*, vol. 4, no. 3, pp. 498–500, 2013.
- [2] C. Perkins, E. Belding-Royer, S. Das *et al.*, "Ad hoc on-demand distance vector (AODV) routing," 2003.
- [3] E. M. Belding-Royer and C. E. Perkins, "Evolution and future directions of the ad hoc on-demand distance-vector routing protocol," *Ad Hoc Networks*, vol. 1, no. 1, pp. 125–150, 2003.

- [4] X. Zhang and G. F. Riley, "Performance of routing protocols in very large-scale mobile wireless ad hoc networks," in 13th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems. IEEE, 2005, pp. 115–122.
- [5] J. Edwards and R. Bramante, Networking self teaching guide: OSI, TCP/IP, LANs, MANs, WANs, implementation, management, and maintenance. John Wiley & Sons, 2015.
- [6] G. Tolosa, "Protocolos y modelo osi," Recuperado de <http://www.tyr.unlu.edu.ar/TYRpublica/02> Protocolos y OSI. pdf, 2014.
- [7] A. Awang, K. Husain, N. Kamel, and S. Aissa, "Routing in vehicular ad-hoc networks: a survey on single-and cross- layer design techniques, and perspectives," IEEE Access, vol. 5, pp. 9497–9517, 2017.
- [8] M. Brindha, "A survey on cross layer distributed topology control in mobile adhoc network," Bonfring International Journal of Networking Technologies and Applications, 4 (1), 1, vol. 3, 2017.
- [9] V. Golla, G. Jayanthi, and H. Shivashankar, "Designing energy routing protocol with power consumption optimization in manet," IEEE Transactions on Emerging topics in Computing, vol. 2, no. 2, pp. 192–197, 2014.
- [10] M. Conti, G. Maselli, G. Turi, and S. Giordano, "Cross-layering in mobile ad hoc network design," Computer, vol. 37, no. 2, pp. 48–51, 2004.
- [11] E. M. Belding-Royer, S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, "Routing approaches in mobile ad hoc networks," Mobile ad hoc networking, vol. 1, no. 1, pp. 275– 300, 2004.
- [12] D. Chander and R. Kumar, "Qos enabled cross-layer multicast routing over mobile ad hoc networks," Procedia Computer Science, vol. 125, pp. 215–227, 2018.
- [13] M. A. Wister Ovando, "Arquitectura de descubrimiento de servicios en manet basada en dispositivos de capacidades superiores liderando clusters," Proyecto de investigación, 2009.
- [14] I. D. Chakeres and E. M. Belding-Royer, "Aodv routing protocol implementation design," in 24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings. IEEE, 2004, pp. 698–703.
- [15] U. de California, "The network simulator - ns-2, <https://www.isi.edu/nsnam/ns/>," 1999.
- [16] G. Malkin, "Nortel networks multi-link multi-node ppp bundle discovery protocol," in Technical Report. September 1999. RFC 2701, 1999.
- [17] P. Manickam, T. G. Baskar, M. Girija, and D. D. Manimegalai, "Performance comparisons of routing protocols in mobile ad hoc networks," arXiv preprint arXiv:1103.0658, 2011.
- [18] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking, 1998, pp. 85–97.
- [19] A. Nasipuri, J. Zhuang, and S. R. Das, "A multichannel csma mac protocol for multihop wireless networks," in WCNC. 1999 IEEE Wireless Communications and Networking Conference (Cat. No. 99TH8466), vol. 3. IEEE, 1999, pp. 1402–1406.
- [20] A. P. Pati, K. R. Kanth, B. Sharanya, M. D. Kumar, and J. Malavika, "Design of an energy efficient routing protocol for manets based on AODV," International Journal Of Computer Science Issues (IJCSI), vol. 8, no. 4, p. 215, 2011.
- [21] B. A. Correa, L. Ospina, and R. C. Hincapié, "Survey of clustering techniques for mobile ad hoc networks," Revista Facultad de Ingeniería Universidad de Antioquia, no. 41, pp. 145–161, 2007.

DepProMod: Modelo de Proceso de Despliegue de Sistemas de Software

Doctorado en Ciencias Informáticas. Facultad de Informática. Universidad Nacional de La Plata.

Autora: Marisa Daniela Panizzi, marisapanizzi@outlook.com

Directores: Marcela Genero Bocco (UCLM), marcela.genero@uclm.es

Rodolfo Bertone (UNLP), pbertone@lidi.info.unlp.edu.ar

Fecha de exposición: 7 de julio de 2022.

Resumen

DepProMod (en inglés, *Deployment Process Model*) es un modelo de proceso de despliegue de sistemas de software que se desarrolló para realizar el despliegue de sistemas software desarrollados a medida, de manera sistematizada y controlada en las Pequeñas y Medianas Empresas (PyMES) de Argentina.

En este artículo se presenta una síntesis del proceso de investigación riguroso llevado a cabo en la tesis doctoral. En primer lugar, se enuncia la motivación y la problemática del despliegue de sistemas de software junto con el marco metodológico definido para el desarrollo de la tesis. En segundo lugar, se presenta una síntesis de los pasos realizados para la construcción del estado del arte que evidencia las necesidades del proceso de despliegue de sistemas de software que permitieron el diseño de DepProMod. A continuación, se presentan los elementos que componen el modelo, la forma de representación de estos y la manera de aplicación escalonada del modelo. Luego se plantean los estudios empíricos realizados para el refinamiento, así como también para la validación de DepProMod y las conclusiones obtenidas. Por último, se enuncian los principales aportes de la tesis doctoral, se esboza el trabajo futuro y las publicaciones obtenidas durante el desarrollo de la tesis.

Palabras clave: procesos de software; despliegue de sistemas de software; modelo de proceso; PyMES; estudios de casos; design science.

1. Introducción

En Argentina, la industria del software se compone mayoritariamente por PyMES,

representando casi el 80% del sector, lo que constituye un eslabón fundamental, en el sector, para el país y refuerza la necesidad de llevar adelante iniciativas que contribuyan con el desarrollo y mejora de competitividad de dichas empresas [1].

En el contexto internacional se refleja la misma realidad respecto a que las PyMES ocupan una gran porción de la industria del software [2]. El porcentaje de representatividad de las PyMES en el sector de la industria del software se mantiene desde las últimas dos décadas [3], [4].

Estas organizaciones descubrieron que es crucial mejorar sus procesos y métodos para ser competitivas en su negocio, aunque no cuentan con recursos y conocimiento suficiente para lograrlo. La única manera de lograr proyectos exitosos es mediante la definición, implementación y estabilización de los procesos de desarrollo [5].

El despliegue de sistemas de software es un proceso crucial del ciclo de vida de desarrollo de software porque que el sistema de software finalmente estará operativo para que el cliente pueda beneficiarse económicamente de su uso [6]. En el despliegue se presentan inconvenientes como por ejemplo la falta de componentes (externos), descargas incompletas, la infraestructura heterogénea e incompatible, incumplimiento de calendarios, retrabajo por falta de pocas capacidades y competencias tecnológicas, capacitación y preparación inadecuadas de los usuarios finales entre otros puntualmente en PyMES [7], [8].

En la actualidad, la aplicación más importante de la automatización en los procesos de software se presenta en las fases finales del desarrollo de software [9]. Si bien es cierto que existen nuevas técnicas / prácticas

como DevOps [10], [11] y Despliegue Continuo [12] en el contexto de metodologías ágiles. Estas técnicas y prácticas tienen como objetivo asegurar que el despliegue sea ágil y automatizado; éstas son utilizadas principalmente por grandes empresas, como Flickr, Netflix, Easy y Amazon, entre otras [9], [10]. Este tipo de empresas cuentan con los recursos humanos y la infraestructura para aplicarlas con éxito. Pero existe una gran cantidad de PyMES en Argentina que desarrollan sistemas de software que no utilizan metodologías ágiles, y además no cuentan con los recursos suficientes para afrontar el uso de estas técnicas emergentes.

En consecuencia, dada la relevancia del proceso de despliegue, los inconvenientes que conlleva un despliegue inadecuado más la necesidad de las PyMES de disponer de procesos estabilizados y controlados para la entrega de sistemas de software que satisfagan las necesidades de sus clientes surge el objetivo de esta tesis doctoral. Este objetivo es **“La definición y validación de un modelo de proceso para realizar el despliegue de sistemas software desarrollados a medida de manera sistematizada y controlada en PyMES de Argentina denominado DepProMod (en inglés, *Deployment Process Model*)”**. Este ha sido creado, refinado y evaluado, siguiendo el marco metodológico *Design Science* (Ciencia del Diseño) [13].

2. Estado del arte.

Para conocer el estado del arte, en primer lugar, se analiza la literatura existente sobre el proceso de despliegue de sistemas de software, a través de un mapeo sistemático de la literatura (en inglés, *Systematic Mapping Study* o SMS). Para el desarrollo del SMS se siguieron las directrices propuestas en [14], [15]. Los resultados del SMS confirmaron la existencia de dos modelos de procesos y una metodología que sirven como una guía para que las empresas de software lleven a cabo todo el proceso de despliegue. Estas propuestas cuentan con la limitación de que delegan en las organizaciones la responsabilidad de tomar decisiones sobre una serie de aspectos del despliegue. Estos

aspectos incluyen los artefactos, las técnicas, los métodos, las herramientas, y la definición de los roles, sólo incluyen las tareas. La delegación de estas decisiones ocasiona que estas propuestas resulten más difíciles de implementar en las PyMES, dado que requieren de procesos que sean más detallados o descriptivos para que les resulten más fácil de implementar [16].

A continuación, y para complementar el SMS, se realiza un estudio exploratorio para recolectar evidencia sobre la práctica actual del proceso de despliegue de sistemas de software en PyMES basado en una encuesta. Para el llevar a cabo el proceso de la encuesta se siguieron las directrices propuestas en [17]. Este estudio permitió confirmar necesidad de un modelo de proceso de despliegue de sistemas de software que ayude a las PyME, a realizar el despliegue de manera sistematizada mediante: a) la ejecución de actividades y tareas bien definidas, b) el uso de plantillas orientadoras, c) la asignación de roles específicos que cuenten con las competencias necesarias para ejecutar el despliegue y d) el uso de herramientas para automatizar algunas de las actividades del proceso, con el propósito de darle agilidad al proceso [18].

Por último, se revisan soluciones existentes en la literatura para el despliegue de sistemas de software sugeridas por expertos en Ingeniería de Software. Dentro de esta revisión se encuentran soluciones emergentes como DevOps y Despliegue Continuo, las cuales han sido adoptadas de manera exitosa por empresas multinacionales y/o grandes y esto se debe a que este tipo de empresas cuentan con los recursos necesarios para afrontar este tipo de inversión tecnológica, así como también de equipos de trabajo con un nivel de conocimientos actualizados y que no son aplicables en PyMES. Dentro de esta revisión de literatura sugerida por los expertos de Ingeniería de software también se realiza un análisis comparativo de estándares, prácticas y tecnologías emergentes que resuelven el despliegue de sistemas de software sugerida por expertos en Ingeniería de Software. Este análisis comparativo se focalizó en el Proceso Unificado (en inglés, *Rational Unified Process*

o RUP) [19], el estándar ISO IEC/IEEE 12207 [20], el estándar ISO/IEC/IEEE 24748-3 [21],

e ITIL 4 [22], [23] y permitió identificar el valor agregado que se propone en DepProMod.

Concretamente, la contribución de esta tesis doctoral se focaliza en el proceso de despliegue de sistemas de software proponiendo y validando un modelo de proceso para el despliegue de sistemas de software desarrollados a medida, denominado DepProMod, que servirá de guía para su realización de manera sistemática y controlada. Este modelo de proceso de despliegue de sistemas de software considera las necesidades evidenciadas en el SMS, en la encuesta sobre el estado actual de la práctica del proceso de despliegue en PyMES de Argentina y, además, contempla las necesidades detectadas en el análisis comparativo de las soluciones sugeridas por los expertos de Ingeniería de Software.

3. Definición de DepProMod.

DepProMod tiene un modelo de ciclo de vida que adopta los 5 grupos de procesos del PMBOK [24]: Planificación, Ejecución, Seguimiento y Control y Cierre por tratarse de un estándar reconocido a nivel internacional, así como también de aplicación en la industria del software en Argentina. Cada uno de estos procesos en DepProMod se denomina "subprocesos" [25], [26].

Para la definición de las actividades de DepProMod se consideraron un conjunto de procesos del estándar ISO/IEC/IEEE 12207 [20] por ser reconocido internacionalmente. Los procesos extraídos de este estándar son los procesos de gestión técnica: gestión de riesgos, gestión de configuración, gestión del proyecto, y otros procesos técnicos: verificación y validación. En DepProMod a estos procesos se los denomina actividades.

A nivel de "tareas", el modelo adopta un grupo de tareas propuestas en la metodología Metrica v3 [27] por considerarse una de las metodologías más completas por los elementos que la componen (fases, actividades, tareas, entradas, salidas, roles y herramientas) y se realizaron algunas adaptaciones por tratarse de

un modelo de despliegue para PyMES. Además, se consideraron una serie de actividades propuestas en el proceso técnico de "transición" de la norma ISO/IEC/IEEE 12207 [20] por su reconocimiento a nivel internacional, así como también las tareas que se realizan en las fases del proceso de Gestión técnica "*Deployment Management*" de ITIL4 [22].

DepProMod cuenta con un total de 16 actividades distribuidas en 5 subprocesos. En la Figura 1. se presentan las relaciones entre los subprocesos de DepProMod y se detallan las actividades de cada subproceso. Las relaciones que se presentan entre los subprocesos representan que las salidas de un subproceso son entradas en otro subproceso [28].

Además, DepProMod propone 6 roles para la ejecución del despliegue. En la Tabla 1. se detallan los roles junto con su acrónimo, nombre, descripción y competencias.

Otro de los elementos que define DepProMod, son las herramientas que se utilizarán en las tareas, que incluyen la gestión de la configuración, la instalación, la documentación, la gestión del plan de despliegue, la formación y las herramientas de diseño. Las herramientas se proponen en función de su utilidad, y las PYMES son libres de usar las herramientas que tienen.

Para la clasificación de los productos de DepProMod se utilizó parte de la clasificación definida en el modelo Competisoft [29], con algunas adaptaciones. La clasificación utilizada para DepProMod se resume en: productos del despliegue, productos del proyecto y otros productos. Los productos de despliegue son los que se generan como documentación en los subprocesos de DepProMod a los que denominaremos salidas, aunque en algunos casos se requieren como entradas y sufren alguna modificación.

Los productos del proyecto son la documentación provista por el cliente y los denominaremos entradas. Los otros productos que se utilizan en el proceso de despliegue son provistos por el cliente, pero no forman parte de la documentación del proyecto.

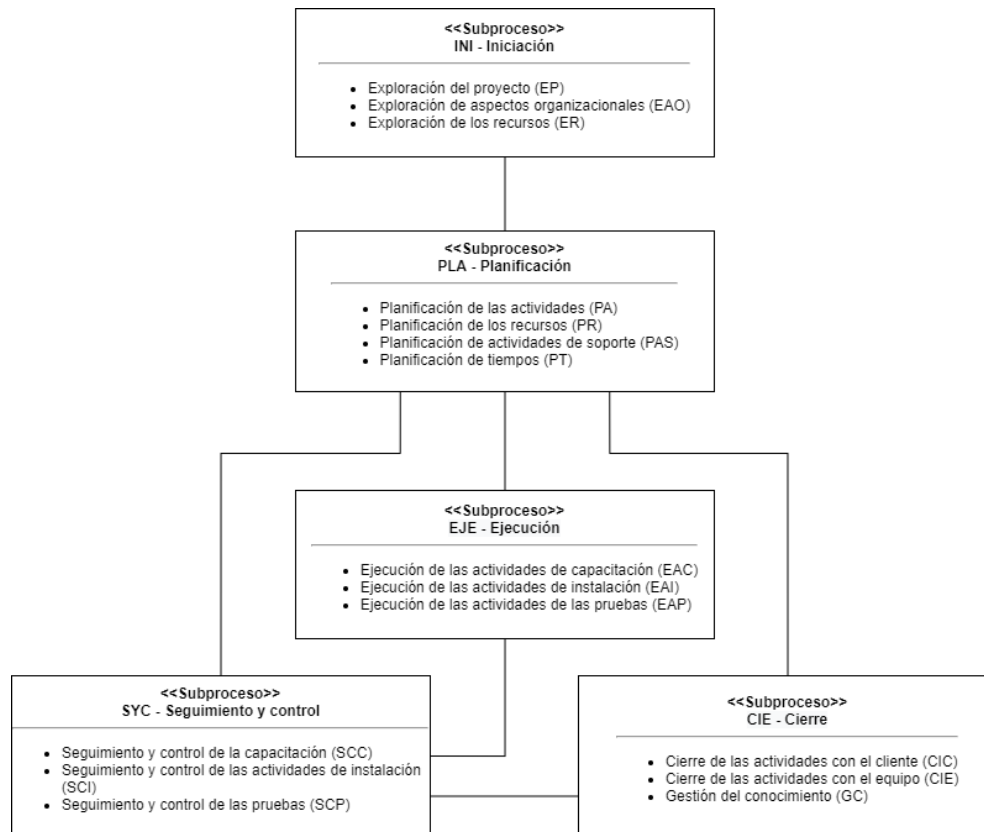


Figura 1. Relación entre los subprocesos de DepProMod.

Tabla 1. Clasificación de roles de DepProMod.

Acrónimo	Nombre del rol	Descripción y competencias
LP	Líder de proyecto.	Es el responsable del proyecto por parte del cliente que requiere el proceso de despliegue del sistema de software. Puede tratarse de un cliente externo o interno. Tiene conocimiento para solicitar cambios.
USC	Usuario clave	Persona o personas que conocen el negocio o procesos, validan las pruebas, validan el plan de capacitación y aceptan el cierre del despliegue.
US	Usuario.	Persona o personas que operan o interactúan directamente con el Sistema de Software.
RD	Responsable de despliegue.	Es el encargado de la realización del despliegue y del cumplimiento de sus objetivos. Tiene conocimiento y experiencia en la gestión de proyectos, toma de decisiones, técnicas de revisión, desarrollo de software y técnicas de estimación y costos.
Involucrados: INS	Instalador.	Es responsable de realizar las actividades vinculadas a la preparación del lugar de instalación, instalación, la carga y/o migración de los datos y las pruebas. Tiene conocimiento y experiencia en infraestructuras, desarrollo de software y gestión de configuración.
CAP	Capacitador.	Es el responsable de llevar a cabo las actividades de capacitación, así como de preparar los esquemas de capacitación (plataformas, estrategias y modalidades). Tiene conocimiento y experiencia en desarrollar actividades de capacitación y gestión de configuración.

Para representar cada uno de los elementos de DepProMod, en general, se emplea el patrón para la descripción de procesos propuesto en el modelo Competisoft [29] dado que es un modelo de mejora de procesos para PyMES de software de Iberoamérica, introduciendo algunas modificaciones para para cubrir las necesidades específicas de DepProMod. Para la representación del flujo de trabajo entre las actividades de cada subproceso se utiliza un diagrama de actividad de alto nivel para representar el flujo de una actividad a otra [19]. Para complementar el diagrama de actividad de alto nivel se presenta un diagrama de actividad para cada una de las actividades del subproceso que permite visualizar el flujo de trabajo entre las tareas que componen una actividad y los roles que realizan las tareas. Además, se proponen una serie de plantillas para estructurar la información de las salidas, 27 plantillas en total para los 5 subprocesos que componen el modelo.

Para que una PyME pueda implementar el modelo paso a paso, y en la medida que logre estabilizar el proceso en un determinado nivel de aplicación y que alcance los conocimientos necesarios para escalarlo al siguiente nivel, se define una arquitectura por niveles adoptando tres de los niveles de capacidad del estándar CMMI-DEV [30]. Estos niveles son: nivel 1 = Realizado (representado en color amarillo), nivel 2 = Gestionado (representado en color azul) y nivel 3 = Definido (representado en color verde). Estos niveles fueron analizados y definidos a nivel de granularidad de las tareas consideradas en el modelo.

4. Refinamiento de DepProMod.

El primer estudio es un estudio de caso (EC1) realizado para examinar la viabilidad de la aplicación de la versión preliminar de DepProMod en un entorno real, una PyME Mediana de tramo 1(55 empleados) [31] desarrolladora de software de Argentina. Para la realización de los estudios de casos se siguieron las directrices propuestas en [32], [33]. En el segundo, se mantuvieron entrevistas con dos expertos de la industria del software con el propósito presentar

DepProMod para conocer su impresión sobre la definición de cada elemento del modelo y su utilidad. En la Tabla 2 se muestra una síntesis de los estudios realizados y los hallazgos encontrados que permitieron refinar y completar DepProMod [28], [34].

5. Validación de DepProMod.

Se presentan tres estudios de casos (EC2, EC3, EC4) realizados en PyMES desarrolladoras de software de Argentina con el propósito de validar DepProMod para detectar sus fortalezas y debilidades. Para la realización de los estudios de casos se siguieron las directrices propuestas en [32], [33]. En estos estudios se evaluaron las tres variables de percepción tomadas del *Technology Acceptance Model* (TAM) [35] son: “Facilidad de Uso Percibida” (en inglés, *Perceived Ease of Use* o PEOU), “Utilidad Percibida” (en inglés, *Perceived Usefulness* o PU) e “Intención de Uso” (en inglés, *Intention to Use* o ITU).

En la Tabla 3 se presentan las características de los estudios de casos.

En los estudios de casos participaron en total 11 profesionales de la industria del software, 2 en EC2, 5 en EC3 y 4 en EC4 respectivamente. A continuación, se detallan las principales conclusiones obtenidas:

- Se constató, que los profesionales que participaron en los estudios de casos en su mayoría con más de 10 años de experiencia en la industria del software consideran que el modelo impacta en la mejora de la calidad de los despliegues de los sistemas de software de sus PyMES.
- Dentro de las fortalezas de DepProMod se encuentra su la flexibilidad que aporta al poder aplicarse por niveles, permitiendo a las PyMES llevar a cabo sus despliegues de sistemas de software desarrollados a medida de manera sistematizada y controlada, adecuándose a sus recursos y necesidades.
- Los usuarios de DepProMod, en los tres estudios de casos resaltan que las fases/actividades/tareas que ofrece el modelo les sirva de guía para realizar el proceso de despliegue y de esta manera no

Tabla 2. Síntesis de los estudios empíricos realizados y los hallazgos encontrados (Refinamiento de DepProMod).

Estudio	Propósito	Descripción del estudio	Hallazgos para refinar y completar DepProMod
Estudio de caso (EC1).	Evaluar la aplicabilidad de la versión preliminar de DepProMod con el propósito de refinarlo y completarlo (si fuese necesario). Además, analizar los requerimientos de información para el proceso de despliegue del sistema de software y así avanzar hacia el diseño de las plantillas necesarias para DepProMod.	EC1 involucró a una PyME Mediana de tramo 1 [31] de 55 empleados, ubicada en Argentina, que ofrece productos y servicios de consultoría. Se trabajó con la documentación del proceso de despliegue del módulo “Creación de empresas” de un sistema de gestión para agencias de publicidad para Latinoamérica.	Los requerimientos de información de la documentación del EC1 permitieron avanzar con el diseño de las plantillas de DepProMod.
Entrevistas.	Presentar DepProMod a los expertos para conocer su opinión sobre la definición de cada elemento del modelo y su utilidad y refinarlo en caso de ser necesario.	Experto1. Cuenta con 25 años de experiencia en la industria del software y se mantuvieron 5 sesiones virtuales de una hora y media cada una.	Se renombraron roles. A partir del feedback del experto respecto a la denominación de algunas entradas y salidas del Subproceso INI - Iniciación, se modificaron algunas de ellas, se eliminó una tarea de una actividad y además se introdujo un cambio en una plantilla del mismo Subproceso. En relación con Subproceso PLA- Planificación, se unificaron dos tareas, se incorporó una acción a una tarea y, además, se eliminó la estimación de costos del despliegue dentro del alcance del modelo. Con relación al Subproceso EJE – Ejecución, se unificaron y renombraron plantillas, se incorporó una plantilla relacionada a la capacitación de usuarios finales y además se replanteó el alcance de algunas tareas relacionadas a las pruebas. Respecto al Subproceso SYC – Seguimiento y control, se unificaron y se simplificaron plantillas relacionadas a los riesgos del despliegue y al seguimiento. En el Subproceso CIE – Cierre, se modificó una tarea.
		Experto2. Cuenta con 17 años en la industria del software y se mantuvieron 4 sesiones virtuales en total, las primeras tres de una hora y la cuarta sesión de dos horas.	Se renombró un elemento del modelo, originalmente denominado “infraestructura” y finalmente se lo denominó “recursos”. Se realizó un cambio en el alcance de una tarea relacionada a la capacitación del Subproceso EJE – Ejecución. Además, se introdujo un cambio en dos plantillas relacionadas a la capacitación.

Tabla 3. Contexto, caso y unidad de análisis de los estudios de casos (EC2, EC3, EC4).

Estudios de casos	Contexto	Caso	Unidad de análisis
EC2	Micro PyME de 3 empleados, ubicada en Argentina, que ofrece productos y servicios de consultoría.	Despliegue de un sistema web del tipo “Tienda Virtual”, dedicado al comercio en el rubro de los repuestos y accesorios para motos y afines.	Aplicación del nivel 1 de DepProMod.
EC3	PyME Mediana de tramo 1 , de 45 empleados aproximadamente, ubicada en Argentina, que ofrece productos y servicios de consultoría de desarrollo web y consultoría de marketing digital.	Despliegue de un sistema web del tipo catálogo de software de mantenimiento que ofrece el cliente final para diferentes industrias.	Aplicación del nivel 2 de DepProMod.
EC4	PyME Mediana de tramo 2 de aproximadamente 430 empleados, ubicada en Argentina. La empresa desarrolla sistemas de software a medida para clientes de diversos rubros.	Despliegue de funcionalidades de un Portal de Recursos Humanos realizado para una entidad bancaria de Argentina.	Aplicación del nivel 3 de DepProMod.

omitir ninguna tarea, así como también la ventaja de la libre elección de las herramientas para utilizar en las tareas.

- Además, los usuarios encontraron adecuados los roles propuestos en DepProMod y que tuvieron que desempeñar y que no han tenido inconveniente con el rol que han desempeñado en DepProMod.
- Si bien los participantes de los estudios resaltaron la importancia de la gestión de los despliegues, percibieron de manera desfavorable las entradas, salidas y las plantillas por el esfuerzo que conlleva la documentación siendo más acentuado en la micro PyME.

6. Conclusiones y trabajos futuros.

En esta sección se presentan los principales aportes de tesis doctoral, se esboza el trabajo futuro y las publicaciones obtenidas

Los principales aportes de esta tesis son:

• DepProMod.

Un modelo de proceso de despliegue de sistemas de software que le permita a las PyMES realizar el despliegue de manera sistematizada y controlada presentado en la sección 3. DepProMod propone fases, actividades y tareas, roles, entradas y salidas, plantillas y recursos y además cuenta con una estructura de tres niveles de aplicación para facilitar su implantación en PyMES.

DepProMod se definió y se evaluó por diferentes profesionales de la industria del software para asegurar que fuera viable y útil. Este modelo también es útil para los investigadores interesados en ampliar la investigación presentada en esta tesis doctoral. En el contexto de la academia, sería útil incorporarlo en los programas de las asignaturas para que los profesores puedan explicarlo como una solución alternativa para resolver los inconvenientes del proceso de despliegue de sistemas de software.

• Conocimiento generado a través de los estudios empíricos.

El conocimiento creado a través de varios estudios empíricos realizados durante el desarrollo de la tesis doctoral. El SMS presentado en la sección 2 junto con la encuesta sobre la práctica actual del despliegue de sistemas de software en PyMES de Argentina. El estudio de caso (EC1) llevado a cabo en una PyME desarrolladora de software de Argentina que permitió examinar la viabilidad de aplicación de la versión preliminar de DepProMod para refinarlo y completarlo que se describe en la sección 4. Las entrevistas realizadas a dos expertos de la industria del software (EE) para conocer su opinión sobre los elementos del modelo y su utilidad descriptas en la sección 4. Los estudios de casos (EC2, EC3 y EC4) realizados en PyMES desarrolladoras de software de

Argentina que permitieron evaluar la percepción de los usuarios de DepProMod medida a través de tres variables, la “Facilidad de Uso Percibida”, la “Utilidad Percibida” y la “Intención de Uso” y que además permitieron identificar las fortalezas y debilidades de DepProMod informadas por los usuarios del modelo, presentados en la sección 5.

Los resultados de la investigación presentados en esta tesis doctoral han abierto una serie de trabajos futuros, tanto desde el punto de vista de la investigación como desde el punto de vista práctico, para permitir la transferencia de DepProMod hacia los profesionales de la industria del software. A continuación, se muestran algunas posibles líneas de trabajo futuras:

- **Refinar DepProMod.**

La evaluación de DepProMod mediante los estudios de casos (EC2, EC3 y EC4) han demostrado que los valores de percepción de las tres variables “Facilidad de Uso Percibida”, “Utilidad Percibida” e “Intención de Uso” presentadas por los usuarios deben ser mejoradas en especial las relacionadas con las “plantillas”. Estos estudios, además confirmaron las fortalezas y debilidades de DepProMod. Dentro de las debilidades más significativas del modelo se evidenció la necesidad de revisión, simplificación y en algún caso, eliminación de las plantillas como futura mejora de DepProMod.

- **Continuar la investigación empírica.**

De esta línea de trabajo se desprenden las siguientes sub líneas de investigación:

- Evaluar el modelo en PyMES de otros sectores industriales que desarrollen software para su propio uso.
- Incorporar en los estudios de caso la recolección de datos cuantitativos como por ejemplo métricas que se desprendan del proceso de despliegue.
- Realizar más estudios de caso que sean llevados a cabo por otros investigadores.
- Extender la evaluación de DepProMod a un contexto internacional, más allá de su aplicación en PyMEs de Argentina.

- **Extender el alcance de DepProMod.**

De esta línea de trabajo se desprenden las

siguientes sub líneas de investigación:

- Evaluar la viabilidad de poder aplicar el modelo en otros tipos de empresas, como por ejemplo empresas emergentes (*Startups*) o empresas de mayor tamaño para recolectar feedback por parte de los usuarios y contrastarlo con los resultados obtenidos en PyMES.
- Evaluar la viabilidad de poder aplicar DepProMod en el despliegue de otro tipo de sistemas de software, por ejemplo, los sistemas de software configurables.

Las publicaciones producidas durante el desarrollo de esta tesis doctoral son:

- Panizzi, M., Genero, M. and Bertone, R. (2023). Refining a Software System Deployment Process Model through Empirical Studies. *Journal of Computer Science and Technology* 23(1).
- Panizzi M., Genero M. y Bertone R. (2021). Refining a software system deployment process model: A case study. *XXVII Congreso Argentino de Ciencias de la Computación (CACIC '2021)*, pp. 439–448. Universidad Nacional de Salta, Salta.
- Panizzi M., Genero M. y Bertone R. (2021). Encuesta para analizar las necesidades con respecto al proceso de despliegue de las PyMES en Argentina. *XXIV Congreso Iberoamericano de Ingeniería de Software (CibSE '2021)*. Universidad de Costa Rica, San José Costa Rica.
- Panizzi M.; Genero M. y Bertone R. (2020). Software system deployment process: A systematic mapping study. Brasil. Curitiba. 2020. *XXIII Congreso Iberoamericano en Ingeniería de Software (CibSE '2020)*, pp. 138-151. Pontificia Universidade Católica do Paraná.
- Panizzi M. (2019). Propuesta de un modelo de proceso de implantación de sistemas informáticos (MoProIMP). *XXII Congreso Iberoamericano en Ingeniería de Software (CibSE '2019)*, pp. 617-624. Universidad de La Habana. ISBN 978-1-5108-8795-4.
- Panizzi M., Bertone R. y Hossian A. (2019) Proposal for a model of a computer systems implantation process (MoProIMP). Pesado P., Aciti C. (eds).

Computer Science – CACIC '2018. C. Communications in Computer and Information Science, vol 995, pp. 157-170. Springer.

- Además, se está trabajando en un artículo con la validación completa de DepProMod, es decir presentando los estudios de casos realizados (EC2, EC3 y EC4). Este artículo se enviará a una revista indexada en el *Journal Citation Report*.

Publicaciones sobre temas relacionados con esta tesis, que corresponden a los resultados de dos estudiantes de la Maestría en Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Buenos Aires. El listado de publicaciones se detalla a continuación:

- Vázquez P., Panizzi M. y Bertone R. (2021). Refinamiento de métricas para proceso de despliegue de sistemas de software: estudio de caso. 8 va. Conferencia Ibero Americana Computación Aplicada (CIACA '2021), pp. 188-192.
- Vázquez P., Panizzi M. y Bertone R. (2021). Métricas para el proceso de despliegue de sistemas software: un mapeo sistemático. *Revista Desarrollo e Innovación en Ingeniería*, Edgar Serna M., Edición 6 Vol. II, pp. 386-396, Medellín – Antioquia.
- Vázquez P.; Panizzi M. y Bertone R. (2019). Estimación del esfuerzo del proceso de implantación de software basada en el método de puntos de caso de uso / Estimating the effort of the software implantation process based on the use case points method. *Brazilian Journal of Development*. Curitiba: Brazilian Journals Publicação de Periódicos Ltda. (CNPJ 32.432.868/0001-57. 2019 vol.5 n°1. pp.1809 - 1822.
- Ortiz F., Dávila M.; Panizzi M.; Bertone R. (2019). State of the art determination of risk management in the implantation process of computing systems. *Advances in Emerging Trends and Technologies*. Estados Unidos: Springer. 2019. pp. 23 - 31.
- Vázquez P.; Panizzi M.; Bertone R. (2018).

Estimación del esfuerzo del proceso de implantación de software basada en el método de puntos de caso de uso. 6to Congreso Nacional de Ingeniería en Informática/Sistemas de Información (CoNaIISI '2018), pp. 761-768. Argentina. Mar del Plata.

Referencias

- [1] Cámara de la industria Argentina del software. (2019). *Reporte anual del sector de software y servicios informáticos de la República Argentina del año 2018*. <https://www.cessi.org.ar/opssi-reportes-949/index.html>.
- [2] Abushama H. (2016). PAM-SMEs: process assessment method for small to medium enterprises. *Journal of Software: Evolution and Process*, 28(8), pp. 689-711.
- [3] Felderer M., Ramler R. (2016). Risk orientation in software testing processes of small and medium enterprises: an exploratory and comparative study. *Software Quality Journal*, 24, pp. 519-548.
- [4] Sharma P., Sangal A. (2019). Building a hierarchical structure model of enablers that affect software process improvement in software SMEs-A mixed method approach. *Computer Standards & Interfaces*, 66, pp. 1-23.
- [5] Ianzen A., Mauda E., Paludo M.A., Reinehr S., Malucelli A. (2013). A. Software process improvement in a financial organization: An action research approach. *Computer Standard & Interfaces*, 36(1), pp. 54-65.
- [6] Subramanian N. (2017). The software deployment process and automation. *CrossTalk*, 30(2), pp. 28-34.
- [7] Tyndall J. (2012). Building an effective software deployment process. *SIGUCCS '12: ACM SIGUCCS Annual Conference*, pp.109-114.
- [8] Carrizo D., Sanchez L. (2017). Benchmarking to adopt an Asap-based methodological. Guideline for software systems deployment. *30th International Business Information Management Association Conference, (IBIMA '17): Sustainable Economic de Economic development, Innovation Management, and Global Growth*, pp. 183-193.
- [9] Fuggetta A., Di Nitto E. (2014). Software process. *36 th International Conference on Software Engineering - Future of Software Engineering (FOSE '14)*, pp. 1-12.
- [10] Erich F., Amrit C., Daneva M. (2017). A qualitative study of DevOps usage in practice. *Software: Evolution and Process*, 29(6), pp. 1-20. <https://doi.org/10.1002/smr.1885>.

- [11] Bass L., Weber I., Zhu L. (2015). *DevOps: A software architect's perspective*. SEI series in software engineering.
- [12] Agile Alliance. (2022). *Agile Alliance. Agile Glossary*.
<https://www.agilealliance.org/agile101/agile-glossary/>
- [13] Wieringa R. (2014). *Design science methodology for information systems and software engineering*. Springer.
- [14] Kitchenham B., Budgen D., Brereton P. (2015). *Evidence-based software engineering and systematic reviews*. New York: Chapman and Hall/CRC.
- [15] Petersen K., Feldt R., Mujtaba S., Mattsson M. (2008). Systematic mapping studies in software engineering. *20th International Conference on Evaluation and Assessment in Software Engineering (EASE '08)*, pp. 68-77.
- [16] Panizzi, M., Genero, M. and Bertone, R. (2020). Software system deployment process: A systematic mapping study. In *Proceedings of the XXIII Iberoamerican Conference on Software Engineering, CIbSE '20*, pages 138-151, Curitiba, Paraná, Brazil.
- [17] Molléri J., Petersen K., Mendes E. (2020). An empirically evaluated checklist for surveys in software engineering. *Information and Software Technology*, 119.
- [18] Panizzi, M., Genero, M. y Bertone, R. (2021). Encuesta para analizar las necesidades con respecto al proceso de despliegue de las PyMES en Argentina. In *Proceedings of the XXIV Iberoamerican Conference on Software Engineering, CIbSE '21*, pages 84-91, San José, Costa Rica.
- [19] Jacobson I., Booch G., Rumbaugh J. (1999). *The Unified software development process*. Addison Wesley - Professional.
- [20] ISO/IEC/IEEE. (2017). *ISO/IEC/IEEE 12207. Systems and software engineering — Software life cycle processes*.
- [21] ISO/IEC/IEEE FDIS. (2020). 24748-3:2020. *Systems and software engineering — Life cycle management — Part 3: Guidelines for the application of ISO/IEC/IEEE 12207 (Software life cycle processes)*.
- [22] Berclaz D. (2019). *Deployment management in ITIL4*. Recuperado el 23 de 05 de 2020, de <https://www.apwide.com/deployment-management-in-til4/>.
- [23] Anand A., Buffington P., Buchanan I., Fok T. (2020). *A practical guide to ITIL 4 in an age of agile*, www.atlassian.com/whitepapers/itil4#
- [24] Institute Project Management. (2013). *A Guide to the Project Management Body of Knowledge (PMBOK® Guide) - Fifth Edition*. USA: Project Management Institute Publisher.
- [25] Panizzi M., Bertone R. y Hossian A. (2019) Proposal for a model of a computer systems implantation process (MoProIMP). Pesado P., Aciti C. (eds). *Computer Science – CACIC '2018. C. Communications in Computer and Information Science*, vol 995, pp. 157-170. Springer.
- [26] Panizzi M. (2019). Propuesta de un modelo de proceso de implantación de sistemas informáticos (MoProIMP). Simposio Doctoral del XXII Congreso Iberoamericano en Ingeniería de Software (CIbSE '2019), pp. 617-624. Universidad de La Habana.
- [27] Portal de administración electrónica. Gobierno de España. (2001). Métrica versión.3. Obtenido:<https://administracionelectronica.gob.es/pae/Home>.
- [28] Panizzi, M., Genero, M. and Bertone, R. (2023). Refining a Software System Deployment Process Model through Empirical Studies. *Journal of Computer Science and Technology* 23(1). https://journal.info.unlp.edu.ar/JCST/inpress_articles
- [29] Competisoft. (2008). Mejora de Procesos para Fomentar la Competitividad de la Pequeña y Mediana. Industria del Software de Iberoamérica. <https://alarcos.esi.uclm.es/competisoft/web/completo/index.htm>.
- [30] CMMI Institute. (2010). *CMMI Model versión 2.0*. Pittsburgh.
<https://stage.cmmiinstitute.com/cmmi>.
- [31] Ministerio de Desarrollo Productivo. (2018). *Nuevas categorías para ser PyMES*. www.argentina.gob.ar/noticias/nuevas-categorias-para-ser-pyme.
- [32] Runeson P., Höst M., Rainer A., Regnell B. (2012). *Case study research in software engineering: Guidelines and examples*. Hoboken: Wiley.
- [33] Runeson P. Höst M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14(2), pp. 131-164.
- [34] Panizzi, M., Genero, M., Bertone, R. (2021). Refining a Software System Deployment Process Model: A Case Study. In *Proceedings of the 27th Argentine Congress of Computer Science*, in CACIC '21, pp. 439-448, 2021.
- [35] Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS Quarterly*, 13(3):319-340.

Tesis Doctoral
Métricas de calidad para validar los conjuntos de datos
abiertos públicos gubernamentales

Autora: Roxana Martínez

Directora: Rocío Rodríguez
Co-Directora: Claudia Pons
Asesor Científico: Pablo Vera

Doctora en Ciencias Informáticas
Facultad de Informática.
Universidad Nacional de La Plata (UNLP)
Calle 50 &, Av. 120, La Plata, Provincia de Buenos Aires, Argentina

Fecha de exposición: 29 de marzo de 2022. Publicación: 2022

Jurado de Tesis Doctoral: Dr. Mario Piattini (España);
Dr. Pablo Fillottrani (Argentina); Dra. Elsa Estévez (Argentina).

ing.roxana.martinez@gmail.com; maria.martinez@info.unlp.edu.ar

RESUMEN

En estos tiempos, los datos son un recurso indispensable para cualquier actividad de gestión pública, por lo que es necesario mantenerlos actualizados, claros y completos. Este trabajo se enfoca en el contexto de Gobierno Abierto en aspectos del tratamiento del dato público abierto que disponibilizan las entidades estatales.

Con el fin de identificar mejoras en los aspectos de calidad de los datasets abiertos, esta tesis plantea métricas críticas y no críticas para su análisis y validación de contenido, por lo que, como parte de la propuesta, se presenta un prototipo de desarrollo propio, llamado HEVDA (Herramienta de Validación de Datos Abiertos). A modo de caso de estudio, se extrae una muestra de datasets públicos estatales que son validados con HEVDA, para obtener un análisis sobre las mediciones utilizadas y realizar así, un estudio cuantitativo sobre los resultados arrojados. Esta herramienta de validación permite detectar en forma sencilla, las falencias y

errores en las fuentes de datos abiertas que podrían complicar la interoperabilidad para su utilización en diversos orígenes de bases de datos y softwares externos de otros organismos. Para evaluar la calidad de datos es necesario tener en cuenta determinadas características en el conjunto de datos analizados, por lo que se realiza un relevamiento detallado de los aspectos más notables en cuestiones de calidad de datos sobre criterios estándares de norma ISO/IEC 25012 [1], estándares universales de calidad de datos [2], dimensiones de la calidad de los datos [3], trabajos relevados y estudios realizados [4] en esta temática. En el estudio desarrollado, se puede analizar si es factible definir métricas de calidad de datos públicos gubernamentales en un formato abierto para efectuar un análisis cuantitativo a través de una herramienta amigable y sencilla.

Palabras clave: Datos Abiertos, Datos Públicos, Gobierno Abierto, Métricas de Calidad de Datos.

1. CONTEXTO

Actualmente se utilizan las TIC (Tecnologías de la Información y la Comunicación) para la colaboración abierta en cuestiones asociadas a la resolución de diversos problemas o temas ciudadanos, que son llevados a cabo a través de sitios web gubernamentales, el enfoque principal se centra en la colaboración y participación de los ciudadanos y organismos públicos. “Esa composición la transforma en un instrumento único para impulsar la cooperación horizontal, el apoyo a la elaboración de políticas de gobierno digital centradas en el ciudadano, la formación de los funcionarios públicos, fomentar el conocimiento de aspectos claves de la construcción de una transformación digital, y el intercambio de soluciones y expertos entre los países” [5].

El Estado Nacional Argentino se encuentra promoviendo la participación ciudadana digital a través de herramientas de software en forma on-line, esta idea parte del concepto de Gobierno Abierto en la que la calidad de la información compartida por las entidades gubernamentales permite “promover y facilitar su uso siendo un instrumento que apoya el cumplimiento de los tres pilares del gobierno abierto: Transparencia, Participación y Colaboración” [6]. Si bien, el objetivo principal es la colaboración, existe otro punto a destacar que es el tratamiento de los datos colaborados, los cuales implican que, al ser compartidos, el gobierno es abierto y “permite una mayor transparencia, ofrece servicios públicos más eficientes, y alienta un mayor uso público y comercial, en la reutilización de la información del gobierno. Algunos gobiernos incluso han creado catálogos o portales (como data.gov) para facilitar al público la búsqueda y el uso de esta información” [7]. La gestión colaborativa permite la administración de los datos públicos compartidos a través de plataformas informáticas. Esta apertura de datos e información a la comunidad fortalece a la eficiencia en la gestión y optimización de las solicitudes de los ciudadanos.

El concepto de la apertura de datos

públicos permite que la información de las entidades gubernamentales se presente a los ciudadanos en un formato abierto para que sea reutilizado por otros, es decir, aplicar el concepto de RISP (Reutilización de la Información del Sector Público). Si bien el Estado Nacional administra grandes cantidades de datos públicos que son propiedad de todos los ciudadanos, y puede abrir estos datos y facilitar su acceso, es importante aclarar que esto ocurrirá siempre y cuando no exponga ningún tipo de información confidencial o privada, por lo que es necesario un adecuado análisis de los conjuntos de datos que serán considerados públicos y abiertos.

Se considera que “resulta necesario aumentar la calidad de los servicios provistos por el Estado incorporando TIC (Tecnologías de la Información y de las Comunicaciones), simplificando procedimientos, propiciando reingenierías de procesos y ofreciendo al ciudadano la posibilidad de mejorar el acceso por medios electrónicos a información personalizada, coherente e integral” [8].

Según el Decreto 434/2016, la estrategia Nacional se estructura en cinco ejes: a) Plan de Tecnología y Gobierno Digital; b) Gestión Integral de los Recursos Humanos; c) Gestión por los Resultados y Compromisos Públicos; d) Gobierno Abierto e Innovación Pública; y e) Estrategia País Digital. De estos cinco ejes mencionados, existen dos que priorizan el uso de las TIC, como ser: a) Plan de Tecnología y Gobierno Digital: los cuales pretenden incorporar infraestructura tecnológica y redes que son la condición necesaria para agregar nuevos modelos de gestión basados en la tecnología de Big Data; b) Gobierno Abierto e Innovación pública: el cual no sólo se complementa con el anterior eje fundamental, sino que, además, permite la evaluación y control de las políticas públicas del Estado Nacional. Es decir que, la administración de grandes volúmenes de información es un recurso valioso y necesario. Por lo que, partiendo de esta premisa, hay una gran necesidad de gestionar esta estructura mediante la utilización e incorporación de nuevas herramientas de análisis de datos.

2. DEFINICIÓN DEL PROBLEMA

El desafío del Gobierno es brindar a los ciudadanos información de manera eficiente y transparente. En este contexto, los nuevos retos que se plantean pueden ser considerados como una oportunidad para replantear las metodologías de diseñar e implementar políticas públicas e impulsar un Estado con funcionarios colaboradores de la ciudadanía. Todo esto conlleva a la necesidad de nuevas formas de liderazgo colaborativo, y un nuevo paradigma en el tratamiento de la información abierta y pública. Este escenario, necesita promover la innovación como una “nueva actitud” y hacer frente a los nuevos desafíos de la ciudadanía digital. Es importante identificar que los pilares para la solución de esta problemática son 3 conceptos fundamentales: a) Comprender el contexto de Gobierno Abierto: apoyo del Estado Nacional y Ciudadano; b) Conjunto de datos públicos: organismos estatales deberían ofrecer una mayor cantidad de fuentes de datos de diversos temas gubernamentales con criterios preestablecidos; c) Datos abiertos: mediante la apertura de los diversos datasets en los portales de datos, utilizados como insumo fundamental de información y servicios.

Según [9] sostiene que algunos problemas que surgen en el Estado, son: a) “Procesos mal definidos”; b) “Ritmo y la magnitud del cambio estatal”; c) “Ausencia de capacidad para el aprendizaje organizacional”; d) “La resistencia del público al cambio”; e) “Ausencia de recursos”: falta de apoyo financiero o de carencia de conocimientos para soporte a las iniciativas innovadoras; f) “Obstáculos técnicos”: Ausencia de soluciones tecnológicas para el problema en cuestión a analizar. Es decir, es necesario implementar una serie de criterios o aspectos consensuados y sostenibles en el tiempo para mejorar la calidad de los datos públicos abiertos que se ofrecen, lo que brindará mayor poder de análisis de la información dada, como así también, nuevos escenarios sobre posibles mitigaciones de inconvenientes ciudadanos.

A modo de resumen, se puede decir que

mantener la calidad y la apertura de los datos públicos, ayuda a los gobiernos y a los diferentes actores de la sociedad civil a tomar mejores decisiones, ya que tienen una visión e información de la realidad. También permite monitorear procesos, construir indicadores y hacer transparente la forma en que se gasta el dinero público [10].

Uno de los puntos a considerar en los portales de datos abiertos es que la disponibilidad de los datos no necesariamente coincide con que tengan calidad, lamentablemente, hoy sigue siendo una dificultad y es un gran desafío para las políticas públicas. El análisis de muchos de los conjuntos de datos públicos representa un problema crucial, ya que está disperso, no estandarizado y en muchos casos desactualizado. El problema en este contexto también tiene que ver con la actualización de la infraestructura técnica para los datos abiertos, ya que las prácticas de gestión de datos son débiles e inconsistentes y cambian con demasiada frecuencia. Los gobiernos deberían trabajar en la transformación técnica y organizacional de los datos abiertos, es decir, invertir en herramientas de hardware y software, como así también en estándares técnicos, capacitación, transformación organizacional y procesos de toma de decisiones para dar un soporte adecuado a la gestión de datos [11]. Como se mencionó anteriormente, aún existen algunas barreras que van desde la calidad insuficiente de los datos publicados hasta la falta de mantenimiento de los portales donde se publican [12]. Varias de estas investigaciones [13], están orientadas a la apertura de un dataset, pero no a un análisis de los problemas que puedan tener los datos abiertos que están implementados y publicados en los portales gubernamentales.

3. OBJETIVO E HIPÓTESIS

El objetivo principal de la presente tesis es proponer una serie de métricas de calidad para validar los conjuntos de datos abiertos públicos gubernamentales.

Objetivos específicos: Para llevar a cabo dicho objetivo se plantean los siguientes objetivos específicos: Estudiar el alcance de los conceptos implicados en este contexto de estudio; Analizar los trabajos relacionados en la temática de Gobierno Abierto relativos al tratamiento de datos abiertos y públicos; Relevar los trabajos relacionados en cuanto a los aspectos de calidad de datos abiertos y públicos; Analizar los criterios de los portales de datos abiertos para efectuar la publicación de éstos; Relevar los distintos tipos y formatos disponibles de datos abiertos que existen en los dataset actuales de los sitios más relevantes y gubernamentales de Argentina; Analizar las falencias de los datasets disponibles en los sitios de Gobierno; Establecer criterios estándar de calidad de datos; Proponer métricas críticas y no críticas de calidad de datos abiertos para validar diversos datasets; Desarrollar una herramienta de validación de datasets gubernamentales; Definir y recolectar la muestra de datasets para ser testeados por la herramienta propuesta como validadora de calidad; Definir un análisis cuantitativo en la medición de las métricas propuestas a través de la herramienta desarrollada; Analizar los resultados arrojados con la herramienta programada para las métricas de calidad propuestas; Validar las métricas sugeridas en base a criterios estándares.

Hipótesis: La hipótesis de esta tesis es la siguiente: *Es posible definir métricas de calidad de datos públicos gubernamentales en un formato abierto para efectuar un análisis cuantitativo a través de una herramienta amigable y sencilla.*

4. ENFOQUE METODOLÓGICO

Para el proceso metodológico de investigación de la tesis doctoral, se define un proceso sistemático cualitativo, que implementa una forma evolutiva incremental en cada una de las etapas involucradas, siendo éstas:

1. *Identificación del Problema:* Definición del problema a tratar;

2. *Revisión Teórica:* Búsqueda, selección, recopilación, revisión y análisis del marco teórico y los trabajos relacionados pertinentes al contexto del tema de la tesis;

3. *Recolección de Datos:* Mediante el estudio de la revisión teórica, se definió un alcance de muestra y se extrajeron datasets de los distintos portales de datos abiertos gubernamentales en Argentina;

4. *Clasificación y Análisis de Datos:* Se realiza una segmentación de los datasets de los portales abiertos y se analizan los resultados en aspectos de métricas de calidad de los datos.

5. *Formulación de la Solución:* Se definió el alcance de la propuesta de métricas críticas y no críticas en base al estudio general realizado, identificando niveles de criticidad para éstas.

6. *Aplicación de la Solución:* Se desarrolló una herramienta de software para obtener de forma automática, un análisis cuantitativo para cada una de las métricas de calidad propuestas, a través del relevamiento de datasets extraídos de los portales gubernamentales de datos abiertos;

7. *Validación de la Solución:* Se realizó la validación de las métricas sugeridas en base a los distintos aspectos definidos que debe reunir un conjunto de métricas para la calidad de datasets abiertos. Esta definición parte del estudio realizado sobre la ISO/IEC 25012, dimensiones de calidad, el estándar universal de calidad de datos de dos capas, trabajos de estudio propio en artículos de investigación y la bibliografía relevada en esta tesis.

Cabe destacar que la estrategia / metodología formal que permite identificar la investigación es del tipo “investigación-acción”, ya que este tipo de metodología se encarga del estudio de una problemática social específica, siendo en este caso el enfoque de la calidad del contenido de los datasets con información pública del estado nacional gubernamental (Información relevante a obras públicas, salud, educación, tecnologías, etc.). Este tipo de investigación propone una solución a un problema detectado, en este caso, las falencias reveladas en la calidad del contenido de los

datos públicos abiertos. En lo que respecta al planteo de una solución, esta metodología utilizada en la tesis presenta una propuesta de herramienta de desarrollo propio que permite validar los datos abiertos y, además, brindar un estudio detallado sobre lo encontrado en dicha verificación a modo de sugerencia de buenas prácticas.

5. PROPUESTA

Para analizar la calidad de datos es necesario cuantificar características en el conjunto de datos analizados. Para el estudio de la calidad de datos, esta tesis se orienta en los aspectos más relevantes en cuestiones de calidad de datos sobre criterios estándares de la Norma ISO/IEC 25012 [1], estándares universales de calidad de datos [2], dimensiones de la calidad de los datos [3], trabajos relevados y estudios realizados [4] en esta temática.

La utilización de métricas de calidad favorece al encuadre de indicadores que permitan obtener un dato más limpio para facilitar el análisis final. “Los indicadores para calidad de datos son herramientas importantes que debemos tomar en cuenta en los procesos de análisis, ya que permite medir y controlar la eficiencia de nuestros procesos que derivarán en análisis y toma de decisiones dentro de una estructura organizacional” [15].

Esta propuesta de métricas de calidad se basan en el estudio de guías y buenas prácticas de publicaciones de datos abiertos gubernamentales [16], en el formato recomendado para los tipos de datos que está mayormente basado en las especificaciones de la W3C [17], en la experiencia de estudios realizados con datasets relevados de diversos sitios gubernamentales de portales de datos abiertos [4], varios criterios enfocados en la calidad del dato (Normas y estándares) y guías elaboradas por el Gobierno de la Ciudad de Buenos Aires, las cuales plantean una serie de criterios para tener en cuenta en la apertura y tratamiento del contenido de los datasets públicos. Estas [18] fueron elaboradas por el equipo de la Dirección Nacional de Datos e Información Pública

junto con la Iniciativa Latinoamericana por los Datos Abiertos (ILDA), y su implementación fue posible gracias al financiamiento del fondo de transparencia del BID, que se beneficia de las generosas contribuciones de los gobiernos de Canadá, Noruega y Suecia [19]. La propuesta de métricas para el tratamiento de la calidad de los conjuntos de datos en formatos abiertos surge a raíz de un trabajo de investigación propio y publicado en la IEEE [20]. Las métricas se clasifican de la siguiente manera:

- **Métricas Críticas:** Contienen aquellas métricas que permiten detectar problemas de datos de una índole prioritaria para un correcto análisis de resultados con datasets, como ser: cuestiones de redundancia, contenido faltante en registros o bien datos erróneos. Es decir, es necesario tener en cuenta estos aspectos, ya que su presencia no favorece a un correcto estudio de los datos disponibilizados. Las métricas críticas propuestas: [MÉTRICA 1] Tratamiento de Números Decimales; [MÉTRICA 2] Registros Duplicados; [MÉTRICA 3] Datos Faltantes y/o Completos; [MÉTRICA 4] Caracteres inválidos.

- **Métricas No Críticas:** Contienen aquellas métricas que pudieran representar problemas de contenido en el dataset. Su detección está enfocada a posibles estimaciones de casos de errores y datos triviales, como así también, descubrimientos de datos redundantes combinados (entre campos y/o registros del dataset) que podrían conducir a inconvenientes en el análisis de un conjunto de datos. Las métricas no críticas propuestas: [MÉTRICA 5] Redundancia para el dominio de una columna; [MÉTRICA 6] Redundancia entre campos de una misma fila; [MÉTRICA 7] Detección de valores ID; [MÉTRICA 8] Campos Triviales.

6. VALIDACIÓN

Cada métrica propuesta, se la relacionó con un criterio de calidad que surgen del estudio de: a) Norma ISO/IEC 25012 [1]; b) Estándar Universal de Calidad de Datos [2]; c) Dimensiones de la calidad de los datos

(CDDQ) propuestas por Dan Myers [3]. El equipo de trabajo “Total Data Quality Management Program” [21], definieron un conjunto de atributos y/o dimensiones para medir y gestionar la calidad de los datos que enfocadas en la evaluación que puede ser automatizada para valorar la idoneidad y adecuación de los datos en orden a objetivos de negocio o bien necesidades. Estudios posteriores han ido modificando esta clasificación y fueron modificando diversas dimensiones. Uno de los estudios más recientes engloba algunas terminologías ya conocidas y proponen otras, como ser [3], que presenta una lista de dimensiones para la calidad de los datos, y lleva a cabo encuestas anuales con el fin de medir el uso de las dimensiones de la calidad de los datos por parte de las organizaciones.

Para este trabajo, por cada métrica propuesta, se la relacionó con una dimensión o criterio de calidad, que son el resultado de tomar en consideración, distintas fuentes:

a) Norma ISO/IEC 25012 [1], que especifica un modelo general de calidad de datos que se encuentran definidos en un formato estructurado dentro de un sistema informático. Para este se presentan los criterios del modelo de calidad de datos definido por el estándar ISO/IEC 25012, de las 15 características que lo componen, se tomaron para este trabajo, los criterios de: ***Exactitud, Completitud, Consistencia y Precisión.***

b) Estándar Universal de Calidad de Datos [2], son los criterios que debe contener un conjunto de datos para que puedan ser de calidad e interoperable y que son definidos por el estándar universal de la calidad de los datos de 2 capas. Para este aspecto se trabaja con el estándar universal en español extraído de la guía de estándares de calidad e interoperabilidad de los datos abiertos del gobierno de Colombia [22], para la validación de las métricas propuestas, se consideran los criterios de: a) Confidencialidad, siendo éstos: ***Precisión, Integridad, Consistencia y Completitud;*** b) ***Presentación, siendo éste el enfoque de estructura.***

c) Dimensiones de la calidad de los datos (CDDQ) propuestas por Dan Myers en DQMatters [3]. Los criterios considerados para la comparativa de validación son: ***Completo, Exactitud, Consistencia, Integridad, Precisión y Representación.***

d) Trabajos relevados y estudios de desarrollo propio para la tesis [4], [14], [20], y [23].

En base a la relación de cada criterio aplicado a las métricas propuestas, se identifica que éstas influyen en uno o más criterios de validación. el criterio. Por lo que, se observa que el criterio más representativo para validar la calidad de datos abiertos es el aspecto de Consistencia con un 92,85%, seguida por el aspecto de Integridad con un 85,71%. Luego, con un doble empate en porcentaje, el criterio de Precisión y Exactitud con el 57,14%. A éstos les sigue el criterio de Estructurales/Representación con el 50%, continuando con el 42,86% para el criterio de Relación entre valores de campos. Finalmente, concluye el criterio Redundancia con el 35,71%.

Esto conlleva a que las métricas propuestas se enfocan en un mayor porcentaje en métricas de validación orientadas a la “Consistencia”, como así también a la “Integridad” de los valores que contienen los datasets, por lo que su importancia radica en la coherencia general de los datos y el valor fiable que se encuentra almacenado. Por ejemplo, un conjunto de datos que contiene información sobre las personas internadas actualmente en un hospital determinado, se la puede considerar inconsistente si el recuento de las personas es mayor que el número de camas registradas. En lo que respecta a la Integridad, reúne un conjunto de características de calidad de datos mayores, relacionando más parámetros, como ser la precisión, fiabilidad y coherencia de los datos.

7. IMPLEMENTACIÓN

Herramienta Desarrollada - HEVDA:

Como parte de la propuesta de este trabajo, se presenta una herramienta desarrollada que

permite la validación de las distintas métricas sugeridas para un conjunto de datos abiertos en formatos CSV (valores separados por comas). Si bien la herramienta HEVDA permite obtener un análisis automático, no modifica el dataset de origen, sino que brinda un estudio detallado que sirve como guía práctica orientativa para la corrección de este. Esta herramienta de desarrollo propio fue parte del trabajo de investigación elaborado y publicado en una revista internacional [14].

La funcionalidad general del software HEVDA consiste en seleccionar un archivo dataset del tipo de formato CSV y efectuar la validación de las métricas propuestas. Para la elección del tipo de formato a considerar se realizó un estudio producto del cual se publicó un artículo [4], en el que se tomó como caso de muestra el portal gubernamental Argentina Unida [24] con sus 973 datasets a julio del año 2020. Sus resultados concluyeron que el formato más utilizado es el tipo CSV con un 61,6% de uso, de allí dicha elección.

Las funcionalidades detalladas de la herramienta programada son: Identificación del cumplimiento de las métricas críticas; Detección y detalle de los casos que no cumplen con el formato válido para el tipo de dato decimal; Cálculo estimativo de los tipos de datos de los campos del dataset validado; Cálculo de la cantidad y porcentaje de los registros duplicados; Detalle de los registros duplicados; Cálculo de la cantidad y porcentaje de los registros completos; Cálculo de la cantidad de casos que poseen campos con registros Nulos (Sin Datos ni espacios en los campos); Cálculo de la cantidad de casos que poseen campos con registros Vacíos (Sin Datos y con espacios en los campos); Cálculo de la cantidad de casos que poseen campos con registros No Disponibles (Con datos que indique N/D, N/A, NULL, -, - -, --); Visualización del detalle de los casos con registros Nulos, Vacíos y No Disponibles; Cálculo de la cantidad de columnas afectadas con caracteres especiales y su correspondiente detalle; Cálculo de la cantidad de columnas afectadas con valores repetidos en un mismo

campo (dominio de valores); Detalle con filtros de búsqueda para los campos del dataset y palabras de los campos detectados/as en los casos que hubo registros con valores repetidos para un mismo campo (dominio de valores); Cálculo de la cantidad y porcentaje de casos detectados con redundancia entre los valores de los campos para un mismo registro; Filtro de búsqueda para los casos detectados con datos redundantes entre los valores de los campos para un mismo registro y su correspondiente detalle; Estimación de la cantidad de ID identificados en las columnas del dataset, y la visualización de éstos; Cálculo de la cantidad de columnas afectadas con posibles campos triviales y su identificación.

Por otra parte, cabe destacar que al efectuar un análisis en un conjunto de datos se pueden encontrar una serie de aspectos que impiden el correcto análisis de validaciones, estos aspectos son de carácter bloqueante. Previo al comienzo del análisis con la herramienta HEVDA, un archivo dataset no deben cumplir con una serie de condiciones a las que se denominan “bloqueantes”, esto es, en caso de que exista algún tipo de bloqueante en la estructura del archivo, éste no podrá ser validado por el prototipo, ya que impedirá el procesamiento y análisis de las métricas propuestas del conjunto de datos. Los tipos de bloqueantes propuestos: a) El archivo posee doble caracter " (comilla); b) El archivo no cumple con la misma cantidad de columnas en cada uno de sus registros; c) El archivo no cumple con el formato CSV de separador/delimitador (coma); d) El archivo no posee una primera fila de títulos/nombres de las columnas del dataset; e) El archivo no es del tipo CSV.

8. CONTRIBUCIÓN CIENTÍFICA

Las contribuciones principales de este trabajo de tesis son las siguientes:

- Relevamiento del estado situación actual de los aspectos más relevantes en el tratamiento de datasets públicos abiertos a nivel nacional como internacional.
- Propuesta de un conjunto de métricas

críticas y no críticas para analizar la calidad de datos abiertos;

- Desarrollo propio de una herramienta para validar desde varios aspectos estándares, la calidad de los datasets publicados en portales de datos abiertos estatales;

- Detección de falencias en los datasets gubernamentales disponibilizados.

- Contribuir en el análisis, verificación y comprensión del estado actual de los valores que contienen los datasets generados por las entidades gubernamentales más relevantes de la Argentina.

- Aporte sobre las mejoras en la calidad del dato y concientización de su importancia para una correcta divulgación del contenido público tanto a nivel nacional como internacional.

9. CONCLUSIONES: REFLEXIONES FINALES

El paradigma de Gobierno Abierto, hoy por hoy desempeña un rol fundamental entre la conexión de distintos organismos gubernamentales y los ciudadanos. Gracias a este nuevo concepto es posible una mejor transparencia en base al tratamiento de rendición de cuentas por parte de entidades estatales, como así también, de una participación más comprometida con los ciudadanos de un país. Disponer de diversas guías que orienten la mejora constante de calidad de los datos abiertos es fundamental, pero, además, es vital contar con herramientas que permitan una rápida validación para tener una mejor visualización en cuanto a falencias o falta de integridad en los conjuntos de datos con el fin de aplicar las mejoras correspondientes en estos. Tener a disposición datos públicos abiertos de calidad, permitirá a los ciudadanos, una mejor confianza en las fuentes de datos y seguimiento de procesos administrativos del Estado Nacional.

Los resultados de los sondeos realizados evidencian que es posible definir métricas de calidad de datos públicos gubernamentales en un formato abierto para efectuar un análisis cuantitativo a través de una herramienta

amigable, sencilla e intuitiva para tratar aspectos de buenas prácticas.

Existen varias carencias en aspectos de calidad de los datos que se brindan en los datos abiertos, por lo que compartir buenas prácticas, guías de aprendizaje, foros o herramientas que permitan un mejor análisis y tratamiento de los datos es un beneficio para el ciudadano que desea acceder a la información pública. La propuesta de la herramienta HEVDA permite validar rápidamente el estado de “salud” de un conjunto de datos y analizar algunas de las métricas propuestas en aspectos de calidad, esto ayudará a saber si un conjunto de datos debe ser modificado o si está en condiciones de servir como base y analizar aspectos más profundos de los problemas de interoperabilidad del software para obtener valor agregado.

10. LÍNEAS FUTURAS

Algunas de las líneas futuras de investigación son las siguientes: a) Continuar trabajando en la ampliación de la herramienta de validación HEVDA; b) Incorporar a la herramienta desarrollada más tipos de formatos abiertos, como ser: XML, JSON, etc. lo que llevará a establecer nuevos aspectos de control en las estructuras de los formatos que serán implementados; c) Analizar las opciones de aplicaciones gráficas que podrían ser utilizadas con los datasets para ser embebidos en el código fuente de la herramienta HEVDA. Incorporar una herramienta de gráficos estadísticos con el análisis de las métricas críticas y no críticas; d) Analizar los datasets orientados a la geolocalización para el tipo de contexto: Coordenadas de longitud y latitud, formatos geoespaciales, archivos del tipo WKT (puntos de coordenadas), SHP (datos geográficos), etc.; e) Efectuar un estudio detallado de los datasets geoespaciales para proponer métricas de calidad de datos para este entorno de trabajo; f) Analizar la posibilidad de incorporar la utilización del prototipo HEVDA, en una entidad gubernamental.

11. PUBLICACIONES

Editorial: Editora Artemis. ebook “Ciências Socialmente Aplicáveis: Integrando Saberes e Abrindo Caminhos”. Curitiba, Brasil. 2022.

Título: “Detección de errores ortográficos para la validación de la calidad en datos abiertos gubernamentales para la métrica del factor syntactic correctness”.

Revista Abierta de Informática Aplicada (RAIA). 2022. Título: “Tipos de Métricas de calidad para validar datasets gubernamentales argentinos”

Journal of the School of Engineering of the Antioquia University, Colombia. 2022. Título: “Quality Study of open government data related to COVID-19 in Latin America”.

Journal of Science and Research: Revista Ciencia e Investigación. 2021. Título: “Quality evaluation of government open data sets in Argentina using the HEVDA Validation Tool”.

Journal: IEEE Latin America Transactions, 100(XXX). 2021. Título: “Metrics proposal to measure the quality of governmental datasets”.

Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI). 2022.

Trabajo I: Título: “Análisis de la Interoperabilidad en la República Argentina. Propuesta: Prototipo para el estudio de Catálogos de Datos Abiertos”.

Trabajo II: Título: “Tecnologías de la información facilitadoras para la interoperabilidad de software en gobierno abierto”.

Workshop de Investigadores en Ciencias de la Computación (WICC). 2022. Título: “Propuesta de técnicas de validación para la calidad de datos abiertos e identificación de patrones para predicciones con Machine Learning”.

Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI). 2021. Título: “Análisis de la Apertura de datos gubernamentales en los portales provinciales de la República Argentina: Aplicación HEVDA”.

50 JAIHO. Jornadas Argentinas de Informática. 2021. Título: “Validación de Métricas propuestas de Calidad para el estudio de datos abiertos en base a criterios estándares: Aplicación HEVDA”

V Biennial Congress of IEEE Argentina Section.IEEE. UTN Argentina, Resistencia. Virtual. 2020. Título: “Analysis of datasets and catalogs in government open portals of the Argentine Republic”.

8o Congreso Nacional de Ingeniería Informática – Sistemas de Información CoNaIISI. 2020. Trabajo I: Título: “Validación de la Calidad en Datos Abiertos con respecto a la detección de errores ortográficos utilizando la métrica del factor Syntactic Correctness”. Trabajo II: Título: “Propuestas de Buenas Prácticas para la Implementación de Smart City en el contexto de Datos Abiertos para la Salud Pública”.

14º Simposio Argentino De Informática en el Estado - SIE (49º JAIHO). 2020. Título: “Análisis del procedimiento para la solicitud de información pública y tiempos de respuesta. Caso de Estudio: Ciudad Autónoma de Buenos Aires (Argentina)”.

III Congreso Internacional de Ciencias de la Computación y Sistemas de Información (CICCSI). 2019. Título: “Criterios para el análisis preliminar de datos extraídos mediante web scraping”.

Congreso Argentino de Ciencias de la Computación (CACIC). 2019. Título: “Análisis de técnicas de raspado de datos en la web – aplicado al portal del Estado Nacional Argentino”.

BIBLIOGRAFÍA

[1] ISO 25012 (2008), “Ingeniería de software - Requisitos de calidad y evaluación de productos de software (SQuaRE) - Modelo de calidad de datos”, Disponible en:

<https://www.iso.org/obp/ui/es/#iso:std:iso-iec:25012:ed-1:v1:en>

[2] Cai, L., & Zhu, Y. (2015), “The challenges of data quality and data quality assessment in the big data era”, Data science journal, 14.

[3] Conformed Dimensions of Data Quality (2018), “Annual Survey about Use of Dimensions of Data Quality”, Disponible en:

http://dimensionsofdataquality.com/dims_survey

[4] Martínez, R., Rodríguez, R., & Vera, P. (2020), “Analysis of datasets and catalogs in

government open portals of the Argentine Republic”, In 2020 IEEE Congreso Bienal de Argentina (ARGENCON) (pp. 1-8). IEEE.

[5] BID – Banco Interamericano de Desarrollo (2021), “Gobierno Digital”, Disponible en:

<https://www.iadb.org/es/dialogo-regional-de-politica/la-red-de-gobierno-electronico-de-america-latina-y-el-caribe>

[6] Olaya, Y. E. L. (2018), “Estudio sobre minería y visualización de datos abiertos del gobierno de Colombia”, Working papers, Maestría en Ingeniería de Sistemas, 2(2).

[7] W3C (2009), “Publishing Open Government Data”, W3C Working Draft 8 September 2009, Disponible en:

<https://www.w3.org/TR/gov-data/>

[8] Modernización del Estado (2018), “Serie de Investigaciones: Gobierno Abierto, Gobierno Digital y País Digital”, Instituto Nacional de la Administración Pública (INAP), Disponible en:

https://www.argentina.gob.ar/sites/default/files/inap_serie_investigaciones_empleo_publico_modernizacion-estado-2018.pdf

[9] Ramírez-Alujas, Á. V. (2010), “Innovación en la gestión pública y open government (gobierno abierto): Una vieja nueva idea”, Innovation in Public Management and Open Government: An Old New Idea, Revista Buen Gobierno, (9)

[10] Open Contracting Partnership, Colman R. (2020), “Women win one in four contracts in the Dominican Republic thanks to inclusive procurement reforms”, Disponible en:

<https://www.open-contracting.org/2020/09/23/women-win-one-in-four-contracts-in-the-dominican-republic-thanks-to-inclusive-procurement-reforms/>

[11] WebFoundation.org (2018), “El barómetro de los datos abiertos. Edición de los líderes de la promesa al progreso”, Open Data Barometer, Disponible en:

https://webfoundation.org/docs/2018/09/WF_ODB_Report2_Spanish_Screen.pdf

[12] Cadena-Vela, S., Fuster-Guilló, A., & Mazón, J. N. (2019), “Publicando datos abiertos considerando criterios de calidad”.

[13] Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Herederó, C. (2018), “Indicadores de calidad de datos abiertos: el caso del portal de datos abiertos de Barcelona”, El profesional de la información (EPI), 27(2), 375-382.

[14] Martínez, R., Pons, C., Rodríguez, R., & Vera, P. (2021), “Quality evaluation of

government open data sets in Argentina using the HEVDA Validation Tool”, Journal of Science and Research: Revista Ciencia e Investigación. ISSN 2528-8083, 6(2).

[15] Graph Everywhere (2021), “Principales indicadores para Calidad de Datos”, Disponible en: <https://www.grapheverywhere.com/principales-indicadores-para-calidad-de-datos/>

[16] Datos.gob.ar (2021), “Estándares según el tipo de Datos”, Disponible en:

https://datosgobar.github.io/paquete-apertura-datos/guia_abiertos/#estandares-segun-el-tipo-de-datos

[17] W3C (2015), “Modelo para datos tabulares y metadatos en la Web”, Disponible en: <https://www.w3.org/TR/tabular-data-model/>

[18] Secretaría de Modernización. Presidencia de la Nación, “Paquete de Apertura de Datos de la República Argentina”, Disponible en: <https://datosgobar.github.io/paquete-apertura-datos/guia-subnacionales/#1-que-son-los-datos-abiertos>

[19] Datos.gob.ar (2021), “¿Por qué es importante estandarizarlos”, Disponible en: <https://datosgobar.github.io/paquete-apertura-datos/guia-interoperables/#por-que-es-importante-estandarizarlos>

[20] Martínez, R., Rodríguez, R. A., & Vera, P. M. (2021), “Metrics proposal to measure the quality of governmental datasets”, IEEE Latin America Transactions, 100(XXX), Disponible en: <https://latam.ieee9.org/index.php/transactions/article/view/5642>

[21] The MIT Total Data Quality Management Program (2002), “MIT TDQM Program Highlight”, Disponible en:

<http://web.mit.edu/tdqm/www/index.shtml>

[22] Gobierno de Colombia (2020), “Guía de datos, LEILA – Librería de calidad de datos”, Unidad de Científicos de Datos Dirección de Desarrollo Digital, Disponible en:

https://colaboracion.dnp.gov.co/CDT/Desarrollo%20Digital/Big%20Data/2020/00_LEILA/LEILA_Presentacion.pdf

[23] Martínez, R., et al. (2020), “Validación de la calidad en Datos Abiertos con respecto a la detección de errores ortográficos utilizando la métrica del factor Syntactic Correctness”, Congreso Nacional de Ingeniería Informática y Sistemas de la Información, Universidad Nacional de Tecnología (UTN), Facultad Regional San Francisco.

[24] Argentina unida, “Datos Argentina”, Disponible en: <https://datos.gob.ar/>

Modelo de madurez para servicios de gobierno electrónico en el ámbito universitario.

Tesis de Doctorado en Ciencias Informáticas
Facultad de Informática, Universidad Nacional de La Plata
Tesista: Ariel Pasini
apasini@lidi.info.unlp.edu.ar

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 La Plata Buenos Aires

Directoras: Patricia Pesado (UNLP) - Elsa Estévez (UNS)

Fecha de exposición: 30 de marzo de 2022

RESUMEN

Con la evolución de las Tecnologías de la Información y las Comunicaciones (TIC), el concepto de gobierno electrónico ha ido evolucionando hasta la etapa de madurez que se denomina gobernanza digital contextualizada, o dirigida por políticas. En esta etapa, los esfuerzos buscan la transformación digital en áreas de políticas o contextos específicos, como podría ser educación, salud o gobiernos locales. Generalmente se utiliza el término “gobierno” en la estructura gubernamental de un estado (municipios, provincias, países), pero también existen gobiernos dentro de diferentes contextos, como agencias impositivas, agencias de control gubernamental, agencias de administración de justicia, etc., que tienen cierto grado de autonomía del estado. Dentro de este tipo de estructuras se encuentran las universidades nacionales, que poseen su propio gobierno democrático interno y gozan de autonomía del gobierno político del Estado. El gobierno universitario está compuesto por alumnos, docentes, no docentes, graduados, alumnos de posgrado, etc., dependiendo de la estructura definida en su estatuto. Todos en su conjunto representan a los “ciudadanos”

universitarios que desarrollan sus actividades en el marco de las reglamentaciones que dispone dicho gobierno. Para llevar a cabo el cumplimiento de las reglamentaciones que exige el gobierno universitario, la universidad pone a disposición de cada tipo de ciudadano un gran número de servicios. Hoy en día, estos servicios son brindados a través del uso de las TICs, lo que conlleva a clasificar a estas prestaciones como servicios de gobierno electrónico universitario, que a su vez constituyen un ejemplo de gobierno electrónico dirigido por políticas; i.e. las políticas universitarias. En este contexto, dada la escasez de literatura relacionada con gobierno electrónico en el ámbito universitario y tratándose de una comunidad reducida, la temática brinda una oportunidad para investigar la forma de evaluar la calidad de las prestaciones de los servicios universitarios de forma controlada y certera.

El objetivo principal de esta tesis es definir un modelo de evaluación, basado en los conceptos aplicados en estructuras de gobiernos masivas, que permita clasificar en niveles de madurez a los servicios prestados por las unidades académicas y,

en consecuencia, permita definir un escalafón de unidades académicas en la prestación de servicios con el fin de ofrecer recomendaciones para mejorar la prestación de servicios a su comunidad.

CONTEXTO

El trabajo se realiza en el marco del Doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP.

El plan de trabajo se enmarca en el proyecto “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” (11/F023) (2018-2022) y continua en el nuevo proyecto “Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital” (2023-2026), UNLP, Ministerio de Educación de la Nación.

Asimismo, la temática se encuentra abordada en proyectos aprobados por la Facultad de Informática UNLP (2022-2023).

El Instituto posee diversos acuerdos de cooperación con varias Universidades de Argentina y del exterior

1 SINTESIS DE TESIS

La evolución del concepto de gobierno electrónico en los últimos veinte años ha sido notable desde todos los puntos de vista, pero los cambios más destacables se relacionan con: la mejora de procesos gubernamentales, la mayor interacción con el ciudadano y la construcción de nuevos canales de comunicación con los diferentes receptores de los servicios de gobierno. Los tres aspectos tienen como objetivo común mejorar las relaciones entre las partes interesadas con el uso de la tecnología, logrando una prestación de servicios, mediante procesos más eficientes y una utilización más eficaz de los recursos.

A lo largo de los años, el gobierno electrónico ha pasado por diferentes instancias: desde fines de los 90 a mediados de los 2000, el concepto se basaba en una presencia institucional en la web, con un sentido principalmente unidireccional, proporcionando información al ciudadano. Desde mediados de los 2000 a inicios del 2010, con el avance de la web 2.0 y la creación de las redes sociales, se inicia una etapa de socialización de servicios públicos, donde aparece la comunicación bidireccional entre el gobierno y los ciudadanos: los receptores de servicios dejan de tener un rol pasivo, donde sólo reciben información, y pasan a tomar un rol más participativo en el uso de los servicios públicos en línea. En el mismo período crece el uso de dispositivos móviles, obligando a los gobiernos a innovar en ese tipo de tecnologías para lograr un mayor compromiso de y con los ciudadanos. Desde el 2010, las comunicaciones móviles, y en particular el uso de datos móviles, pasan a ser parte de la vida cotidiana de los ciudadanos, lo que genera una importante demanda de la comunidad para que los gobiernos aumenten la prestación de servicios públicos en línea. Todo el proceso evolutivo generó un alto nivel de información digitalizada e importantes bases de datos que, con el tiempo, se fueron convirtiendo en uno de los principales insumos para el análisis de datos que hoy en día utilizan los gobiernos para apoyar la toma de decisiones.

Si observamos el avance del concepto de gobierno electrónico desde el punto de vista del uso de las TICs, en (Janowski, 2015) se presenta un modelo de cuatro etapas que va creciendo en complejidad tecnológica y política:

- 1. Digitalización o tecnológico:**
Incorpora el uso de tecnologías. No hay una transformación real del gobierno. Se utilizan los mismos procedimientos que se utilizaban sin

tecnología. La necesidad de mejorar los procesos internos para ser un organismo más eficiente es el inicio de la etapa de transformación.

2. **Transformación o gobierno electrónico:** Cambios estructurales de procesos internos. Mejora de estos procesos para un uso más eficiente de tecnología generando nuevas herramientas. Los ciudadanos hacen uso de estas herramientas y se ven beneficiados. Aumenta la confianza en el gobierno que permitirá avanzar hacia la etapa de compromiso.
3. **Compromiso o gobernanza electrónica:** Aumenta el compromiso entre la comunidad y el gobierno. Mejora la prestación de servicios públicos. Incrementa la participación ciudadana. Mejora la toma de decisiones, en consecuencia, se requiere un análisis sectorial para lograr una mejora real, lo que da origen a la etapa de contextualización.
4. **Contextualización o gobernanza dirigida por políticas:** Se generan políticas de gobernanza dirigidas a las necesidades de un contexto particular. Los contextos pueden ser regionales, territoriales, organizacionales, etc.

A medida que los servicios públicos evolucionan van alcanzando diferentes niveles de madurez. Según el modelo propuesto por la ONU (Concha & Naser, 2012) se definen cuatro niveles, basados en el nivel de automatización: 1) emergente, 2) mejorado, 3) transaccional y 4) integrado.

Retomando la idea de contextualización de gobernanza digital, el ámbito universitario se presenta como uno de los contextos de gobierno que se ve alcanzado por el gobierno electrónico, donde la Ley de Educación Superior (LES) (República

Argentina, 1995) da políticas explícitas sobre cómo debe ser gobernada una universidad pública en nuestro país.

Las universidades nacionales poseen su propio gobierno democrático interno y gozan de autonomía del gobierno político del Estado. El gobierno universitario está compuesto por alumnos, docentes, no docentes, graduados, alumnos de posgrado, etc., dependiendo de la estructura definida en su estatuto, y todos, en su conjunto, representan a los ciudadanos que desarrollan sus actividades en el marco de las reglamentaciones que dispone dicho gobierno.

El sistema de educación pública argentino posee, al 2015, 53 universidades nacionales (SPU - Secretaría de Políticas Universitarias, 2015). La Ley de Educación Superior Argentina (República Argentina, 1995), en su artículo 29, establece que las instituciones universitarias tendrán autonomía académica e institucional y definirán su estatuto y órgano de gobierno. Cada una de las universidades define, entonces, su estructura académica (facultad, departamento, escuela, etc.) y su cuerpo de gobierno (claustros que forman el consejo) de forma independiente.

Los gobiernos que dirigen las universidades, para cumplir con el objetivo principal de brindar una formación académica de alto nivel, prestan un amplio conjunto de servicios a su comunidad, dentro de los que se encuentra, por ejemplo, registrar a un alumno en una asignatura o registrar la nota de un alumno. Estos servicios son actualmente brindados a través del uso de las TIC. En línea con el concepto de gobierno electrónico presentado anteriormente, podemos inferir que los gobiernos universitarios también pueden beneficiarse de este concepto, y hacer uso del modelo de madurez de servicios públicos.

Tomando como base los conceptos de gobierno electrónico y de gobierno universitario se realizó una investigación

en bases de datos científicas, llegando a la conclusión de que no hay publicaciones, al menos como resultado de las búsquedas realizadas, que describan la aplicación de los conceptos de gobierno electrónico en el ámbito universitario. En consecuencia, se propuso relacionar dichos conceptos y generar un modelo de evaluación de gobiernos universitarios en función de los servicios prestados a su comunidad.

El modelo de evaluación propuesto (EGOV-U) busca estandarizar la identificación de los servicios prestados por las universidades, a fin de obtener una denominación común en todas las estructuras de unidades académicas analizadas. Para esto, como punto de partida, se tomaron 24 servicios: ocho correspondientes a alumnos, ocho a docentes, cuatro a graduados y cuatro a no docentes. La selección de los servicios se enfocó en las actividades básicas que realiza una unidad académica, buscando que representen a gran parte de la comunidad universitaria.

Seleccionados los servicios, se analizó el alcance de los mismos dentro de la unidad académica y se los clasificó según: los receptores de los servicios, el tipo de servicios (informacional, autorización, certificación y control) y el tipo de comunicación, identificando la combinación óptima para que cada servicio pueda alcanzar el máximo nivel de madurez. En base a los mismos criterios, se le otorgará una calificación al servicio que permitirá realizar una evaluación cuantitativa de los mismos.

Por otro lado, se definió una escala formada por cinco niveles que representan la evolución de la unidad académica en la prestación de estos servicios. Dicho nivel se obtendrá en función de los valores cuantitativos obtenidos por cada uno de los servicios prestados, permitiendo determinar el rango en el que se encuentra cada una de las unidades académicas.

Dado que el modelo se basa en 24 servicios y el modelo de madurez apunta a un

proceso de mejora continua, se describen, además, los pasos a seguir para una extensión del modelo que permita asegurar el crecimiento de la unidad académica más allá de los servicios planteados en el modelo presentado.

Se invitó a participar en el proceso de validación a 20 unidades académicas, de las cuales aceptaron 18. La evaluación se realizó mediante un instrumento de recopilación de información, publicado en Google Form, que contenía 115 preguntas agrupadas en cinco secciones. Las respuestas recibidas se procesaron en base a dos perspectivas: 1) según los tipos de servicios universitarios prestados para cada grupo de receptores y 2) por unidad académica.

Luego del análisis se pudo determinar el nivel de madurez de los servicios prestados por las unidades académicas y la clasificación obtenida, ubicándola dentro de uno de los cinco rangos propuestos por el modelo. Todas las unidades académicas participantes se encontraron por encima del mínimo aceptable y se generaron, para cada una de ellas, un conjunto de sugerencias con el fin de mejorar la prestación de sus servicios universitarios.

2 ENUNCIADO DEL PROBLEMA

La hipótesis de investigación planteada es:

Es posible aplicar los conceptos de gobierno electrónico en el contexto del gobierno electrónico universitario, identificando los servicios brindados por las unidades académicas, determinando su nivel de madurez y estableciendo una clasificación de la unidad académica junto a un conjunto de sugerencias de mejora que le permita evolucionar la prestación de sus servicios.

En base a la hipótesis planteada se definió el siguiente objetivo general:

Proponer un modelo de madurez para servicios de gobierno electrónico en el ámbito universitario, que aplique a las diferentes estructuras académicas e involucre servicios universitarios a diferentes receptores de la comunidad universitaria.

De la hipótesis y el objetivo surgen las siguientes preguntas de investigación:

¿Cuáles son los servicios que presta a su comunidad el gobierno universitario?

¿Cómo prestan los servicios las unidades académicas?

¿Cómo se podría mejorar la prestación de servicios?

3 CONTRIBUCIONES

La tesis presenta cuatro contribuciones principales:

1. La contextualización del concepto de gobierno electrónico al ámbito universitario, en particular, para la función de prestación de servicios
2. La definición del modelo EGOV-U, que permite evaluar la madurez de las unidades académicas en la prestación de servicios universitarios
3. Una evaluación sobre el uso estratégico de las tecnologías digitales que hacen las unidades académicas en Argentina para mejorar las interacciones con los miembros de su comunidad en relación a la prestación de servicios
4. Una herramienta práctica y de uso sencillo que las unidades académicas en Argentina pueden utilizar para avanzar en sus esfuerzos de transformación digital.

Relacionados con esas contribuciones se han obtenido las siguientes cinco publicaciones:

1. Pasini, A., Estevez, E., Pesado, P., ***Assessment Model for Digital Services provided by Higher Education Institutions***, en Proceedings of the 20th Annual International Conference on Digital Government Research, DGO. 2019, ACM, junio 2019, pp. 468-477, <https://doi.org/10.1145/3325112.3325268>. (indexada)
2. Pasini, A., Estevez, E., Pesado, P., Boracchia, M., ***Methodology for Assessing the Maturity Level of University Services***. Computer Science & Technology Series XXII Argentine Congress of Computer Science. Selected papers. La Plata: REDUNCI. 2017. p169 - 180. isbn 978- 987-41-2728-0. (indexada)
3. Pasini, A., Pesado, P., ***Quality Model for e-Government Processes at the University Level: A Literature Review***, en Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance, ICEGOV '15-16, ACM, marzo 2016, pp. 436-439, <https://doi.org/10.1145/2910019.2910106>. (indexada)
4. Pasini, A. C., Estévez, E. C., Pesado, P. M., Boracchia, M., ***Una metodología para evaluar la madurez de servicios universitarios***. Argentina. San Luis. 2016. Libro. Artículo Completo. Congreso. XXII Congreso Argentino de Ciencias de la Computación. REDUNCI
5. Pesado, P. M., Esponda, S., Pasini, A., Boracchia, M., Estevez, E., ***Normas y modelos de calidad para la mejora de productos y procesos de software, y de procesos de gestión***. Argentina. Salta.

2015. Libro. Artículo Completo. Workshop. XVII Workshop de Investigadores en Ciencias de la Computación. REDUNCI

4 CONCLUSIONES GENERALES

- **Evolución del gobierno electrónico:** Se describió la evolución del gobierno electrónico hasta la etapa de contextualización o gobernanza dirigida por políticas y la estructura de los gobiernos universitarios motivando el desarrollo de esta tesis.
- **Conceptos relacionados de gobierno electrónico y gobierno universitarios** Se presentaron los conceptos de: gobierno electrónico, servicios públicos, tipos de servicios y niveles de madurez. Por otro lado, las diferentes estructuras y formas de gobierno que puede tener una universidad. Generando el concepto de gobierno electrónico universitario (EGOV-U).
- **Revisión de literatura:** Se realizó una revisión de literatura en dos periodos, con escasos resultados en la temática planteada.
- **Se generó el modelo EGOV-U:** Se describió el modelo EGOV-U, que cuenta con 24 servicios básicos, según el tipo de servicio y el canal de comunicación utilizado, se estableció un nivel de madurez en la prestación del mismo que deriva en una calificación al servicio y un nivel de prestación de servicios de la UA. Seguida por recomendaciones de mejora para a las UA
- **Se validó el modelo en 18 UA:** Se definió un instrumento de recopilación de información de 115 preguntas que fue respondido por 18 UA de diferentes regiones del país.

- **Las respuestas obtenidas se procesaron en dos dimensiones:** 1) en base a los tipos de servicios por grupo de receptores. 2) por unidad académica.

• Limitaciones

- **Modelo de 24 servicios:** El modelo se propuso tomando un conjunto de 24 servicios universitarios representando a cuatro grupos de la comunidad universitaria.
- **Orientado a unidades académicas públicas:** El modelo está orientado principalmente a unidades académicas públicas.
- **Se contactó a docentes principalmente con actividades de gestión:** Para la recopilación de información se contactó a docentes de las unidades académicas, en lo posible con gestión administrativa.
- **Solo 4 receptores:** Para el modelo se tuvieron en cuenta solo cuatro tipos de representantes de la comunidad universitaria.
- **Especificidad del contexto:** Con respecto a las estructuras de las unidades académicas, el relevamiento se realizó sobre cada una de estas en particular; resta generalizar los resultados a las universidades.

5 LÍNEAS DE TRABAJOS FUTUROS

Se plantean tres posibles líneas de trabajos futuros:

- 1 **Extensión del modelo**
Se propone la ampliación del modelo básico. La extensión del mismo para analizar cualitativamente el uso de los canales de entrega por servicio.
- 2 **Desarrollo de herramientas**
Se evaluará la posibilidad de desarrollar herramientas de software para la autoevaluación del estado de

madurez de los servicios prestados por la UA

3 Generalización del modelo

Una posible área de estudio es analizar la generalización del modelo y su aplicación a otras áreas de contextos educativos – como escuelas primarias, secundarias, terciarias, e instituciones privadas.

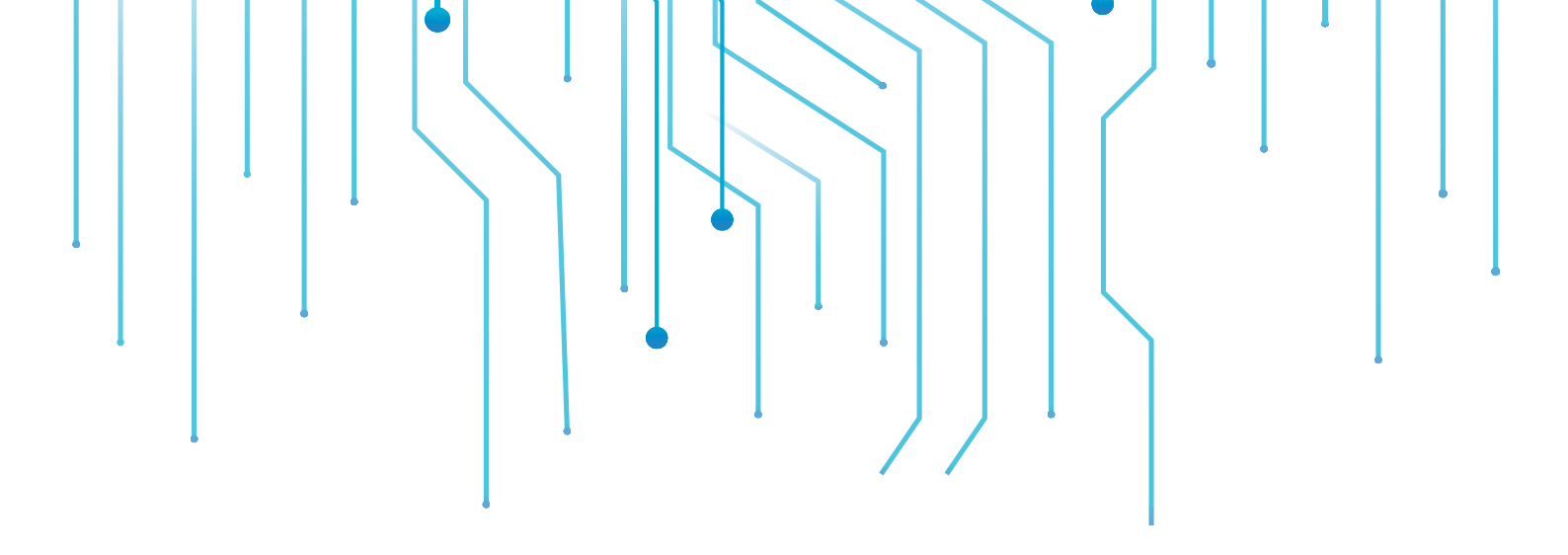
6. BIBLIOGRAFÍA

El informe de tesis completo se encuentra disponible en el repositorio SEDICI link: <https://doi.org/10.35537/10915/134522>

Se mencionan parte de la bibliografía utilizada en la tesis:

- Albalushi, T. H. and Ali, S. (2015) ‘Evaluation of the quality of E-government services: Quality trend analysis’, in *2015 International Conference on Information and Communication Technology Research, ICTRC 2015*, pp. 226–229. doi: 10.1109/ICTRC.2015.7156463.
- Andrian, R., Hendradjaya, B. and Sunindyo, W. D. (2016) ‘Software assessment model using metrics products for e-Government in the G2B model’, in *2016 4th International Conference on Information and Communication Technology, ICoICT 2016*. doi: 10.1109/ICoICT.2016.7571931.
- Audytra, H., Hendradjaya, B. and Sunindyo, W. D. (2017) ‘A proposal for quality assessment model for software requirements specification in Indonesian language for e-Government’, in *Proceedings of 2016 International Conference on Data and Software Engineering, ICoDSE 2016*. doi: 10.1109/ICODSE.2016.7936125.
- Balushi, T. A. and Ali, S. (2016) ‘Road map of service quality to E-government service quality’, in *Proceedings of the 27th International Business Information Management Association Conference - Innovation Management and Education Excellence Vision 2020: From Regional Development Sustainability to Global Economic Growth, IBIMA 2016*, pp. 3175–3180.
- Concha, G. and Naser, A. (2012) ‘CEPAL - El desafío hacia el gobierno abierto en la hora de la igualdad’. Available at: http://www.eclac.cl/cgi-bin/getProd.asp?xml=/publicaciones/xml/9/46119/P46119.xml&xsl=/publicaciones/ficha.xsl&base=/publicaciones/top_publicaciones.xsl.
- Fan, J. and Yang, W. (2015) ‘Study on E-Government services quality: The integration of online and offline services’, *Journal of Industrial Engineering and Management*, 8(3), pp. 693–718. doi: 10.3926/jiem.1405.
- Fath-Allah, A. et al. (2018b) ‘Towards an E-government portals quality framework based on ISO 25010’, in *2018 6th International Conference on Control Engineering and Information Technology, CEIT 2018*. doi: 10.1109/CEIT.2018.8751906.
- FIELD, T., MULLER, E. and LAW, E. (2003) *The e-Government Imperative*. OECD Publishing (OECD e-Government Studies). doi: 10.1787/9789264101197-en.
- Janowski, T. (2015) ‘Digital government evolution: From transformation to contextualization’, 32, pp. 221–223.
- Margariti, V. et al. (2020) ‘Assessment of organizational interoperability in e-Government: A new model and tool for assessing organizational interoperability maturity of a public service in practice’, in *ACM International Conference Proceeding Series*, pp. 298–308. doi: 10.1145/3428502.3428544.
- Papastylianou, A. et al. (2020) *Blended Learning and Open Courseware for Promoting Interoperability in Public Services, Communications in Computer and Information Science*.

- doi: 10.1007/978-3-030-37545-4_6.
- Pasini, A Estevez, E Pesado, P Boracchia, M. (2017) 'A Methodology for Assessing the Maturity Level of University Services', *Computer Science & Technology Series – XXII Argentine Congress of Computer Science. Selected papers.*, pp. 169–180.
 - Pasini, A., Estévez, E. and Pesado, P. (2018) 'Assessment Model for Higher Education Institution Digital Services'.
 - Pasini, A. and Pesado, P. (2016) 'Quality Model for e-Government Processes at the University Level – A Literature Review', pp. 1–4.
 - Peristeras, V. and Tarabanis, K. (2004) 'The governance enterprise architecture (GEA) high-level object model', *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 3035(May 2014), pp. 85–94. doi: 10.1007/978-3-540-24683-1.
 - Prakash, S. and Gunalan, I. (2020) 'An empirical framework for digital government assessment', in *ACM International Conference Proceeding Series*, pp. 271–278. doi: 10.1145/3428502.3428541.
 - Republica Argentina (1995) 'Ley Nro 24521 - Ley de Educacion Superior'.
 - Sandoval-Almazán, R. (2016) *Open justice in latin America? An assessment framework for judiciary Portals in 2015, Achieving Open Justice through Citizen Participation and Transparency.* doi: 10.4018/978-1-5225-0717-8.ch012.
 - Sandoval-Almazan, R. and Gil-Garcia, J. R. (2016) 'Toward an integrative assessment of open government: Proposing conceptual lenses and practical components', *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), pp. 170–192. doi: 10.1080/10919392.2015.1125190.
 - SPU - Secretaria de Politicas Universitarias (2015) 'Síntesis de Información Estadísticas Universitarias Argentina'.
 - Srimuang, C., Cooharajanane, N., Tanlamai, U. and Chandrachai, A. (2018) 'Development of an open government data assessment model: User-centric approach to identify the weighted components in Thailand', *International Journal of Electronic Governance*, 10(3), pp. 276–295. doi: 10.1504/IJEG.2018.095952.
 - Srimuang, C., Cooharajanane, N., Tanlamai, U., Chandrachai, A., *et al.* (2018) 'The study of public organization's intention to use an open government data assessment application: Testing with an applicable TAM', in *2017 12th International Conference for Internet Technology and Secured Transactions, ICITST 2017*, pp. 231–236. doi: 10.23919/ICITST.2017.8356389.
 - Supriyanto, A. and Mustofa, K. (2017) 'E-gov readiness assessment to determine E-government maturity phase', in *Proceeding - 2016 2nd International Conference on Science in Information Technology, ICSITech 2016: Information Science for Green Society and Environment*, pp. 270–275. doi: 10.1109/ICSITech.2016.7852646.



TIAE

Tecnología Informática Aplicada en Educación

Tecno pedagogía en el Proceso Formativo

Mg. Roberto Bertone¹, Mg. José Luis Filippi², Lic. Guillermo Lafuente³, Mg. Carlos Ballesteros⁴,
Lic. Gustavo Lafuente⁵, Pablo Etcheverry⁶, Analía Schpetter⁷

LIAU⁹ - Facultad de Ingeniería – UNLPam.

pbertone@ada.info.unlp.edu.ar¹

{filippij², lafuente³, balleste⁴, gustavo⁵, etcheverry⁶}@ing.unlpam.edu.ar

anaschpetter@gmail.com⁸

⁹Laboratorio de Investigación de Ambientes Ubicuos

Resumen

La tecno pedagogía es la tecnología aplicada a la educación. A través de ella se busca dar respuesta pedagógica al impacto de las herramientas digitales en el proceso formativo.

La tecno pedagogía utiliza estrategias didácticas como el aprendizaje significativo y colaborativo, conjuntamente incorpora recursos tecnológicos permitiendo el desarrollo de actividades de enseñanza virtuales que permitirán a los estudiantes accionar de forma autónoma.

Los centros educativos del siglo XXI se han de enfrentar a las nuevas demandas de una sociedad cambiante y con una clara vocación tecnológica. Aparecen nuevas formas de abordar lo tecnológico: aprendizaje adaptativo y móvil, inteligencia artificial, realidad virtual y aumentada y surgen las interfaces de uso natural (Johnson et al., 2016; Becker et al., 2017). [1]

El propósito del proyecto, de naturaleza teórico-práctico, consiste en indagar las posibilidades que ofrecen los dispositivos tecnológicos de última generación, como instrumentos aplicados a diferentes situaciones de aprendizaje, considerando los aspectos pedagógicos y las teorías del aprendizaje como marcos flexibles de la organización pedagógica. Las actividades que se proponen tienen lugar en el contexto de la UNLPam.

Palabras claves: Tecno pedagogía. Aprendizaje Móvil. Aprendizaje Colaborativo.

Contexto

Tipo de Investigación: Aplicada

Campo de Aplicación Principal: 7 1802 Computación, 7 1803 Comunicaciones.

Campos de Aplicación posibles: 13 1040 Ciencia y Tecnología, 7 4399 Otras – Educación – Tecnología Aplicada a la Educación

Institución que Coordina el Proyecto: Facultad de Ingeniería, Universidad Nacional de La Pampa.

Introducción

Con la aparición del coronavirus a comienzos del 2020, los cambios tecno pedagógicos que se realizaban de manera natural en los centros educativos, se precipitaron como resultado de la situación de encierro ciudadano.

A partir del 2022 con la vuelta a la formación presencial, se visualiza un modelo educativo donde lo presencial y lo virtual van a convivir de forma permanente, y en este punto el accionar docente es vital para desplegar su actividad profesional según los modelos que la sociedad moderna exige.

Aparecen nuevas estrategias formativas como el microlearning [2] y el aprendizaje inversivo. Surge el desafío educativo, definir metodologías y procesos para desplegar una formación pedagógica eficaz. Por ello se requiere un diseño tecno pedagógico instruccional adaptado al contexto en el que se va a desplegar la tarea educativa, con

docentes capacitados en desempeñar su profesionalidad de forma híbrida. [3].

A partir del estrepitoso cambio de paradigma educativo que se articuló en marzo del 2020, donde se determina el paso de la formación presencial a la virtual, numerosos actores se pronuncian en priorizar el desarrollo de habilidades integrales además de las cognitivas. Por ello es necesario "...construir otras formas de hacer educación, y que los cambios curriculares den cuenta de la posibilidad, la necesidad y la urgencia de vincular los contenidos disciplinares con habilidades de aprendizaje socioemocionales en las aulas...". [4]

Aquí radica el objetivo primario del proyecto: indagar las posibilidades que ofrecen los dispositivos tecnológicos de última generación, como instrumentos aplicados a diferentes situaciones de aprendizaje, considerando los aspectos pedagógicos y las teorías del aprendizaje como marcos flexibles de la organización pedagógica.

Situación Actual del Problema.

La educación influenciada por ésta realidad, acepta la necesidad de reconfigurar el proceso educativo, partiendo de algunas premisas:

- a. Redefinir nuevas acciones para docentes y estudiantes.

El creciente número de dispositivos móviles y la conectividad a internet de forma permanente, permite a los estudiantes adquirir nuevos contenidos en línea y configurar sus conocimientos a partir de intereses propios. El estudiante es artífice de su futuro, asume un papel activo, decide que aprender y cómo.

El rol tradicional del docente es superado por nuevas exigencias, requiere nuevas capacidades y conocimientos para desempeñar su actividad en entornos virtuales de aprendizaje.

Frente a la realidad del mundo digital que rodea a los estudiantes de hoy en día, es justo que los docentes replanteen su rol en la

construcción del conocimiento sobre ellos. De allí que la competencia exigida a un profesor del Siglo XXI es precisamente que se enfoquen "en las necesidades de los alumnos, supervisando su búsqueda de información e intentando facilitar la búsqueda de información individual de los alumnos ya que el papel de suministrador de conocimiento ha sido superado por las TIC" [5].

Mucho se habla de cuáles serían las competencias digitales para un docente, un acercamiento a ello lo propone Núñez-Torrón Stock [6], quien plantea cinco competencias básicas para el docente en la educación, y entre ellas se pueden destacar dos:

1. Usar herramientas de trabajo en línea.
2. Utilizar dispositivos móviles en el aula.

En la mayoría de las universidades de todo el mundo se observa la incorporación de los sistemas de Gestión de Aprendizaje (LMS), para implementar entornos virtuales que permiten la gestión del aprendizaje. [7] Sin embargo algunas investigaciones indican que las motivaciones de los estudiantes son vistas como una obligación al participar en entornos de aprendizaje institucionales cerrados, mientras que la motivación es mayor cuando el proceso formativo es abierto a diferentes tecnologías en contextos informales. Estos espacios son conocidos como PLE (Personal Learning Environments) donde los estudiantes configuran su propio entorno. [8]

Algunos autores ya incorporan un nuevo concepto, el de mPLE (Mobile Personal Learning Environments) como una nueva estructura de aprendizaje para la generación que hace uso de los dispositivos móviles. [9]

La agregación del modelo tecno pedagógico, que integra las tecnologías en todas las áreas del conocimiento, va a permitir desplegar las competencias digitales de los docentes en el aula, influenciando de manera positiva el desempeño de los estudiantes.

La tecno pedagogía utilizada como herramienta de planificación pedagógica expone diferentes actividades formativas que

conectan lo que se enseña en la práctica con lo que aprende el estudiante.

Es un modelo que viabiliza la construcción del conocimiento de forma colaborativa, donde los docentes efectúan su aporte según su área del conocimiento específico, tecnología, pedagogía y/o contenidos. Así, los docentes aprenden unos de otros, lo que puede derivar en el desarrollo colaborativo del currículo. [10]

La organización de la tecno pedagogía en el aula, integra las tecnologías con la pedagogía y el conocimiento científico. Para que la integración sea efectiva, los docentes no solo deben poseer conocimientos en tecnología y pedagogía, se debe poseer además el conocimiento propio de la disciplina, y desde allí integrar el conocimiento tecnológico y el conocimiento pedagógico. [11].

b. Facilitar la continuidad del aprendizaje. El aprendizaje móvil permitirá a los estudiantes dar continuidad a su formación a lo largo de toda su vida.

Los estudiantes universitarios no solo deben dominar las disciplinas correspondientes a su carrera, además deben incorporar habilidades transversales como el pensamiento crítico, resolución de problemas, persistencia y trabajo colaborativo. [12]

En la actualidad se está transitando tiempos de pos pandemia, y si algo bueno se puede rescatar de esta situación, es la consolidación del uso de los dispositivos tecnológicos de última generación durante los últimos dos años, escenario que alcanzo a todos los sectores sociales incluidas las instituciones educativas en sus diferentes niveles. Entre los instrumentos tecnológicos que sobresalieron en las actividades formativas virtuales que se desplegaron durante la pandemia y que van a prevalecer pos pandemia, podemos indicar [13], aulas colaborativas, educación personalizada e integrar en las actividades

pedagógicas las tecnologías emergentes como nuevos ambientes de enseñanza y aprendizaje. Se trata no solo de afianzar los conocimientos adquiridos de los estudiantes, sino el de desarrollar en ellos las habilidades cognitivas y meta cognitivas para la realización de tareas en su quehacer diario, objetivo primario del diseño tecno pedagógico. [14]

c. Grandes volúmenes de datos. Big data. Los investigadores que estudian el aprendizaje en línea, los sistemas de tutoría inteligente, los laboratorios virtuales, las simulaciones y los sistemas de gestión del aprendizaje están explorando maneras de entender y utilizar mejor la analítica del aprendizaje, a fin de mejorar la actividad del docente y seguir avanzando en la educación para todos. [15]

En éste contexto el punto de partida debe ser siempre tecno pedagógico. No se trata de innovar por innovar, hay que explotar el potencial que ofrecen las nuevas herramientas tecnológicas que hacen posible el aprendizaje móvil, a partir de una planificación docente previa dentro de un marco teórico formalizado.

El BigData trae beneficios de cara a la personalización de la educación en el alumno/a de modo que estudiantes como profesores sepan en qué punto del aprendizaje se encuentran, cuáles son sus debilidades y fortalezas. [16]

Línea de Investigación y Desarrollo

El plan de actividades corresponde al proceso de investigación aplicada, con objetivos de corto, mediano y largo plazo, y una duración prevista de cuatro años.

Primer año.

- Identificar el porcentaje de inserción de dispositivos móviles en la institución (teléfonos y/o tablets).
- Analizar la inserción del aprendizaje móvil en las universidades de todo el mundo.

- Estudiar como el aprendizaje móvil puede cerrar la brecha entre el aprendizaje formal y el informal.
- Identificar los escenarios educativos que muestren aspectos móviles en las actividades formativas del ámbito académico propio.
- Examinar herramientas tecnológicas móviles orientadas al proceso educativo de acceso libre y gratuito.

Segundo y Tercer año.

- Definir el enfoque tecno pedagógico que posibilite un aprendizaje móvil de calidad.
- Estudiar las tecnologías disponibles para implementar acciones de aprendizaje móvil en el contexto educativo universitario.
- Definir prioridades según necesidades de inmediatez en la virtualización de las disciplinas que se han de impartir.
- Gestionar el uso de plataformas de formación virtual que posibilite una doble modalidad, presencial y mediada por las nuevas tecnologías móviles.
- Incorporar el uso de redes sociales que favorezcan la práctica educativa.
- Confeccionar objetos de aprendizaje acorde a las herramientas disponibles en la nube y a las características de los dispositivos móviles que van a operar el producto final.
- Diseñar y/o desarrollar aplicaciones móviles a partir de las necesidades que se presenten durante el transcurso del proceso educativo virtual.
- Capacitar a la comunidad educativa en general en el desarrollo de objetos de aprendizaje a partir de los requerimientos de carácter institucional.
- Difundir los avances a toda la comunidad universitaria los progresos a medida que van transcurriendo, a través de jornadas, congresos y/o revistas científicas.
- Propiciar el intercambio de información permanente con grupos de investigación que den valor agregado a nuestra actividad profesional.

Cuarto año.

- Instituir las aplicaciones desarrolladas para dispositivos móviles que hayan alcanzado buen nivel de aceptación.
- Conformar un repositorio de objetos de aprendizajes de acceso libre.
- Registrar y difundir los resultados alcanzados con la finalidad de que se puedan utilizar, ampliar y mejorar a través de trabajos futuros.
- Disponer de un catálogo de trabajos realizados en el sitio web del grupo de investigación GIAU (Grupo de Investigación de Ambientes Ubicuos) a partir del cual los interesados podrán acceder al material requerido.
- Presentar en jornadas, congresos y/o revistas de todo el mundo los resultados alcanzados.

Resultados Obtenidos/Esperados

El proyecto da inicio a su actividad de investigación y desarrollo durante el 2023, por lo que sus integrantes se encuentran efectuando las tareas enmarcadas en el primer año.

Al finalizar el proyecto se espera contribuir en la incorporación de aplicaciones tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las necesidades de los usuarios que transitan en el ámbito educativo de la UNLPam y la UNLP, profundizando en el aprendizaje móvil, considerando los aspectos pedagógicos y las teorías del aprendizaje como marcos flexibles de la organización pedagógica.

Formación de Recursos Humanos

Director de Proyecto

Co-Director de Proyecto

6 Investigadores

Referencias

[1] Integración de Tecnologías de la Información y la Comunicación a la Docencia
- 2017

<https://nuestroscursos.net/mod/url/view.php?id=1857>

[2] Que es un Nanodegree. Webinar. 9 de febrero de 2016

<https://www.itmadrid.com/que-es-un-nanodegree/>

- [3] Tendencias eLearning 2022. <https://www.itmadrid.com/wp-content/uploads/2021/12/itmadrid-tendencias-elearning-2022.pdf>
- [4] [UNESCO (2021), Educación pos pandemia: salud mental y prevención de la violencia con la comunidad educativa, Memoria, UNESCO, en: UNESCO (2021), Educación pos pandemia: salud mental y prevención de la violencia con la comunidad educativa, Memoria, UNESCO, en: <https://es.unesco.org/sites/default/files/doc-memoria-foro-educacion-post-pandemia.pdf>.
- [5] webdelmaestrocmf.com, “Estas son las capacidades digitales de un docente moderno.” [Online]. Available: http://webdelmaestrocmf.com/portal/5-competencias-digitales-que-deben-tener-los-profesores-actuales/?utm_source=blogsterapp&utm_medium=Facebook. [Accessed: 27-Oct-2017].
- [6] A. Núñez-Torrón Stock, “5 competencias digitales que deben tener los profesores actuales.” [Online]. Available: <http://www.ticbeat.com/educacion/5-competencias-digitales-profesores/>. [Accessed: 27-Oct-2017].
- [7] M. P. Prendes, “Plataformas de campus virtual con herramientas de software libre: Análisis comparativo de la situación actual en las universidades españolas,” Murcia, Project Report EA-2008-0257, Jun. 2009. [Online]. Available: http://www.um.es/campusvirtuales/informe_final_CVSL_SF.pdf
- [8] F. J. García-Peñalvo, “Docencia,” in Libro blanco univ. digital 2010, J. Laviña Orueta and L. Mengual Pavón, Eds. Barcelona, Spain: Ariel, 2010, pp. 29-61.
- [9] G. Attwell, J. Cook, and A. Ravenscroft, “Appropriating Technologies for Contextual Knowledge: Mobile Personal Learning Environments,” in Best Practices for the Knowledge Society. Knowledge, Learning, Development and Technology for All, M. D. Lytras, P. O. de Pablos, E. Damiani, D. Avison, A. Naeve, and D. G. Horner, Eds. Springer Berlin Heidelberg, 2009, pp. 15-25.
- [10] Technological Pedagogical Content Knowledge (TPACK) in Action. https://www.researchgate.net/publication/271937971_Technological_Pedagogical_Content_Knowledge_TPACK_in_Action
- [11] [Kiray, S. A., Çelik, I., & Çolakoğlu, M. H. (2018).TPACK Self-efficacy perceptions of science teachers:A structural equation modeling study. Egitim ve Bi-lim, 43(195), 253-268. <http://egitimvebilim.ted.org.tr/index.php/EB/article/view/7538>]
- [12] World Economic Forum, 2015.
- [13] Tendencias educativas que marcaran el 2022 y el futuro de las aulas. <https://www.entrestudiantes.com/2022/01/tendencias-educativas-marcaran-2022-futuro-aulas/>.
- [14] Pedroza, O., & Crespo, M. (2017). Importancia del diseño tecno pedagógico basado en el enfoque de la acción, para reforzar el dominio del idioma inglés como segunda lengua. Revista Colombiana de Computación, 18(2), 7-21, e-ISSN: 2539-2115. Recuperado de: <https://doi.org/10.29375/25392115.3214>
- [15] Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief.US Department of Education, Office of Educational Technology, 1-57.
- [16] Otero, A. (2020). El Big Data. Como se usaría en la educación, vermislab. Recuperado de <https://www.vermislab.com/big-data-que-es-como-nos-afecta-y-su-uso-en-educacion/>

ESTRATEGIAS DE ACCESIBILIDAD PARA UNA EDUCACIÓN INCLUSIVA EN LA UNSAdA.

Adó Mariana, Ahmad Tamara, Russo Claudia, Sarobe Mónica, Nicolás Alonso

Universidad Nacional de San Antonio de Areco
(UNSAdA)

Rivadavia y Quetgles (CP 2760) - TE 2326 421167

San Antonio de Areco, Bs. As.

{mado, ccrusso, msarobe, nalonso}@docentes.unsada.edu.ar,
tamaraahmad@unnoba.edu.ar

RESUMEN

Pensar en una verdadera inclusión y democratización de la educación superior, implica incluir y representar a todas las minorías dentro de la Universidad, entre ellas, las personas con discapacidad.

En nuestro país, este colectivo goza de un amplio abanico de protección y legislación tanto internacional como nacional, que regula entre otras cosas, su acceso y permanencia en la educación superior.

En los últimos años, las universidades argentinas, así como los espacios interuniversitarios, han puesto sus esfuerzos en el derecho a una educación de calidad para todas las personas, partiendo de los principios universales de inclusión, diversidad, igualdad y equidad, estableciendo distintos lineamientos y políticas educativas, entre ellos, aquellos que buscan achicar la brecha digital existente al acceso a la información y al conocimiento para personas en situación de discapacidad.

Esta línea de investigación propone, por un lado, hacer un relevamiento de la legislación nacional, convenios y tratados internacionales que reglamentan la accesibilidad y el derecho a la educación de las personas con discapacidad y el estado del arte, que

permitan a la UNSAdA tener un marco de referencia para definir estrategias que posibiliten impartir una educación inclusiva mediada por tecnología, aplicando principios de accesibilidad y diseño universal; y, por otro lado, capacitar a docentes y no docentes con el fin de sensibilizar a la comunidad educativa en materia de discapacidad, accesibilidad y educación inclusiva.

Palabras clave: Accesibilidad, Discapacidad, Inclusión, Equidad, Educación inclusiva

CONTEXTO

Esta línea de investigación se circunscribe en el marco del proyecto de investigación “Educación a distancia e innovación tecnológica” perteneciente a la convocatoria Núcleos de Organización y Ejecución de Actividades Científicas y Tecnológicas (NACT) de la Secretaría de Investigación e Innovación de la UNSAdA.

1. INTRODUCCIÓN

En los últimos años, la educación superior ha puesto como una de sus preocupaciones principales el derecho a una educación de calidad para todas las personas, partiendo de los principios

universales de inclusión, diversidad, igualdad y equidad [1].

No puede pensarse en una verdadera inclusión y democratización de la educación superior, si hay minorías que no son representadas dentro de la Universidad, entre ellas, las personas con discapacidad. Por este motivo, en la actualidad la mayoría de las universidades del país cuentan con un área específica dedicada a la inclusión y accesibilidad de las personas con discapacidad. Sus formas de organización varían dependiendo de cómo se articule en cada Institución de Educación Superior (IES), entre ellas podemos encontrar, áreas, comisiones o programas con el propósito de coordinar y articular acciones comunes [2].

En el año 2003 se crea la Comisión Interuniversitaria: Discapacidad y Derechos Humanos, este espacio interuniversitario ya funcionaba desde el 1992 pero con otras denominaciones. Luego, en el año 2017 pasa a llamarse Red Interuniversitaria de Discapacidad (RID) y pertenece a la Comisión de Extensión, Bienestar Universitario y Vinculación Territorial del Consejo Interuniversitario Nacional (CIN) desde el año 2008.

El objetivo de la RID es:

“constituirse en un actor social de la Política Pública de Educación Superior, en pos de asegurar el ejercicio operativo al derecho a la educación, generando un espacio político para el diseño de estrategias que amplíen las capacidades institucionales contribuyendo al logro de Universidades accesibles y no excluyentes” [3].

La RID adscribe al modelo social de la discapacidad, a partir de la premisa de que es el entorno quien actúa como barrera creando situaciones discapacitantes para las personas y que la discapacidad no es una barrera en sí, sino

que esta queda evidenciada por la interacción con un entorno discriminante. Contemplando que “no hay una persona naturalmente discapacitada, sino que tal situación se produce por un entorno y actitudes discriminatorias” [2] que evitan su participación plena y efectiva en la sociedad, en igualdad de condiciones con los demás.

En este sentido, parafraseando a Mareño Sempertegui [4], se habla de personas en situación de discapacidad para referirse a aquellas minorías poblacionales tradicionalmente denominadas como lisiados, impedidos, minusválidos, deficientes, discapacitados o personas con discapacidad, según diversos momentos históricos y perspectivas teóricas de análisis. Se concibe a la discapacidad como una situación construida social y culturalmente, una creación humana que todos los actores sociales producimos día a día a partir de actitudes, decisiones, acciones y omisiones fundadas en la naturalización de criterios de normalidad. Es la situación construida, que involucra aspectos físicos, sociales, culturales, políticos, económicos y actitudinales la que limita o no limita, la que permite o no permite, la que posibilita o no posibilita; independientemente de las características personales o corporales de las personas que participen de ella.

En relación con lo anterior, en el año 2011, la RID presenta el documento llamado “Programa Integral de Accesibilidad en las Universidades Públicas. Profundización y avances en su implementación”, resultante de la Reunión Extraordinaria de la Comisión Interuniversitaria: Discapacidad y Derechos Humanos, realizada entre los días 5 y 6 de septiembre de 2011, con sede en la Universidad Nacional de Tucumán, que se inscribe en el marco del Programa Integral de Accesibilidad en las Universidades Públicas, aprobado

por Resolución N° 426/07 del CIN, en el cual se promueve una “Universidad para todas y todos sin distinción, desde la perspectiva de considerar a la Educación como derecho y como bien público social” [5] y en el cual se venía trabajando sobre el tema de accesibilidad física en las universidades.

En este contexto, se define accesibilidad universal como:

“el requisito fundamental que deben cumplir los espacios, ámbitos, servicios, bienes, así como los objetos o instrumentos, herramientas y dispositivos, para poder ser utilizables por todas las personas de manera: autónoma, es decir, en forma independiente, no subordinada al auxilio de otra; segura, o sea, libre de todo peligro o riesgo; confortable, es decir, de manera cómoda sin inconvenientes, obstáculos o restricciones en su uso, y con un mínimo de esfuerzo; y equitativa, es decir, con equidad entendida como la cualidad que consiste en atribuir a cada persona aquello que le corresponde por derecho” [6].

La incorporación de las tecnologías de la información y comunicación (TIC) en las universidades, así como los Entornos Virtuales de Enseñanza y Aprendizaje (EVEA), abrieron una gran variedad de posibilidades de acceso al conocimiento para una gran cantidad de personas, pero al mismo tiempo, han aumentado la brecha digital respecto a la “gran minoría de personas para quienes la desigualdad de oportunidades en el acceso a la información y al conocimiento no les permite participar en igualdad de condiciones en la sociedad de la información” [7].

En tal sentido, la Convención Internacional sobre los Derechos de las Personas con Discapacidad (CDPD), aprobada el 13 de diciembre de 2006 por la Asamblea General de las Naciones

Unidas, ratificada por la República Argentina por medio de la Ley N° 26.378 en junio de 2008, define como su propósito “promover y proteger los derechos y la dignidad de las personas con discapacidad, sobre la base de un enfoque holístico de la labor realizada en las esferas del desarrollo social, los derechos humanos y la no discriminación” [8].

Asimismo, la CDPD, sobre la comunicación necesaria para las personas con discapacidad define:

La comunicación incluirá los lenguajes, la visualización de textos, el Braille, la comunicación táctil, los macrotipos, los dispositivos multimedia de fácil acceso, así como el lenguaje escrito, los sistemas auditivos, el lenguaje sencillo, los medios de voz digitalizada y otros modos, medios y formatos aumentativos o alternativos de comunicación, incluida la tecnología de la información y las comunicaciones de fácil acceso [8].

Además, se entenderá por lenguaje el lenguaje oral, así también como la lengua de señas y otras formas de comunicación no verbal [8].

Determina a los ajustes razonables como las modificaciones y adaptaciones necesarias y adecuadas que no impongan una carga desproporcionada o indebida, cuando se requieran en un caso particular, para garantizar a las personas con discapacidad el goce o ejercicio, en igualdad de condiciones con las demás, de todos los derechos humanos y libertades fundamentales [8].

Igualmente, define el concepto de diseño universal como “el diseño de productos, entornos, programas y servicios que puedan utilizar todas las personas, en la mayor medida posible, sin necesidad de adaptación ni diseño especializado”, y qué éste además, “no excluirá las ayudas técnicas para grupos particulares de

personas con discapacidad, cuando se necesiten” [8].

En relación con lo anterior, la ley 26.653, sobre “Accesibilidad de la Información en las Páginas Web”, obliga a todas las entidades públicas, incluidas las IES a desarrollar páginas web accesibles con el objetivo de la democratización del acceso a la información, el conocimiento y, por ende, a la educación [9].

Finalmente, es necesario que en las universidades se reconozca “la importancia de incorporar condiciones de accesibilidad como condición necesaria para favorecer la expansión de la equidad y el acceso a la educación superior y, por ende, fortalecer los procesos institucionales de democratización del conocimiento y la información” [10].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta línea de investigación se busca definir estrategias que permitan llevar adelante acciones educativas para la inclusión de las personas con discapacidad en la universidad. Para ello, se establece como objetivo principal definir y desarrollar estrategias de accesibilidad para una educación inclusiva en la UNSADA.

Como objetivos específicos se definen:

1. Analizar el marco legal, el marco teórico y el estado del arte referente al tema de accesibilidad y discapacidad en educación.
2. Definir estrategias que permitan impartir una educación inclusiva mediada por tecnología, aplicando principios de accesibilidad y diseño universal.

3. Capacitar a todos los actores responsables de la acción formativa.

3. RESULTADOS OBTENIDOS / ESPERADOS

Con esta investigación, en primer lugar, se pretende analizar el marco teórico, el marco legal y estado del arte referente al tema de accesibilidad y discapacidad en educación, en particular, para educación superior, con el fin de contar con un marco de referencia que dé lugar a la definición de estrategias que permitan impartir una educación inclusiva mediada por tecnología, aplicando principios de accesibilidad y diseño universal.

Además, se aportará a la creación de una guía que posibilite la producción, administración y almacenamiento de contenidos digitales bajo la forma de Recursos Educativos Digitales (RED). La guía se propondrá desde una mirada interdisciplinaria que permita determinar y sistematizar aspectos determinantes en el desarrollo de materiales didácticos digitales [11], siguiendo principios de accesibilidad web y diseño universal.

Asimismo, dentro del marco de esta línea de investigación, se está programando una capacitación para el plantel docente y no docente de la universidad con el fin de comenzar a introducir el tema de la discapacidad, con el fin de sensibilizar a la población universitaria.

Por último, se aspira a la difusión y generación de conocimiento mediante la presentación y participación en diferentes congresos, jornadas y workshops de carácter nacional e internacional relacionados con la temática.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por cuatro investigadores formados del área de Sistemas y Tecnología aplicada a la educación, un investigador del área de Sistemas, dos investigadoras en formación y dos becarios de grado.

Asimismo, se espera que, durante el transcurso del año, se concrete la definición de dos Tesinas de grado relacionadas con el tema.

5. BIBLIOGRAFÍA

[1] Cicerchia L., Addante P., Ahmad T., Lencina P., Russo C. (2014). *Accesibilidad Web con herramientas de Software Libre en UNNOBA Virtual*. Recuperado de: https://digital.cic.gba.gob.ar/bitstream/handle/11746/7087/Poster_Accesibilidad_Web_UNNOBA_CICERCHIA.pdf?sequence=1

[2] Seda, J. A. (2015). *Discapacidad y universidad: Interacción y respuesta institucional. Estudio de casos y análisis de políticas y legislación en la Universidad de Buenos Aires*. Eudeba. Libro digital, EPUB.

[3] Sitio web de La Red Interuniversitaria de Discapacidad, Consejo Interuniversitario Nacional. <https://rid.cin.edu.ar/>

[4] Mareño Sempertegui, M. (2010). *Inclusión Educativa en la Universidad Nacional de Córdoba: el desafío de aplicar los principios del Diseño Universal en la gestión de políticas universitarias*. Ponencia presentada en: Primeras Jornadas Nacionales de Accesibilidad y Diseño Universal. Córdoba, Argentina: Universidad Nacional de Córdoba.

[5] RID-CIN. (2011). *Programa Integral de Accesibilidad en las Universidades*

Públicas Profundización y avances en su implementación. RID - Red Interuniversitaria de Discapacidad. <https://rid.cin.edu.ar/>

[6] Mareño Sempertegui, M., Torrez, V. (2013). *Accesibilidad en los entornos virtuales de las instituciones de educación superior universitarias*. Virtualidad, Educación y Ciencia. Vol. 4, Núm. 7. Recuperado de: <https://revistas.unc.edu.ar/index.php/vesc/article/view/6176>

[7] Sáenz Espitia, J.G. (2008). *Accesibilidad en las aulas de la web 2.0*. Bdigital Portal de Revistas UNAL. Volumen. <https://revistas.unal.edu.co/index.php/email/article/view/13970/14737>

[8] Convención Internacional sobre los Derechos de las Personas con Discapacidad, aprobada el 13 de diciembre de 2006 por la Asamblea General de las Naciones Unidas.

[9] Ley 26.653, *Accesibilidad de la Información en las Páginas Web*, 03 de noviembre de 2010. <https://www.argentina.gob.ar/normativa/nacional/ley-26653-175694>

[10] Cambours de Donini, A., Gorostiaga J. (2016). *Hacia una universidad inclusiva. Nuevos escenarios y miradas*. Aique Grupo Editor.

[11] Proyecto de investigación "Educación a distancia e innovación tecnológica". Convocatoria Núcleos de Organización y Ejecución de Actividades Científicas y Tecnológicas (NACT). Secretaría de Investigación e Innovación de la UNSAdA (Expte. 044/2022)

DISEÑO DE HERRAMIENTAS TECNOLÓGICAS EN MODELOS EDUCATIVOS HÍBRIDOS

Mg. Ing. G. Sergio Navas¹, Dra. Consuelo Escudero²,
Prof. Paola Beatriz Baiutti³, Diego A. Bustos⁴

Gabinete de Computación / Facultad de Ingeniería / Univ. Nacional de San Juan¹
Dpto. de Física / Facultad de Ingeniería / Univ. Nacional de San Juan^{2,4}
Facultad de Ciencias Sociales / Universidad Nacional de San Juan³

Av. Libertador Gral. San Martín 1109 (oeste) - San Juan
0264-4211700 (Int. 285¹ / 435¹)
snavas@unsj.edu.ar¹, cescudero@unsj-cuim.edu.ar², paolabaiutti@gmail.com

RESUMEN

La búsqueda de mejores estrategias didácticas y con la incorporación de aulas virtuales a los cursos regulares, nos lleva a plantear y generar recursos educativos más adecuados, que utilicen al máximo los medios tecnológicos actuales, que respondan a temas que presentan dificultades de aprendizaje y que además procuren flexibilidad para ser aplicados en distintos cursos.

Dentro de todo el conjunto de nuevas tecnologías, rescatamos dos herramientas como las más adecuadas para ofrecer un enfoque al actual modelo educativo, atravesado por una perspectiva desde la hibridación. Ya desde principios del año 2022 habíamos incursionado en la tarea de construir un Objeto Virtual de Aprendizaje (OVA) amplio, abarcativo, que contempla una unidad completa de un tema de física introductoria. En el contexto de las Tecnologías Digitales de la Información y la Comunicación (TDIC) también habíamos irrumpido en el diseño y puesta en marcha de un prototipo que a futuro será un Laboratorio Remoto (LR) en función en la UNSJ, específico para realizar experiencias en Física enlazadas a la Red Argentina Colaborativa de Laboratorios de Acceso Remoto.

Palabras clave: Aprendizaje híbrido - Objeto virtual de aprendizaje - Laboratorio remoto.

CONTEXTO

El presente tratado expone nuevas metodologías propuestas y en curso de implementación, a nivel de software & hardware, con enfoque orientado a la educación en ambientes híbridos. Incluye el desarrollo de software y de equipos tecnológicos ad-hoc, generando herramientas educativas innovadoras. Lo que se describirá se encuentra enmarcado en dos proyectos de investigación, uno en terminación, 2020/2022, más uno recién propuesto y en progreso para 2023/2024, en el ámbito de la Fac. de Ing. y financiados por la UNSJ.

Tales proyectos están conformados por equipos de investigación interdisciplinarios, de la Facultad de Ingeniería, de la FCFN y de la Fac. de Ciencias Sociales de la UNSJ.

Este trabajo a su vez se acopla con otro proyecto aprobado por la Red Argentina Colaborativa de Laboratorios de Acceso Remoto (CONFEDI-RLAB¹, financiado por SPU).

1. INTRODUCCIÓN

El proceso de digitalización puso en evidencia la escasez de estrategias y recursos para la enseñanza remota de las ciencias naturales en todos los niveles educativos, especialmente en su rama experimental. Derivada de la pandemia COVID 19 que afectó el desa-

¹ <https://confedi.org.ar/red-argentina-colaborativa-de-laboratorios-de-acceso-remoto-confedi-r-lab/>

rrrollo de las actividades educativas presenciales, durante el año 2020, se trabajó intensamente para afrontar la emergencia educativa, con el fin de garantizar la continuidad pedagógica; ha quedado claro la imperiosa necesidad de diseñar nuevos artefactos y dispositivos fundamentados en marcos metodológicos actuales, que incluyan la tecnología como mediadora para resolver problemas de esta nueva realidad donde lo remoto ha llegado para quedarse. Dos años después, el retorno a la presencialidad conlleva nuevos desafíos y también oportunidades.

Lo habitual en la literatura analizada es que los Objetos de Aprendizaje² (OA) se desarrollen independientes del contexto. Nuestra perspectiva es que los ambientes educativos y las tecnologías deben integrarse.

Dentro de todo el conjunto de nuevas tecnologías, se destaca los sistemas hipermediales, particularmente los *Objetos Virtuales de Aprendizaje* (OVA), que son un caso particular del OA, pero aplicado al aprendizaje virtual, hace referencia normal y actualmente a un objeto software reutilizable que cumple determinadas características.

Un producto OVA de desarrollo propio ya se encuentra en fase funcional, en al menos cuatro aulas virtuales (cuatro asignaturas de física, de ingeniería y de Lic. en Biología); y aunque recién se están recabando datos a efectos de obtener cifras más certeras para ponderar sus resultados con propósitos estadísticos-comparativos, ya se conoce que dio muy buenos resultados, dado que se hicieron encuestas ad-hoc a los alumnos, durante primer y segundo semestre de 2022 [1]. Bajo la premisa de tales resultados positivos obtenidos con el prototipo ya implantado, es que vemos como un camino viable y muy apropiado el proseguir realizando análisis, diseños y desarrollos bajo esa misma línea, para otras temáticas álgidas en cuanto a su enseñanza y a su aprendizaje, y propicias de abordar; además sabiendo que se trata de un recurso pro-

metedor pero que, a su vez, está aun absolutamente subutilizado.

Una primera búsqueda ha arrojado poca utilización de OA y OVA para mejorar la comprensión de contenidos y la motivación del alumno. Se evidencia también un alto grado de desconocimiento acerca del tema por parte de los docentes [2]. Por otra parte, los desarrolladores habitualmente hacen los OVA independientes del contexto, y en el desarrollo de sus diseños no se suele trabajar con equipos interdisciplinarios. Efectivamente, se constituye ésta en un área de vacancia.

La característica de “recurso compartible”, que puede ser usado y re-usado en contextos diversos, es una de las cualidades esenciales de los OVA. Su creación se considera una opción atractiva y válida para dar respuesta a las necesidades actuales de enseñanza.

A su vez, y sumado a lo anterior, el abanico de estrategias que se puede desplegar para incorporar *actividades experimentales* a propuestas de enseñanza digital incluye la realización de Actividades Experimentales Simples (Laboratorios Caseros), el uso de teléfonos inteligentes como instrumentos de medición (Laboratorios Móviles), Laboratorios Virtuales y *Laboratorios Remotos*. Estos últimos son de particular interés en la educación superior y se los considera potentes herramientas para promover el aprendizaje de competencias científicas y de procedimientos intelectuales y sensorio-motores propios del quehacer experimental y del ejercicio profesional [3].

Desarrollo de ambientes educativos asociados a tecnologías: OVA y LR

El continuo avance de la tecnología y la mejora en la conexión a Internet están generando nuevos espacios educativos en los que la integración de las Tecnologías Digitales de la Información y la Comunicación (TDIC) tienen un rol protagónico.

La búsqueda de mejores estrategias didácticas que incrementen el rendimiento de los alumnos, y con la incorporación de aulas virtuales a los cursos regulares, nos lleva a plantear y generar recursos educativos más adecuados, que utilicen al máximo los medios

² Un objeto de aprendizaje se define como “una entidad, digital o no digital que puede ser utilizada, reutilizada o referenciada durante el aprendizaje apoyado en tecnología” (IEEE, 2002)

tecnológicos actuales, que respondan a temas que presentan dificultades de aprendizaje y que además procuren flexibilidad para ser aplicados en distintos cursos.

Como docentes de la Fac. de Ingeniería y de la Fac. de Ciencias Exactas, Físicas y Naturales, motivados por el interés de elaborar materiales educativos hipermediados, particularmente software Web, con cualidades interactivas y que incluyan artefactos multimedia, capaces de provocar una verdadera evolución pedagógica, ya desde principios del año 2022, habíamos incursionado en la tarea de construir un Objeto Virtual de Aprendizaje no elemental, es decir amplio, abarcativo, que contempla una unidad completa de un tema de física introductoria, concretamente de cinemática de la partícula [1].

A continuación, se caracteriza sumariamente la creación de dicho OVA estratégico para la enseñanza y el aprendizaje de Física I, más propiamente de Cinemática, aplicado (por ahora) a carreras de ingeniería y de licenciatura en biología de la Universidad Nacional de San Juan. Se destaca que no se trata de un OVA elemental, ya que abarca la mayor parte del contenido de esa unidad de aprendizaje, con determinado nivel de profundidad preestablecido; contiene una importante cantidad de objetos menores reutilizados y recursos tales como: animaciones, videos y simulaciones, estratégica y pedagógicamente insertadas. La maquetación, los conceptos vertidos, ecuaciones, imágenes, juegos de autoevaluación, múltiples tests interactivos, etc. son de diseño propio. Se ha buscado contribuir con la construcción de un camino para otras experiencias y reflexiones acerca de estrategias educativas en un contexto de cultura digital y pos-pandémico.

Su implantación, en fase beta, fue realizada en 2022 en dos aulas virtuales a modo de prototipo, actualmente, ya más depurado y mejorado, se halla en cuatro aulas virtuales, mostrando muy buena aceptación por parte de los alumnos. Se espera ampliar el rango de utilización en los años subsiguientes. Se ofrece así la oportunidad al estudiante de realizar un aprendizaje interactivo, autónomo, dinámico y personalizado [4].

Una segunda herramienta, con la que venimos trabajando y considerando como protagonista innovadora en la inclusión de tecnologías educativas híbridas, son los Laboratorios de Acceso Remoto (LR).

En el marco de la implementación de un LR en la UNSJ vinculado a la Red Argentina Colaborativa de Laboratorios de Acceso Remoto, se ha diseñado y puesto en marcha, en sus etapas iniciales, un prototipo básico para realizar una experiencia en Física³. Representa un experimento que no se ha implementado antes en la Univ. Nacional de San Juan, de ahí su mayor potencial [5].

El laboratorio tradicional de experimentación requiere la presencia física de personas para la manipulación de los sistemas de control e instrumentos de un laboratorio, en un entorno controlado bajo supervisión.

Un laboratorio remoto (LR), en cambio, es un conjunto de dispositivos físicos reales, situados en determinadas instituciones, dotados de instrumentos, sensores, motores, cámaras de video, etc., de modo que puedan ser manipulados a distancia a través de internet

Los primeros pasos dados en la puesta en marcha de R-LAB en la UNSJ en el marco de la RED CONFEDI entrañan la generación de un grupo de trabajo GLABUNSJ que se inicia en agosto de 2021 buscando contribuir a una propuesta abierta y sólida de implementación de una red de laboratorios remotos al interior de la Fac. de Ingeniería de UNSJ, con posibilidad de integración con RLAB. Algunos desarrollos locales se remontan a un servicio integrado con la plataforma Moodle: ReLab (Laboratorio Remoto de Control Automático) - UNSJ en 2011. En este tratado se proveen avances de una actividad experimental que está en implementación en Física I sobre el tópico colisiones reales en la que los docentes desarrollan intervenciones didácticas, materiales y evaluaciones, mientras los estudiantes controlan, miden y grafican diferentes parámetros en contexto real pero remoto [5].

³ <https://drive.google.com/file/d/13sz2CgvfBO2moMI9iYtiJ-8SfTKAdor4/view>

La desvalorización de la medición y los sistemas de unidades es una de las señales que indica la existencia de un aprendizaje superficial. Desde la concepción de magnitud, el estudiante comienza a interactuar informalmente con los objetos y si esta práctica no se continúa, queda en un nivel muy elemental. Es por ello, que el objetivo es identificar las estrategias y obstáculos que se presentan al diseñar un laboratorio remoto.

En una colisión, para poder simular fuerzas de impacto, es preciso especificar un valor para el coeficiente de restitución ε . A tal fin, realizamos mediciones en los rebotes de una pelotita sobre una superficie estática, usando un sensor piezoeléctrico controlado con una placa Arduino NANO y acoplada a un PC. Se desarrolló el software para muestreo, cálculo y graficación requerido y se obtuvieron gráficas de intensidad de la perturbación elástica, que permitieron calcular el tiempo entre rebotes con una buena precisión (0.33%)

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación corresponde a la temática: *Tecnología Informática Aplicada en Educación*. Parte de los integrantes de este equipo de investigación tienen como principal antecedente, el haber participado en proyectos conjuntos o en actividades de formación de posgrado, contando mayoritariamente con artículos en colaboración. Muchos en el marco de la Universidad Nacional de San Juan (UNSJ) –en varios proyectos canalizados a través de CICITCA, CONICET y FONCYT–, con distintas universidades argentinas y extranjeras.

Actualmente mantienen tres proyectos en curso, uno de ellos presentando ya su finalización, todos involucrados al estudio, investigación y desarrollo de tecnologías educativas. Han realizado numerosas publicaciones y presentaciones, un par se ha citado en la bibliografía.

3. RESULTADOS OBTENIDOS

En cuanto al software Web (OVA), ha sido recientemente implantado en aulas virtuales, por tanto es aun breve su período de servicio y

utilización, tan solo una cohorte ha podido accederlo, ergo no hay aún comparativas estadísticas. Sin embargo sí se conocen niveles de virtuosidad valorados por la alta aceptación que ha tenido por parte de los alumnos. Esto contribuirá seguramente a la eficacia del aprendizaje y a la permanencia estudiantil.

En la Fig. 1 se puede observar una captura de pantalla, donde se distingue un menú de opciones navegacional semi-desplegado, en un ítem seleccionado. Para mayores detalles, a modo de muestra, fuera del aula virtual, se ha instalado una versión del prototipo de este OVA que se puede apreciar ampliamente en <https://cinematicaova.000webhostapp.com/>



Fig. 1 - Captura de pantalla de la porción inicial del tratado general de Velocidad en el OVA

En cuanto LR, la posibilidad hoy de uso de tecnologías multimedia desde la misma plataforma Moodle junto al compromiso de la Institución en la adquisición de equipos para realizar streaming desde laboratorios, están dando sustento a nuevas propuestas.

A la fecha, se ha conseguido desarrollar localmente un sistema para experimentar choques reales, que consta de un sensor piezoeléctrico y una placa Arduino Nano, que conectados a una PC pueden medir tiempo e intervalos de tiempo entre rebotes (de una esfera) del orden de los milisegundos o menores, con una buena precisión. Se utilizan esferas de distintos materiales (acero inoxidable, vidrio, madera, PVC, etc.)

Cuando la esfera golpea el piso, la vibración se transmite al sensor que transfiere impulso a un cristal piezoeléctrico. El cristal

genera un voltaje, dependiente de la carga, que se transmite a una computadora a través de un amplificador. Por software se genera un diagrama en la pantalla acerca de la amplitud de la señal en un lapso de 5 segundos o más, en el que picos y sus derivadas determinan máximos y mínimos; y también, la coordenada de cada pico y el intervalo de tiempo que la pelota pasa en el aire entre dos rebotes sucesivos, siendo $\Delta t_n = 2t_0 \varepsilon n$, donde t_0 es el tiempo que demora en golpear por primera vez la superficie cayendo desde una altura h_0 , y ε es el coeficiente de restitución, magnitud adimensional que ofrece un parámetro sobre cuánta energía fue disipada en una colisión.

Se ha desarrollado un software para PC que controla la placa Arduino, toma las muestras digitalizadas, genera tablas de datos, realiza procesamiento y emite la gráfica $\ln(\Delta t_i)$ vs. n_i (número de rebotes) para su interpretación, también calcula y muestra la recta de regresión, véase Fig. 2. A partir de la ecuación de la recta de ajuste, se extraen los parámetros del fenómeno y se calcula el coeficiente de restitución, t_0 y su error.

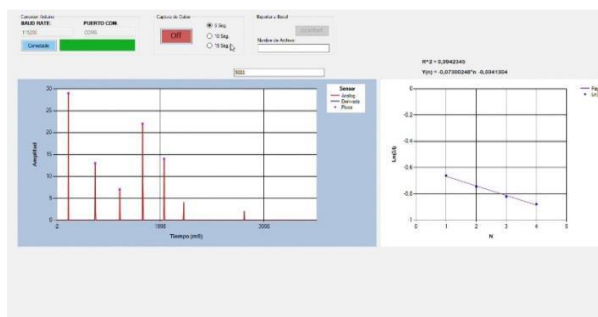


Fig 2 - Señal obtenida en sucesivas colisiones y gráfica de $\ln(\Delta t_i)$ vs n_i .

4. FORMACIÓN DE RECURSOS HUMANOS

En la temática de diseño de herramientas tecnológicas inscriptas en ambientes virtuales, de la mano de innovaciones educativas, los integrantes del grupo de investigación vienen trabajando desde el año 2018.

Desde principios de 2022 incursionamos en la tarea de construir OVAs amplios, abarcativos; y en el contexto de la TDIC. También acometimos en el diseño y puesta en marcha de un prototipo a futuro de LR en la UNSJ

para realizar experiencias en Física, enlazado la RED Arg Colaborativa de Laboratorios de Acceso Remoto.

El grupo de trabajo incluye alumnos avanzados, becarios, investigadores en formación e investigadores formados. Actualmente se avanza habiendo sumado algunos miembros más, que tienen en curso becas y tesis en doctorados, maestrías y especializaciones.

5. BIBLIOGRAFÍA

- [1] Navas, G. S.; Escudero, C. and Zalazar, D. (2022) "Virtual learning object as an enriched resource for teaching kinematics", IEEE Biennial Congress of Argentina (ARGENCON), 2022, pp. 1-8, doi: 10.1109/ARGENCON55245.2022.9939778. <https://ieeexplore.ieee.org/document/9939778>
- [2] Fernández, M., García, D., Erazo, C., & Erazo, J. (2020) "Objetos Virtuales de Aprendizaje - Una estrategia innovadora para la enseñanza de la Física". Revista Arbitrada Interdisciplinaria Koinonía. URL https://www.researchgate.net/publication/349199211_Objeto_Virtuales_de_Aprendizaje_Una_estrategia_innovadora_para_la_ensenanza_de_la_Fisica.
- [3] Moralez R., Ramirez R. et al. (2019). Virtual/Remote Labs for Automation Teaching: a Cost Effective Approach. ELSEVIER. IFAC-PapersOnLine, Volume 52, Issue 9, Pages 266-271. URL <https://www.sciencedirect.com/science/article/pii/S2405896319305567>
- [4] Gil Vera, V. (2019) "Objetos virtuales de aprendizaje en la enseñanza de la física: un caso de aplicación". Revista Internacional de Aprendizaje en Ciencia, Matemáticas y Tecnología 5(1), 9-17 [en línea]. URL <https://journals.gkacademics.com/revEDUMAT/article/view/1856>
- [5] Escudero, C.; Zalazar, D. y Navas, S. (2022) "Hacia la implementación de R-LAB en UNSJ: escenarios en Física y algo más". 9º Seminario Internacional de Educación a Distancia. Escenarios inéditos en la Educación Superior Perspectivas, huellas y emergentes. RUEDA, UNMDP, 14-15 Nov.

RECURSOS EDUCATIVOS ACCESIBLES. ESTUDIO DE DISTINTAS TÉCNICAS Y HERRAMIENTAS BASADAS EN INTELIGENCIA ARTIFICIAL

Javier Díaz, Alejandra Schiavoni, Paola Amadeo, Ivana Harari,
Alejandra Osorio

LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

Facultad de Informática. Universidad Nacional de La Plata

Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528

jdiaz@unlp.edu.ar , ales@info.unlp.edu.ar , pamadeo@linti.unlp.edu.ar, iharari@info.unlp.edu.ar,
aosorio@cespi.unlp.edu.ar

RESUMEN

El presente artículo describe el análisis de diferentes herramientas y técnicas para accesibilizar recursos educativos. En el ámbito educativo, existen los llamados Recursos Educativos Abiertos (Open Educational Resources-OER), sin embargo no siempre están específicamente publicados para estudiantes con discapacidades. Los docentes e investigadores deben poner más énfasis en el desarrollo de recursos textuales y multimediales preparados para que puedan ser accedidos por todas las personas, siguiendo los lineamientos de accesibilidad propuestos por normas internacionales. El objetivo es presentar distintas tecnologías de accesibilización de contenido que es utilizado en ambientes educativos, ya sea a través de plataformas o compartido en forma directa con los estudiantes. Desde hace un tiempo existen un conjunto de herramientas que permiten que un documento textual, una imagen o un video puedan ser accesibles e interpretados por ejemplo por un lector de pantalla. Actualmente comenzaron a emerger otras tecnologías, aplicaciones y servicios, basadas en Inteligencia Artificial que posibilitan hacer accesibles varias aplicaciones y recursos. Estas tecnologías están permitiendo ir más allá del acceso al contenido digital, ofreciendo nuevas formas de compromiso con el mundo.

Palabras clave: accesibilidad, recursos educativos, recursos accesibles

CONTEXTO

En el LINTI, Laboratorio de Investigación en Nuevas Tecnologías Informáticas de la Facultad de Informática de la UNLP, se vienen llevando a cabo varias líneas de investigación sobre Accesibilidad Web y su aplicación en diferentes ámbitos, considerando plataformas de aprendizaje abiertas y recursos educativos accesibles. En la Facultad de Informática se utilizan plataformas de código abierto para las gestiones académicas desde hace más de quince años, incluyendo sistemas de gestión de aprendizaje como Moodle, repositorios abiertos y sistemas de gestión administrativa, como SIU Guaraní. También se realizan continuos mecanismos de integración entre las diferentes plataformas, lo que permite una mayor flexibilidad y aprovechamiento en el uso de las mismas.

Las líneas de investigación que se mencionan en este artículo se desarrollan en LINTI de la UNLP y está enmarcado en el proyecto I+D 11/F028 “De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región”, aprobado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz. Este proyecto hace hincapié entre otros puntos en la formación de competencias y

habilidades digitales para todos los ciudadanos, en una sociedad que se plantea como digital.

Dada su relevancia, el tema de accesibilidad web se viene trabajando en la Facultad desde el año 2002, y se incorporó esta temática en el plan de estudios de las carreras que se dictan en la institución, a través de la asignatura Diseño Centrado en el Usuario. El libro “Guía de recomendaciones para diseño de software centrado en el usuario” registra aproximadamente 10.000 descargas desde el año 2013 [1]. También se institucionalizó su abordaje mediante la creación de una Dirección de Accesibilidad desde el año 2010, se desarrollaron tesinas, trabajos de cátedra al respecto, proyectos de innovación y desarrollo con estudiantes de Informática y diferentes proyectos de extensión acreditados por la Universidad Nacional de La Plata. Los proyectos de los últimos años son “Trabajando por una Web Accesible”, “Por una Web Inclusiva”, “Por una Web inclusiva: abordaje en escuelas secundarias”, que entró en el Programa de promoción de derechos y fortalecimiento de la organización comunitaria [2]. También, se aprobó el Proyecto de Desarrollo e Innovación sobre “Rampas Digitales Innovativas para Personas con Discapacidad”, en la convocatoria Agregando Valor 2018-2019, del Ministerio de Ciencia, Tecnología e Innovación de la Nación Argentina. Todos los años en la semana de la discapacidad se realizan jornadas en las que se exponen los trabajos realizados y se intercambian experiencias entre los diferentes grupos de otras universidades.

1. INTRODUCCIÓN

Dentro del contexto educativo, el concepto de accesibilidad significa que todos los estudiantes tengan las mismas condiciones durante el trayecto que transitan, incluyendo el contenido educativo y el proceso de enseñanza, los que deberían ajustarse a las necesidades de los estudiantes.

En el ámbito educativo, hay un consenso general que es necesario analizar y plantear políticas de accesibilidad en plataformas que albergan OER, lo que significa que los repositorios abiertos deben ser diseñados teniendo en mente el concepto de accesibilidad [3]. Sin embargo, en la actualidad la mayoría de los OER no están desarrollados y preparados en forma específica para estudiantes con algún tipo de discapacidad. En este contexto, Zhang [4] realizó un estudio de la situación de los OER y la accesibilidad y concluyó que los educadores deben prestar más atención en estas cuestiones, siguiendo las guías de la WCAG 2.0 (Web Content Accessibility Guidelines). Lo que ocurre con los recursos abiertos se traslada a todos los recursos educativos que se utilizan en el proceso de enseñanza en los diferentes niveles. Año a año, el número de estudiantes con discapacidad se va incrementando rápidamente en todo el mundo, sin embargo las instituciones educativas, escuelas y universidades no suelen satisfacer las necesidades de dichos estudiantes. En consecuencia, en gran medida abandonan sus estudios.

Si bien en este artículo nos vamos a enfocar en la accesibilidad de los recursos educativos, también es importante tener en cuenta las plataformas de aprendizaje (LMS) y los sistemas de videoconferencia que se comenzaron a utilizar masivamente en el dictado de los cursos durante la pandemia del COVID-19. El estudio y análisis de accesibilidad de estas herramientas viene siendo realizado por nuestro grupo de investigación desde hace varios años, como continuidad de las líneas de investigación propuestas en [5]. Se analizan y testean en forma continua el LMS Moodle en cada una de las nuevas versiones que se van instalando, la plataforma edX para los MOOCs y los sistemas de video conferencia utilizados para el dictado de las clases [6].

Actualmente existen varias herramientas propietarias y de software libre para validar la accesibilidad de los recursos textuales y multimediales, como por ejemplo, los videos;

estas herramientas en general requieren un testeo manual. Además, la corrección de errores y la construcción de recursos accesibles en la mayoría de los casos debe realizarse también en forma manual.

Desde hace ya muchos años han comenzado a realizarse estudios y desarrollarse técnicas y herramientas basadas en Inteligencia Artificial (IA), que permiten cambiar la forma en que aprendemos, interactuamos y nos comunicamos. También en el contexto de accesibilidad, la IA puede ser aplicada. Se trata de democratizar la IA teniendo una mayor injerencia en las prácticas laborales cotidianas, motivadas por la preocupación de que la capacidad de utilizar técnicas de IA no quede en un campo de desarrollo minoritario [7]. En base a esto se están implementando herramientas que ayudan a hacer testeos y validaciones que reducen significativamente el trabajo manual.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Siguiendo con la línea de trabajo que venimos describiendo, las tareas de investigación, desarrollo e innovación que se llevan a cabo en este proyecto abordan los temas relacionados a la accesibilidad en las herramientas utilizadas en educación y en los recursos que se les brinda a los estudiantes de manera que todos transiten una experiencia educativa equitativa, incluyendo a los que cuentan con algún tipo de discapacidad. Esto significa que tanto el contenido como el proceso de enseñanza se adapten a las necesidades de todos los estudiantes. Es importante pensar en crear un futuro digital inclusivo, y para ello hay que considerar las tecnologías basadas en IA y cómo se relacionan con la accesibilidad. La Fig 1 muestra el impacto de la IA en las herramientas que se pueden desarrollar para las distintas

discapacidades y para optimizar la validación de los contenidos.



Fig. 1: El impacto de IA en Accesibilidad: Ilustración del Grupo de UX de LenixDesign

Actualmente, muchas tecnologías están dirigidas a la inclusión digital, incluso desarrolladas específicamente para personas con discapacidad [8]. Algunas de estas tecnologías son:

- control de voz para computadoras y móviles, destinados a personas con movilidad reducida y alguna discapacidad motriz
- agregado automático de traducciones y subtítulos para personas con discapacidad auditiva, Automatic Speech Recognition ASR.
- reconocimiento de imágenes y agregado de texto alternativo para personas que necesitan utilizar un lector de pantalla
- la posibilidad de acortar y resumir artículos para personas con dificultades de lectura
- Incorporación de reconocimiento facial como modo de autenticación para facilitar el acceso a los dispositivos.

Algunas empresas están utilizando Machine Learning para mejorar el testeo automático de accesibilidad e incentivar a los desarrolladores de productos a validar los principios de accesibilidad y crear contenido accesible. Empresas como Amazon, Google y Facebook entre otras han hecho grandes avances en este tema y sostienen que cuando se agregan datos específicos de dominio, es el primer paso que

constituye un gran avance que reduce significativamente la cantidad de pruebas manuales. El objetivo es continuar utilizando estas técnicas para ofrecer cada vez más y de manera constante más automatización y ahorro de tiempo en el futuro.

En esta línea vamos a continuar investigando las distintas formas en que la IA puede intervenir en los tests de accesibilidad tanto de las aplicaciones como de las páginas web. En términos prácticos, eso significa que cuando probamos un sitio web para verificar que cumpla con los requisitos de estándares como WCAG y brindamos orientación para remediarlo, se debe ser conscientes del papel de la IA considerando la posibilidad de que se genere algún problema de accesibilidad como de que se resuelva. En la web, la IA puede resolver problemas de accesibilidad y puede crear problemas de accesibilidad, a veces simultáneamente.

Actualmente, las herramientas para ASR [9] utilizan inteligencia artificial para reconocer y transcribir el lenguaje hablado a texto escrito. En particular se utilizan deep neural networks (DNNs), que se entrenan con un amplio dataset de audios, acentos, estilos y finalmente se predice la transcripción del texto escrito. Es más sencillo y adaptable que métodos anteriores pero suele llevar más tiempo y requerimientos de procesamiento. Entre las herramientas que se han utilizado para realizar transcripciones del material de distintas materias de las carreras que se dictan en la Facultad podemos mencionar:

veed.io [10]: portal que brinda distintos productos para la edición de videos, en particular el agregado de subtítulos en forma automática. Brinda distintos tipos de licencias.

auris AI [11]: portal que brinda transcripciones y subtítulos en línea. Se pueden descargar en distintos formatos. Brinda distintos tipos de licencias.

Subtitle edit [12][13]: editor de subtítulos de videos open source. Brinda la posibilidad de descargar librerías para el subtítulo a través de IA. No es tan potente como los mencionados

previamente pero acelera el proceso manual en forma significativa.

Cada una de ellas posee sus ventajas y desventajas que se están estudiando, junto con otras que aparecen en el mercado.

Finalmente no podemos dejar de mencionar distintos proyectos que utilizan deep learning para el reconocimiento de la lengua de señas. Es un tema complejo que posee múltiples aristas de estudio [14] [15].

3. RESULTADOS Y OBJETIVOS

Continuando con la evaluación de herramientas y tecnologías para la validación y testeo de la accesibilidad en plataformas y recursos educativos, y analizando nuevas propuestas y desarrollos que incluyen la IA como un aporte importante en este tema, se plantean los siguientes objetivos:

- Análisis de las normas de accesibilidad con todas sus actualizaciones para la creación de contenido web, teniendo en cuenta las pautas a cumplir según la WCAG 2.1 [16]
- Estudio de las nuevas normas internacionales para realizar recursos multimediales accesibles [17]
- Continuación con la evaluación de las distintas herramientas para subtítulos de videos, que resultan de suma utilidad y no son comúnmente usadas en la generación de los mismos. En esta línea analizamos las herramientas para la grabación de clases y los distintos formatos de subtítulos.
- Estudio de herramientas existentes que utilizan IA para accesibilidad tales como ChatGPT [18], Seeing AI [19], es una aplicación que cumple con distintas tareas, Axe [20], NavCog [21]

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por tres profesoras de amplia trayectoria en el campo de la investigación, que trabajan en el área de ambientes virtuales de aprendizaje y accesibilidad web. Además, una alumna becaria de la carrera de Licenciatura en Informática de la Facultad de Informática.

La participación en eventos de la especialidad, ha permitido seguir estableciendo canales de comunicación con otros investigadores que trabajan en las mismas áreas.

Las Jornadas de Accesibilidad que se realizan todos los años permiten intercambiar experiencias entre los participantes del curso y exponer los trabajos realizados. El video de la Jornada del año 2022 denominada *Trabajos de estudiantes y experiencias de inclusión en la Facultad de Informática de la UNLP.*, se encuentra disponible en: <https://www.youtube.com/watch?v=NwDaxwXpsCw>.

5. BIBLIOGRAFÍA

[1] Díaz, J., Amadeo, P., Harari, I. “Guía de recomendaciones para diseño de software centrado en el usuario”. Editorial de la Universidad Nacional de La Plata (EDULP). ISBN: 978-950-34-1030-1. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/32172>

[2] Dictamen Proyectos de Extensión Acreditados en la UNLP. https://unlp.edu.ar/proyectosex/proyectos_de_extension_acreditados_y_subsidiados-4708

[3] Law, P., Perryman, L.-A., and Law, A. “Open educational resources for all? Comparing user motivations and characteristics across The Open University’s iTunes U channel and OpenLearn platform.” (2013).

[4] Zhang, X., Tlili, A., Nascimbeni, F. *et al.* Accessibility within open educational resources

and practices for disabled learners: a systematic literature review. *Smart Learn. Environ.* 7, 1 (2020). <https://doi.org/10.1186/s40561-019-0113-2>

[5] Díaz J., Harari I., Schiavoni A., Amadeo A., Gómez S., Osorio, A.: “Herramientas Digitales para Educación. Análisis de su accesibilidad”. Publicado en el Libro de Posters del XXIV Workshop de Investigadores en Ciencias de la Computación - WICC 2022, Abril 28-29, Mendoza, Argentina. La edición - Godoy Cruz : FUSMA Ediciones, 2022. ISBN: 978-987-48222-4-6

[6] Díaz J., Harari I., Schiavoni A., Amadeo A., Gómez S., Osorio, A.: “Aporte para pensar la educación en pandemia desde la accesibilidad.” Proceedings del Congreso Argentino en Ciencias de la Computación CACIC 2021 compilación de Marcia I. Mac Gaul. - 1a ed ISBN 978 -987-633-574-4 (2021).

[7] Wolf C. “Democratizing AI?: experience and accessibility in the age of artificial intelligence”. XRDS: Crossroads, The ACM Magazine for Students, Volume 26, Issue 4, Summer 2020, pp 12–15 <https://doi.org/10.1145/3398370>

[8] Xiong, M. How AI can influence Accessibility. Published in LexisNexis Design (2020). Disponible en: <https://medium.com/lexisnexis-design/how-ai-can-influence-accessibility-aadd6a398996>

[9] What is Automatic Speech Recognition: Our guide to ASR <https://aurisai.io/blog/what-is-automatic-speech-recognition-asr/>

[10] VEED.IO <https://www.veed.io/tools/add-subtitles>

[11] Auris AI <https://aurisai.io/subtitle/>

[12] <https://www.nikse.dk/subtitleedit>

[13] Cómo crear subtítulos con inteligencia artificial y gratis con Subtitle Edit
<https://www.creatubers.com/como-crear-subtitulos-con-inteligencia-artificial-y-gratis-con-subtitle-edit/>

[14] Ramírez Cerna, Lourdes (2021) *Reconocimiento de Lengua de Señas usando Deep Learning*. Woman in Data org
<https://www.youtube.com/watch?v=M62Lt4rzXs4>

[15] Ramirez Cerna, Lourdes; Escobedo Cardenas, Edwin; Garcia Miranda, Dayse; David Menotti, Guillermo Camara-Chavez (2021) *A multimodal LIBRAS-UFOP Brazilian sign language dataset of minimal pairs using a microsoft Kinect sensor*
https://scholar.google.com/citations?view_op=view_citation&hl=es&user=Neh8OZwAAAAJ&citation_for_view=Neh8OZwAAAAJ:2osOgNQ5qMEC

[16] Web Content Accessibility Guidelines (WCAG) 2.1.
<https://www.w3.org/TR/WCAG21/>

[17] W3C. Making Audio and Video Media Accesible.
<https://www.w3.org/WAI/media/av/>

[18] ChatGPT
<https://openai.com/blog/chatgpt/>

[19] Seeing AI
<https://www.microsoft.com/en-us/ai/seeing-ai>

[20] Axe <https://www.deque.com/axe/>

[21] NavCog
<https://www.cs.cmu.edu/~NavCog/navcog.html>

Experiencia de Usuario en Sistemas Interactivos usando reconocimiento de Emociones

L. Aballay¹, I. Millán Tejada², S. Aciar¹, L. Romera², L. Espinosa², S. Ledesma Vila¹

lnaballay@gmail.com, flavia.millan@gmail.com, saciar@gmail.com, lilianaromera@gmail.com, lvilmaespinosa@gmail.com, sergioledesma.9427@gmail.com

¹ Instituto de Informática, Facultad de Ciencias Exactas Físicas y Naturales, Universidad Nacional de San Juan

² Departamento de Informática, Facultad de Ciencias Exactas Físicas y Naturales, Universidad Nacional de San Juan

Resumen

La opinión de los usuarios son una valiosa herramienta de alerta para determinar qué funciona bien y qué no en un sistema. Contar con una metodología de evaluación de Experiencia de Usuario (UX) de sistemas interactivos, que integre mecanismos automáticos de percepción de emociones de los usuarios mientras trabajan en el sistema, ayudará a conocer las emociones de los usuarios mientras usan un sistema interactivo y de esta manera poder tomar acciones correctivas para mejorar la experiencia de usuario.

Palabras claves

Computación Afectiva, Experiencia de Usuario, Emociones

1. Contexto

En este trabajo se presenta la propuesta de investigación dentro del marco de un proyecto postulado a la convocatoria 2023-2024 del Consejo de Investigaciones Científicas y Técnicas y de Creación Artística (CICITCA) de la Universidad Nacional de San Juan (UNSJ). Las actividades de investigación se llevarán a cabo en las instalaciones del Instituto de Informática (IdEI) de la Facultad de Ciencias Exactas físicas y Naturales (FCEFN) de la UNSJ, en el Laboratorio de Tecnologías Adaptativas y Colaborativas (TAC).

2. Introducción

Conocer la opinión de los usuarios de sistemas interactivos se ha vuelto un elemento estratégico fundamental para empresas tales como: comercios, medios de comunicación (diarios, radios, canales de tv), bancos,

instituciones educativas, etc.. Estas organizaciones ofrecen sus productos por medio de sistemas interactivos, ya sean sitios web o aplicaciones móviles. Las percepciones de los usuarios son una valiosa herramienta de alerta para determinar qué funciona bien y qué no en un sistema. Conocer lo que el usuario opina permitirá orientar el diseño del sistema para satisfacer las necesidades del mismo, aumentando su productividad y eficiencia, consiguiendo que el sistema sea de preferencia para el usuario. Dada la importancia de conocer cómo se siente el usuario respecto a la interacción con sistemas interactivos informáticos, el desarrollo de tecnologías capaces de detectar emociones en forma automática a partir de un dispositivo de entrada ha experimentado un gran auge en estas últimas décadas [1]. Por este motivo, es necesario contar con una metodología de evaluación de Experiencia de Usuario (UX) de sistemas interactivos, que integre estos mecanismos de percepción de emociones de los usuarios mientras trabajan en el sistema.

Aplicar la metodología de evaluación UX propuesta ayudará a conocer las emociones de los usuarios mientras usan un sistema interactivo y de esta manera poder tomar acciones correctivas para mejorar la experiencia de usuario.

Se presentan a continuación los conceptos considerados como base teórica para el desarrollo de la investigación del presente proyecto.

2.1 HCI - Interacción Humano Computador

La Interacción Humano Computadora [3], que internacionalmente se conoce como Human-Computer Interaction (HCI), es un área

multidisciplinar que tiene sus raíces en la psicología cognitiva, sociología e informática, es el estudio de la relación que existe entre los usuarios humanos y los sistemas informáticos que usan para realizar diversas tareas [4].

La ACM (Association for Computer Machinery) define la HCI como: “Disciplina relacionada con el diseño, la evaluación e implementación de sistemas informáticos interactivos para el uso de los seres humanos, y con el estudio de los fenómenos más importantes con los que está relacionado”.

2.2 Computación Afectiva

La computación afectiva es el área que se centra en investigar cómo se usan las emociones en los sistemas informáticos, dada la necesidad de tomar en cuenta los estados emocionales de los usuarios a la hora de desarrollar software.

“desarrollo de sistemas inteligentes capaces de dotar a un ordenador con la habilidad de reconocer, interpretar, procesar y simular estados emocionales” (Rosalind Picard – MIT [5])

Las interfaces afectivas, que tratan de inferir las emociones de sus usuarios, se están desarrollando en un gran número de dominios [6] [7].

La computación afectiva describe métodos computacionales que están relacionados, derivados o diseñados a propósito para influir en las emociones ofreciendo una mejor interacción y experiencia de usuario teniendo en cuenta su estado emocional [8]. Por lo tanto, el reconocimiento de la emoción es uno de los pasos clave hacia la computación afectiva [9].

2.3 Experiencia de Usuario - UX

La UX posee varios aspectos que son importantes para evaluar en una aplicación de software. Incluye todas las emociones, creencias, preferencias, percepciones,

respuestas, comportamientos y logros de los usuarios. Que suceden antes, durante y después del uso de un producto [10]. UX extiende el concepto de usabilidad más allá de la eficacia, la eficiencia y la satisfacción.

De acuerdo con la norma ISO 9241-210, UX se puede definir como: “las percepciones y respuestas de la persona resultantes del uso y / o uso anticipado de un producto, sistema o servicio” [11]. Establece que UX “incluye todas las emociones, creencias, preferencias, percepciones, respuestas físicas y psicológicas, comportamientos y logros de los usuarios que ocurren antes, durante y después del uso”. Además, la norma ISO 9241-210 señala que UX “es una consecuencia de la imagen de marca, la presentación, la funcionalidad, el rendimiento del sistema, el comportamiento interactivo y las capacidades de asistencia del sistema interactivo, el estado interno y físico del usuario resultante de una experiencia previa, experiencias, actitudes, habilidades y personalidad, y el contexto de uso” [11].

Varios autores han propuesto aspectos, atributos o factores que definen la UX. Estos factores engloban las diferentes dimensiones y / o características que forman parte de UX.

Morville, propone siete facetas o factores de UX y las presenta en forma de panal de abejas para ayudar a las personas a comprender la necesidad de definir prioridades para ilustrar la UX: útil, utilizable, deseable, localizable, accesible, creíble y valioso [12].

Revang, ha sugerido varios factores que influyen en la UX [13], afirma que la experiencia de usuario es una serie de fases y que es importante centrarse en la positividad en los seis factores principales: usabilidad, utilidad, búsqueda, accesibilidad, credibilidad y deseabilidad. A su vez, cada factor se subdivide en otros factores, obteniendo un total de 30 factores.

Arhippainen y Tähti, han creado un modelo que muestra cinco elementos que influyen en

la UX [14]. Todos estos factores influyen en la experiencia que evoca la interacción del usuario con el producto: factores sociales (como la presión del tiempo), factores culturales (incluido el idioma y los hábitos), factores del producto (como funciones, usabilidad, movilidad, entre otros), factores del usuario (por ejemplo, emociones, habilidades, experiencias previas, motivación, entre otros) y contexto de uso (como tiempo y lugar).

Garret, ha propuesto varios elementos de la UX divididos en cinco planos [15]: superficie, esqueleto, estructura, alcance y estrategia. Estos cinco planos proporcionan un marco conceptual para discutir problemas de UX y las herramientas que se pueden utilizar para resolverlos. Estos planos no son atributos, sino que se presentan como una guía para el diseño de UX.

La UX abarca diferentes facetas relacionadas a la calidad de un producto software. El estándar ISO/IEC 25010 [16] considera de forma general las siguientes facetas de la UX: accesibilidad, dependabilidad, emotividad, jugabilidad, usabilidad, entre otras.

En la presente investigación se estudiará la percepción de los usuarios para evaluar UX según la faceta Emocionalidad.

2.4 Diseño Centrado en el Usuario - DCU

El diseño centrado en el usuario (DCU) es un proceso iterativo que se enfoca en la comprensión de los usuarios y su contexto en todas las etapas de diseño y desarrollo [17]. Busca construir productos involucrando a los usuarios en todo el proceso, resolviendo necesidades concretas, consiguiendo así la mayor satisfacción y mejor experiencia de uso posible.

Dentro de las necesidades que el sistema debe cubrir se encuentra: Eficacia, Eficiencia, Satisfacción (utilidad, confianza [18], placer, comodidad), ausencia de riesgo (seguridad [19], salud, medio ambiente, económica),

Confiabilidad, Seguridad [20], Cobertura de una variedad de contextos (cobertura, flexibilidad), Facilidad de aprendizaje, Accesibilidad.

2.5 Métodos detección de emociones

Reconocer las emociones que las personas pueden expresar durante el uso de los sistemas informáticos es una tarea compleja. Para tratar de automatizar esta detección, se han llevado a cabo diferentes investigaciones [21]. La mayoría de estos métodos se centran en la búsqueda, observación y análisis de las emociones generadas en los usuarios.

Se han llevado a cabo diversas investigaciones para reconocer estas emociones percibidas e interpretarlas con múltiples fines: educación, salud, seguridad, marketing, entre otros.

Así, el reconocimiento de emociones es uno de los pasos clave en su interpretación y se refiere a la capacidad de las máquinas para "percibir" la presencia de emociones en los seres humanos. Los humanos expresan sus emociones a través de varios canales: voz, texto, expresiones faciales, voces, gestos corporales, entre otros.

Cabe aclarar que los métodos de reconocimiento pueden ser de las siguientes categorías

- Detección implícita de emociones [22]: también llamados métodos de reconocimiento automático porque el usuario no interviene en la detección, ésta se realiza casi sin que el usuario se dé cuenta por medio de señales fisiológicas de las emociones. Son los que reconocen las emociones a través de las expresiones faciales [23], la voz [24], el texto [25], los latidos del corazón, las ondas cerebrales [26], etc.

- Detección explícita de emociones [22]: O métodos de reconocimiento no automáticos. Se consulta a la persona sobre sus emociones a través de cuestionarios o entrevistas, por ejemplo.

En este trabajo se hará una revisión de los diferentes métodos de detección para determinar el más adecuado o más factible de usar en la metodología de evaluación UX propuesta.

2.6 Emociones

El significado de “emoción” según diversas fuentes es:

- Del latín “*emotio*”, que significa “movimiento o impulso”, “aquello que te mueve hacia”.
- Reacciones psico-fisiológicas que representan modos de adaptación a ciertos estímulos cuando percibimos un objeto, persona, lugar, suceso, o recuerdo importante.
- Psicológicamente: las emociones alteran la atención, hacen subir de rango ciertas conductas guía de respuestas del individuo y activan redes asociativas relevantes en la memoria. Los sentimientos son el resultado de las emociones y pueden ser verbalizadas (palabras).
- Según el diccionario de la Real Academia Española: Emoción se define como una “alteración del ánimo intensa y pasajera, agradable o penosa, que va acompañada de cierta conmoción somática”.
- Según el Oxford English Dictionary define la emoción como “agitación o perturbación de la mente; sentimiento; pasión; cualquier estado mental vehemente o agitado”.
- Las emociones se experimentan a veces cuando algo inesperado sucede y los efectos emocionales empiezan a tener el control en esos momentos [2].

3. Formación de recursos humanos

En el marco del desarrollo del proyecto propuesto, se planea convocar a alumnos a

iniciar sus trabajos de fin de carrera de las carreras de Licenciatura en Ciencias de la Computación o de la Licenciatura en Sistemas de Información de la FCEFN de la UNSJ. En este sentido, se planea también postularlos en programas de becas de la UNSJ o Nación para alumnos avanzados. Por otro lado, también se propone convocar a alumnos de posgrado de la carrera de Maestría en Informática de la FCEFN de la UNSJ, para que realicen sus tesis en temas relacionados al proyecto. Adicionalmente, durante este proyecto se desarrollará una tesis de doctorado dentro del programa de doctorado en Ciencias de la Informática de la FCEFN de la UNSJ.

4. Resultados Esperados

Con los resultados de la investigación se realizarán publicaciones en congresos nacionales e internacionales, así como también se publicarán resultados más significativos en reconocidas revistas del área. Por otra parte, al momento de implementar la metodología de evaluación UX, es posible elegir como caso de estudio un sistema de cualquier ámbito: transporte, salud, educación, seguridad, marketing, por nombrar algunos. Realizar esta evaluación de la percepción de los usuarios será un aporte significativo para los diseñadores de estos sistemas, ya sea que pertenezcan a organismos públicos como a empresas privadas de desarrollo de software. Una vez hechas las evaluaciones a los sistemas, se le puede brindar a la empresa u organismo propietario del sistema un análisis de los resultados obtenidos para que puedan tomar sus medidas correctivas pertinentes en caso de ser necesarias.

Además, el proyecto está estrechamente vinculado con las carreras de informática de la FCEFN; ya que en la materia “Ingeniería de Software II”, de la Licenciatura en Sistemas y la Licenciatura en Computación, y “Diseño Web II”, de la Tecnicatura en Programación Web, tienen entre sus contenidos mínimos la temática HCI y Experiencia de Usuario, lo que

permitirá hacer participar a alumnos como posibles sujetos en pruebas experimentales con la implementación de la metodología de evaluación UX diseñada.

Referencias Bibliográficas

- [1] M. Imani and G. A. Montazer, "A survey of emotion recognition methods with emphasis on E-Learning environments," *J. Netw. Comput. Appl.*, vol. 147, no. April, p. 102423, 2019.
- [2] Gil-Olarte, P., Palomera, R. y Brackett, M., "Relating emotional intelligence to social-competence and academic achievement in high school students," vol. 18, supl., pp. 118-123.
- [3] T. Granollers, J. Lorés Vidal, and J. J. Cañas Delgado, "Introducción a la Interacción Persona Ordenador," in *La Interacción Persona Ordenador*, J. Lores, Ed. Editorial UOC-AIPO, 2002, p. 45.
- [4] C. Faulkner and Christine, *The essence of human-computer interaction*. Prentice Hall, 1998.
- [5] B. Kort, R. Reilly, and R. W. Picard, "An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion," in *Proceedings - IEEE International Conference on Advanced Learning Technologies, ICALT 2001*, 2001, pp. 43-46.
- [6] S. Baldassarri, "Computación Afectiva: tecnología y emociones para mejorar la experiencia de usuario."
- [7] R. W. Picard, A. Wexelblat, and C. I. N. I. Clifford I. Nass, "Future interfaces," in *CHI '02 extended abstracts on Human factors in computing systems - CHI '02*, 2002, p. 698.
- [8] Y. P. Cruz, C. A. Collazos, and T. Granollers, "The Thin Red Line Between Usability and User Experiences," in *Proceedings of the XVI International Conference on Human Computer Interaction - Interacción '15*, 2015, pp. 1-2.
- [9] L. Masip Ardévol, "User experience methodology for the design and evaluation of interactive systems," University of Lleida, 2013.
- [10] K. Finstad, "The Usability Metric for User Experience," *Interact. Comput.*, vol. 22, no. 5, pp. 323-327, Sep. 2010.
- [11] "ISO 9241-11:2018(en), Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts," 2018. .
- [12] P. Morville, "The 7 Factors that Influence User Experience," *Interaction Design Foundation*, 2018.
- [13] M. Revang, "User Experience Project: The User Experience Wheel." [Online]. Available: <http://userexperienceproject.blogspot.com/2007/04/user-experience-wheel.html>. [Accessed: 25-Aug-2020].
- [14] L. Arhippainen and M. Tähti, "Empirical Evaluation of User Experience in Two Adaptive Mobile Application Prototypes," in *Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia*, 2003, 2003, pp. 27-34.
- [15] J. J. Garrett, *The elements of user experience : user-centered design for the Web and beyond*, 2nd ed. New Riders, 2011.
- [16] ISO/IEC, "ISO/IEC 25010:2011 - Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models," 2011. [Online]. Available: <https://www.iso.org/standard/35733.html>. [Accessed: 03-Aug-2017].
- [17] Á. Villareal, Á. P. Villareal, A. F. Aguirre, C. A. Collazos, and R. Gil, "Propuesta Metodológica para la Inclusión del Aspecto Emocional Dentro del Diseño de Interfaces de un Entorno Virtual de Aprendizaje," *Rev. Colomb. Comput. - RCC*, vol. 15, no. 1, 2016.
- [18] Y. Wang, "Inclusive security and privacy," *IEEE Secur. Priv.*, vol. 16, no. 4, pp. 82-87, Jul. 2018.
- [19] P. C. Realpe, C. A. Collazos, J. Hurtado, and A. Granollers, "Towards an integration of usability and security for user authentication," *ACM Int. Conf. Proceeding Ser.*, vol. 07-09-September-2015, Sep. 2015.
- [20] J. Johnston, J. H. P. Eloff, and L. Labuschagne, "Security and human computer interfaces," *Comput. Secur.*, vol. 22, no. 8, pp. 675-684, Dec. 2003.
- [21] J. M. Garcia-Garcia, V. M. R. Penichet, and M. D. Lozano, "Emotion detection: A technology review," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1311, 2017.
- [22] M. Feidakis, *A Review of Emotion-Aware Systems for e-Learning in Virtual Environments*. Elsevier Inc., 2016.
- [23] A. Cherbonnier and N. Michinov, "The recognition of emotions beyond facial expressions: Comparing emoticons specifically designed to convey basic emotions with other modes of expression," *Comput. Human Behav.*, vol. 118, p. 106689, May 2021.
- [24] Z. Zhang, "Speech feature selection and emotion recognition based on weighted binary cuckoo search," *Alexandria Eng. J.*, vol. 60, no. 1, pp. 1499-1507, Feb. 2021.
- [25] K. Zahra, M. Imran, and F. O. Ostermann, "Automatic identification of eyewitness messages on twitter during disasters," *Inf. Process. Manag.*, vol. 57, no. 1, 2020.
- [26] J. Frey, C. Mühl, F. Lotte, and M. Hachet, "Review of the Use of Electroencephalography as an Evaluation Method for Human-Computer Interaction," 2013.

Un bot virtual en Moodle para asistir y complementar actividades académicas en espacios de educación superior

Gustavo Illescas¹, Exequiel Herrera Rosa², Guillermo Rodriguez³

¹Instituto de Investigación en Tecnología Informática Avanzada (INTIA-Centro Asociado CIC, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA).

²Becario EVC-CIN, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA).

³Instituto Superior de Ingeniería de Software Tandil (ISISTAN-CONICET), Facultad de Ciencias Exactas, UNCPBA.

{illescas, grodri}@exa.unicen.edu.ar, eherreraderosa@alumnos.exa.unicen.edu.ar

RESUMEN

En el presente trabajo se propone realizar un asistente virtual que complemente activamente la utilización de la plataforma de aprendizaje Moodle (Virtual Moodle Bot), contando con diversas funciones adecuadas al rol alumno o al rol docente de acuerdo a quien utilice la plataforma. La implementación de esta herramienta contará con dos instancias fundamentales: una primera instancia de investigación centrada en la intención de uso por parte de los usuarios y las tecnologías de desarrollo y, en una segunda instancia, se realizará la prototipación de un asistente virtual dentro de la plataforma mencionada. Además, La propuesta busca explorar el conocimiento generado y detectar tanto información oculta como así también patrones en los datos obtenidos, que asistan a los tomadores de decisiones como lo son los miembros de equipos de gestión de las unidades académicas, integrando así investigación con procesos de gestión. Este proyecto se encuentra dentro del marco de la beca EVC-CIN, el cual fue presentado y admitido en el año 2022.

Palabras clave: *Asistentes Virtuales, Educación Superior, Moodle, Analítica de Datos.*

CONTEXTO

En el nuevo medio digital se cuenta con una masiva cantidad de asistentes virtuales que

buscan aumentar la productividad de sus organizaciones (Rodriguez, et. al. 2020). Mayoritariamente son implementados en organismos donde sean prioritarias las labores de gestión y trivialidades asociados a la recolección de datos, tales como bancos, portales de compra y venta, redes sociales, entre otros (Ibáñez, et. al. 2021). Pese a su potencial, aún no se logra consolidar una nueva tendencia al cambio tanto en su formato como en su utilidad, por lo que este tipo de tecnología se encuentra estancada desde hace ya varios años. Aun así, gracias a la acelerada transformación digital producto de la pandemia, se han visto indicios de un posible cambio de paradigma, tales como el anuncio del Metaverso de Meta (Andrew, 2021). Meta restablece el concepto de asistentes virtuales a fin de complementar su utilidad conversacional con labores complejas y multifacéticas. Este tipo de proyectos buscan producir interacciones entre usuarios-asistentes en entornos virtuales cotidianos, con tal de ser beneficiados gracias a la presencia de agentes inteligentes.

Pese a las nuevas tendencias, en las plataformas educativas tales como Moodle se visibiliza la falta de este tipo de tecnologías para complementar la capacitación académica, la cual ha sido utilizada en otras instituciones

con grandes resultados (Roy, et. al. 2019; Chinedu, et. al. 2021). Es por ello que Virtual Moodle Bot (VMB) busca aportar un incentivo que acompañe este nuevo futuro de la era digital dentro de los espacios de educación superior (Dimitriadis, 2020; Bocanet, et. al. 2020), y para ello se debe hacer una buena distinción entre las dos labores objetivos que componen este trabajo.

1. INTRODUCCIÓN

Desde sus inicios, el proyecto fue planeado de tal forma que deba pasar por varias etapas de desarrollo, previas al inicio de su implementación e incluso posterior al mismo, con tal de exponer los resultados obtenidos. A modo de exposición, se presenta la Figura 1 con la primera parte del proceso de evolución del proyecto junto con cada una de las etapas involucradas, las cuales serán descritas a continuación.

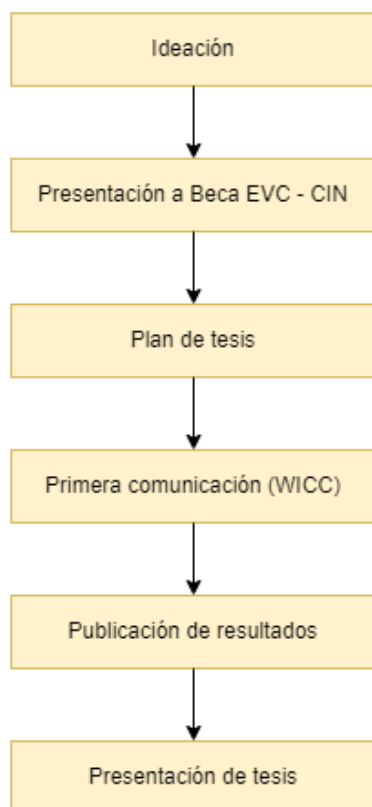


Figura 1: Evolución del proyecto VMB.

La primera etapa del proyecto, siendo el proceso de ideación, surge debido al interés mutuo de los involucrados en cuanto a su

experiencia dentro del entorno de aprendizaje Moodle, y su notable falta de agentes inteligentes. A partir de ello se planteó buscar una solución que logre complementar las actividades académicas de los usuarios de la plataforma, obteniendo una nueva base de conocimientos como resultado de estas interacciones. A partir de allí, se buscó enmarcar esta idea dentro del proyecto de incentivos 03/C314 titulado “Gestión Informática del Conocimiento como soporte para la toma de decisiones Organizacionales” desarrollado en el Instituto de Tecnología Informática Avanzada (INTIA) de la Facultad de Ciencias Exactas (EXA) UNCPBA, siendo parte de los colaboradores de VMB miembros activos dentro del mismo.

Continuando con la segunda etapa, este proyecto fue presentado y finalmente avalado bajo el marco de la beca de incentivo a la investigación EVC-CIN 2022, a manos del becario Exequiel Herrera De Rosa, director Dr. Gustavo Illescas y Co-Director Dr. Guillermo Rodríguez. La presentación obtuvo un puntaje de 82,93 sobre 100, habiendo además obtenido los siguientes distintivos dentro del contexto UNCPBA:

- Primer puesto para Ing. de Sistemas
- Tercer puesto dentro del área de Ingenierías y Tecnologías.
- Tercer puesto dentro de la Facultad de Cs. Exactas.
- Decimotercer puesto a nivel universidad, dentro de 67 postulantes.

En su siguiente etapa, el proyecto fue presentado y aceptado a modo de tesis para la carrera de Ing. de Sistemas en la UNCPBA encontrándose bajo desarrollo por los colaboradores mencionados previamente. Una vez presentado el plan de tesis, se prosigue con el envío de este plan al Workshop WICC 2023. Luego de la finalización del desarrollo se ha planificado una publicación final para dar a conocer los resultados del mismo, eligiendo para ello en el futuro un congreso o revista que se encuentre en la temática. Finalmente, se presentará y defenderá la tesis derivada de este proyecto ante un jurado integrado por docentes de la UNCPBA.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Uno de los focos principales de la primera etapa de desarrollo del proyecto reside en investigar y procesar un gran volumen de datos académicos a fin de evaluar las tendencias de actitud de los alumnos ante la presencia de un asistente interactivo (exponiendo el grado de inserción tentativo de esta nueva tecnología). Continuando con la segunda etapa, se cuenta con una relación estrecha con diversos equipos de desarrolladores académicos de la plataforma Moodle que brindarán todo tipo de consultorías respecto al inicio y continuidad de esta propuesta. Se buscarán alternativas de implementación para el asistente virtual y los medios necesarios para conectar el motor que hay detrás del *bot* con la interfaz de la plataforma, previendo así el logro del prototipado funcional de VMB. Por último, se deberá desarrollar una interfaz bidimensional que pueda servir como vista intermedia entre el alumno y el asistente. La interacción con los *bots* va a generar un nuevo conocimiento, el cual puede ser analizado mediante el uso de varios algoritmos, que permitirán obtener nuevas conclusiones.

Composición del sistema

La primera parte del sistema se compone por la vista Moodle, donde los alumnos y los docentes interactúan con la plataforma (Figura 2). De aquí se generarán las intenciones que deberán ser transmitidas a VMB.

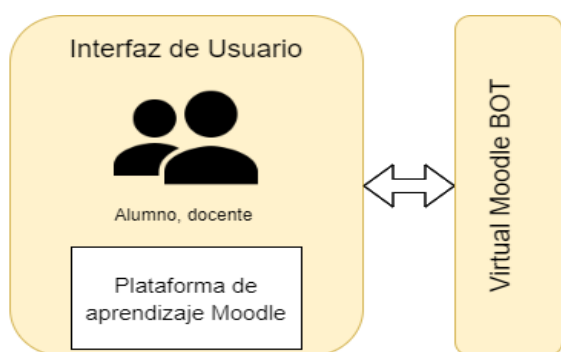


Figura 2: VMB - Interfaz de usuario

Esta interfaz se comunica activamente con el asistente virtual (Figura 3), realizando una interacción bidireccional sobre la información que administra la plataforma. Se cuenta con

una base de datos que gestionará todo tipo de interacciones posibles con el *bot*, además de las distintas interfaces de programación de aplicaciones (APIs) típicas en este tipo de servicios (reconocimiento del lenguaje natural, correctores ortográficos, etc.). Además, se cuenta con una capa de *Machine Learning* para alimentar la inteligencia artificial y las ramificaciones de diálogos disponibles, con tal de que el asistente tenga un mayor número de conversaciones contempladas y disponibles para interactuar con los usuarios.

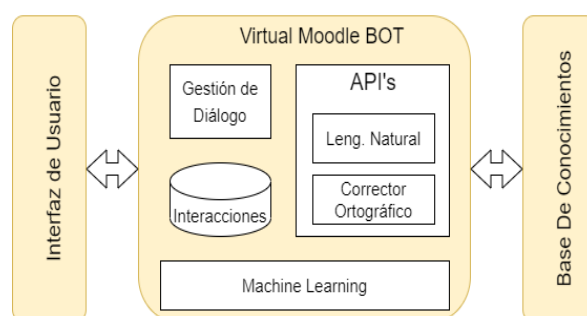


Figura 3: Asistente Virtual

En la Figura 4 se puede observar la base de conocimientos, última etapa del sistema que se encarga de toma de *input* los datos de interés para que puedan ser procesados en el *core* de éste mismo. En esta sección se encuentran todos los algoritmos de analítica de datos y sistemas externos necesarios para analizar y estudiar el conocimiento generado. Además, es de suma importancia que la base de conocimientos se encuentre complementada por la correcta Gestión de Relación con los Clientes (CRM), los cuales son sistemas pensados para controlar la relación con los clientes y mejorar los procesos (en este caso académicos) involucrados.

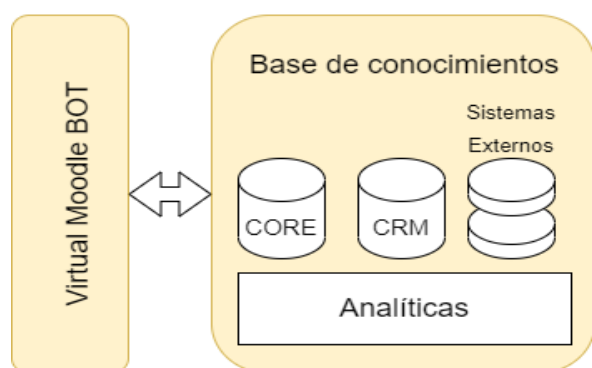


Figura 4: Base de conocimientos

Finalmente, en la Figura 5 se muestra la estructura general del sistema junto con las 3 etapas que lo conforman.

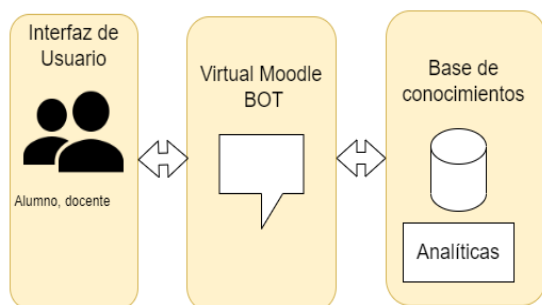


Figura 5: visión general de VMB

Metodología

Siguiendo bajo la misma línea de trabajo que el proyecto acreditado (Illescas G., et. al 2022), se utilizará la misma metodología de desarrollo *Design Thinking* (Brown T., 2009), siendo un método de trabajo en equipo que propicia la colaboración y la entrega frecuente de resultados a través de varias iteraciones.

3. RESULTADOS

OBTENIDOS/ESPERADOS

La interacción con los *bots* generará un nuevo conocimiento, el cual será almacenado en una base de datos. De allí se podrán extraer conclusiones, utilizando algoritmos que interpreten esos datos, tales como las preguntas más realizadas, actividades más progresadas o solicitadas, como así también la posibilidad de aplicar minería de texto. Por otro lado se busca identificar un conjunto de funcionalidades a desarrollar en cuanto a las facilidades que va a tener el alumno para interactuar con la plataforma, dar énfasis a la optimización de los procesos y lo relacionado a la motivación que puede producir esta interacción. Esto último es importante para que el alumno se sienta incentivado a trabajar con la herramienta y que de esa manera le permita una instancia de aprobación diferente a como actualmente lo está haciendo con el entorno base.

Estos dos últimos aspectos tienen una relación estrecha con la búsqueda de la permanencia del alumno dentro de la carrera (Gutiérrez Pizarro 2014; Miranda, et. al. 2017; Illescas, et. al. 2022) tanto para brindarle mayor flexibilidad a su formato de estudio como así también para

ayudar en su experiencia de cursada mediante las nuevas herramientas.

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo se muestra en la siguiente tabla:

Apellido y nombre	Título	Cargo	Funciones
Illescas, Gustavo	Dr.	Prof. UNCPBA	Director
Rodríguez, Guillermo	Dr.	Prof. UNCPBA	Co-Dir.
Herrera De Rosa Exequiel	Pre-grado	Alumno	Becario
Aciti, Claudio	Mg.	Prof. UNCPBA	Integrante
Silvestrini, Pia	Med. Vet.	Prof. UNCPBA	Integrante
Todorovich, Elías	Dr.	Prof. UNCPBA	Integrante

Becario y tesis de grado

- *Virtual Moodle Bot*: Un Bot virtual en Moodle para asistir y complementar actividades académicas en espacios de educación superior. (presentada 2022). Dirección: Illescas G., Rodríguez G. Becario: Herrera De Rosa Exequiel (Ev-CIN 2022)

5. BIBLIOGRAFÍA

- Andrew Bosworth & Nick Clegg (2021). Building the Metaverse Responsibly. VP Facebook Reality Labs, VP Global Affairs.
- Bocanet, Vlad & Fleseriu, Cristina (2020). Modern Technologies Used in Education. 10.4018/978-1-7998-1591-4.
- Carlos, Henriquez & German, Sánchez-Torres & Salcedo, Dixon (2021). Tashi-Bot: A Intelligent Personal Assistant for Users in an Educational Institution. 10.20944/preprints202108.0380.v1.
- Chinedu, Okonkwo & Ade-Ibijola, Abejide (2021). Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence. 2. 100033. 10.1016/j.caeai.2021.100033.
- Dimitriadis, George. (2020). Evolution in Education: Chatbots. Homo Virtualis. 3. 47. 10.12681/homvir.23456.

- Gutiérrez Pizarro C. (2014). Estudio de la permanencia y deserción de los estudiantes de pregrado en una Facultad de Ingeniería, utilizando una metodología mixta de Investigación y árboles de decisión en su predicción. Tesis doctoral Universidad de Alcalá.
- Ibáñez, L. J., & Rodríguez, G. (2020). Agente Conversacional para Servicio al Cliente en Redes Sociales. In 2020 IEEE Congreso Bienal de Argentina (ARGENCON) (pp. 1-6). IEEE.
- Illescas Gustavo, Todorovich Elías, Aciti Claudio, Rodriguez Guillermo, Silvestrini Pia (2022). Aplicación de Analítica de Datos en espacios de Educación Superior. XXIV Workshop de Investigadores en Ciencias de la Computación Universidad Champagnat. Mendoza, Argentina.
- Miranda M y Guzmán J. (2017). Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos-Form. Univ. vol.10 no.3 La Serena. Chile.
- Mogaji, Emmanuel, Balakrishnan, Janarthanan & Nwoba, Arinze & Nguyen, Nguyen. (2021). Emerging-Market Consumers' Interactions with Banking Chatbots. Telematics and Informatics. 65. 101711. 10.1016/j.tele.2021.101711.
- Reyes Ochoa, Carlos & Garza, David & Garrido, Leonardo & Cueva, Víctor & Ramirez Uresti, Jorge. (2020). Virtual Assistants in Education: Model, Implementation and Results.
- Rodriguez, G., Teyseyre, A., Gonzalez, P., & Misra, S. (2020, November). Gamifying users' learning experience of Scrum. In International Conference on Information and Communication Technology and Applications (pp. 497-509). Springer, Cham.
- Roy, Arindam & Singh, Dharmpal, Sahana, Sudipta. (2021). Educational Assistance Bot. Journal of Physics: Conference Series. 1797. 012062. 10.1088/1742-6596/1797/1/012062.
- Vladova, Gergana & Haase, Jennifer & Rüdian, Leo & Pinkwart, Niels. (2019). Educational Chatbot with Learning Avatar for Personalization.

Interacción Persona Ordenador y Educación. Desarrollos y experiencias.

Sanz Cecilia^{1,2} , Artola Verónica¹ , Salazar Mesía Natalí¹ , Iglesias Luciano¹ , Archuby Federico¹ ,
Buffarini Abril¹ , Ibañez Bárbara¹, Bruno Laureano¹, Zeballos Matías¹, Pirono Franco¹, Del Gener
Aldana¹, Astudillo Gustavo³ , Baldassarri Sandra⁴ 

¹Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC.
Facultad de Informática – Universidad Nacional de La Plata

² Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

³ Facultad de Ciencias Exactas y Naturales - Universidad Nacional de La Pampa

⁴GIGA AffectiveLab, Universidad of Zaragoza, España

{csanz, vartola, nsalazar, farchuby}@lidi.info.unlp.edu.ar,
li@info.unlp.edu.ar, abrilbuffarini@gmail.com, astudillo@exactas.unlpam.edu.ar, sandra@unizar.es

RESUMEN

Este artículo presenta una síntesis de las acciones de investigación, desarrollo e innovación abordadas en el año 2022 vinculadas a la línea de interacción persona ordenador y su aplicación al escenario educativo en el marco de un sub-proyecto del Instituto de Investigación en Informática LIDI. En esta línea se vienen investigando diferentes paradigmas de interacción, y su integración en juegos y aplicaciones, que habilitan experiencias educativas innovadoras, donde los estudiantes exploran, analizan y son parte central de la actividad. En este sentido, se ha avanzado en el diseño de experiencias que aprovechan el uso de objetos físicos del entorno en combinación con aplicaciones desarrolladas. Se utilizan modelos de interacción basados en realidad aumentada y virtual, interacción tangible, e hibridaciones de estos paradigmas, así como también se utilizan técnicas relacionadas con el área de computación afectiva. Al mismo tiempo, se ha puesto énfasis en la creación de metodologías específicas para el diseño y desarrollo de juegos serios educativos. Se describen en este trabajo los principales resultados alcanzados en 2022 en estas temáticas.

Palabras clave: interacción persona-ordenador, interacción tangible, realidad aumentada, realidad virtual, entornos inmersivos, computación afectiva, escenarios educativos, juegos serios.

CONTEXTO

Las acciones que aquí se describen forman parte del sub-proyecto titulado “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos”, que pertenece al proyecto general titulado: “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” (período 2018-2022), y su continuidad en el proyecto “Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital”. El primero es un proyecto del Instituto de

Investigación en Informática LIDI, de la Facultad de Informática de la Universidad Nacional de La Plata, acreditado por el Ministerio de Educación de la Nación, mientras que el segundo, se ha presentado como nuevo proyecto que da continuidad a gran parte de las líneas de investigación y desarrollo, agregando algunas y avanzando sobre nuevos desafíos.

1. INTRODUCCIÓN

Las líneas de investigación, desarrollo e innovación que aquí se presentan abarcan el área de interacción persona-ordenador, la ingeniería de software, el desarrollo de juegos, y específicamente, se posicionan en temas de tecnología informática aplicada en educación.

Desde los inicios de la computación, la búsqueda por mejorar las formas de interacción entre las personas y las computadoras ha sido un tema relevante de investigación. Actualmente, se han abierto nuevas posibilidades en este campo, que ofrecen oportunidades para generar experiencias educativas innovadoras, a partir del uso de sensores, actuadores, de dispositivos móviles y del auge de la computación ubicua. El estudio de modelos de interacción basados en interacción tangible, realidad aumentada, realidad virtual y sus combinaciones junto con la generación de entornos inmersivos forman parte de la agenda de investigación actual. Al mismo tiempo, la aplicación de técnicas de computación afectiva se viene abordando en forma creciente y aporta también al escenario educativo, ayudando en la personalización de itinerarios formativos, en la mejora de propuestas de enseñanza y aprendizaje, y en el fortalecimiento de los entornos personales de aprendizaje (Astudillo, Sanz y Baldassarri, 2022; 2021a y b; 2020).

En las líneas siguientes se describen algunos conceptos teóricos que subyacen a esta línea de I+D+I.

Las interfaces de usuario tangibles (TUI: *Tangible User Interface*) proponen una forma innovadora de interacción en la que se reducen las restricciones impuestas por las interfaces GUI (*Graphical User Interfaces*). Las TUI permiten interactuar con el

contenido digital a través de la manipulación de objetos físicos de la vida real, en lugar de usar dispositivos de entrada convencionales como mouse y teclado (Rodić y Granić, 2021). Estos objetos físicos actúan como una representación de la información digital, y sirven como mecanismo de entrada y/o salida, ofreciendo feedback a los estudiantes, a partir de diferentes canales (háptico, visual, auditivo, etc) (Chettaoui et al., 2022; Alvarado, Sanz y Baldassarri, 2021). En varios trabajos se han estudiado los beneficios de la TUI y el uso de materiales físicos en combinación con información digital para guiar la resolución de problemas y abordar conceptos abstractos en áreas como la música, la programación, bioquímica, matemática, entre otras (Artola, Sanz y Pesado, 2020; Artola, Sanz y Baldassarri, 2022). También este tipo de interfaces soportan el aprendizaje combinando la interacción corporizada con el uso de objetos físicos, lo que puede promover el pensamiento y el desarrollo de habilidades motrices en niños y adultos (Horneker, 2011).

En el gradiente de paradigmas de interacción que van de aquellos centrados mayormente en el entorno físico a los centrados en el entorno digital, también han cobrado fuerza las experiencias educativas con realidad aumentada. En este sentido, trabajos como el presentado por Carrizo, Barutti, y Soto (2022) e Ibañez y Kloss (2018), permiten conocer beneficios del uso de la realidad aumentada en procesos educativos. En estos trabajos se resalta la motivación, y el cambio de rol del estudiante en este tipo de actividades, y su utilidad en el aprendizaje de diferentes temáticas. La realidad aumentada permite combinar información del entorno físico con información digital, posibilitando una experiencia combinada y centrada en el contexto. En general, estas actividades habilitan las capacidades exploratorias de los estudiantes para descubrir relaciones e información del contexto (Salazar Mesía y Sanz, 2022; Salazar Mesía, Sanz y Gorga, 2019 a y b; Barroso Osuna y Palacios Rodríguez, 2022; Laurens-Arredondo, 2022).

En cuanto a la realidad virtual, es una tecnología adecuada para el desarrollo de algunas experiencias educativas, debido a su facilidad para captar la atención de los estudiantes mediante su inmersión en mundos virtuales relacionados con las diferentes ramas del saber, lo que puede ayudar en el aprendizaje de los contenidos de diversas materias. Cuando se diseña un sistema con este paradigma de interacción, se requiere atender a determinadas variables importantes para alcanzar los objetivos esperados. Aspectos como: la inmersión, el sentido de presencia, la calidad de las imágenes de las escenas, la interactividad, los sentidos involucrados, y las formas de interacción planteadas influyen en la experiencia del usuario y en el logro de los objetivos que se proponen (Menjivar et al., 2021; Tsaramirsis et al., 2016; Chirinos, Sanz y Dapoto, 2020; Mazza, Sanz y Artola, 2021).

Como parte de las líneas de investigación que se abordan se encuentra el diseño de juegos serios (definidos como aquellos que tienen un objetivo más

allá del entretenimiento), y en su implementación se integran los paradigmas de interacción mencionados. Se han encontrado evidencias de que el uso de los juegos serios puede tener impacto en el aprendizaje, en el disfrute y en generar menor ansiedad durante la experiencia educativa (Kiili y Ketamo, 2018; Archuby, Sanz y Manresa-Ye, 2023).

Estos estudios inspiran la investigación que se desarrolla en el marco de los proyectos mencionados en este trabajo y dan fundamento a las aplicaciones que se implementan, muchas de ellas diseñadas como juegos serios. Así, se analizan para diferentes situaciones educativas, qué tipos de paradigmas de interacción podrían aportar y beneficiar una determinada actividad con estudiantes. En este sentido, el estudio de las posibilidades y de antecedentes vinculados a experiencias educativas con tecnologías digitales iluminan las decisiones para el diseño de estas experiencias, entramándose con las características del contexto y los destinatarios. Preguntas tales como: ¿cuáles son los objetivos de aprendizaje?, ¿cómo se abordan las estrategias para lograr el entretenimiento y el alcance de los objetivos planteados?, ¿cuáles son los resultados esperados?, ¿qué rol tendrán los estudiantes?, ¿qué objetos físicos podrían aportar al aprendizaje y a la experiencia educativa a diseñar, y servirían como dispositivos de entrada/salida de la aplicación a crear?, son algunas de las que orientan los desarrollos que surgen como resultados de esta investigación.

Las líneas de investigación aquí presentadas guardan relación directa con las temáticas de la Maestría y Especialización en Tecnología Informática Aplicada en Educación de la misma Facultad, por lo que se desarrollan tesis y trabajos finales en estos temas. También se cuenta con una agenda de cooperación a nivel nacional e internacional, que permite profundizar el trabajo que se lleva adelante en el III LIDI.

2. LÍNEAS DE INVESTIGACIÓN / DESARROLLO

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Juegos serios educativos basados en Realidad Aumentada, Interacción Tangible y Realidad Virtual. Juegos pervasivos. Metodologías ágiles y herramientas para su diseño, creación y evaluación
- Interacción Tangible. Marcos que fundamentan este paradigma, relación entre objetos físicos y digitales, aplicaciones en educación
- Diseño de Objetos Activos para IT. Tipos de feedback en objetos activos. Uso de actuadores y sensores
- Realidad Aumentada para juegos serios educativos
- Entornos inmersivos e interactivos basados en realidad virtual y aumentada. Embodied Interaction.
- Técnicas de mapeo de proyecciones y modelado 3D para crear entornos aumentados.

- Sistemas educativos adaptativos y sistemas recomendadores para recursos educativos
- Computación afectiva en entornos digitales para el escenario educativo

Estas líneas se sostienen a partir del trabajo conjunto de docentes investigadores, becarios, tesistas, pasantes y estudiantes que participan.

3. RESULTADOS OBTENIDOS

Los resultados que aquí se presentan son los correspondientes al año 2022 e inicios de 2023, pero se dan como un *continuum* del trabajo sostenido a lo largo de los últimos años.

En 2023 se ha avanzado en el diseño y desarrollo de diferentes juegos serios educativos. Para el proceso de diseño se viene aplicando la metodología DIJS, que surge de una tesis de maestría de uno de los miembros del proyecto. La metodología DIJS permite guiar tanto la concepción del juego como sus etapas posteriores hasta alcanzar la implementación y la evaluación (Archuby, Sanz y Manresa-Ye, 2020; Archuby, Sanz y Manresa-Ye, 2023).

Durante el 2022, se ha consolidado el desarrollo de EmpoderAR. Este juego serio permite, a través de una línea de tiempo sobre figuras destacadas de la Informática, “aumentar” la historia de algunas de las mujeres que han sido innovadoras en esta disciplina. El juego presenta trivias que posibilitan ganar trofeos, que luego aparecerán en la cueva de la sabiduría, que es resultado de la dinámica lúdica propuesta. El juego fue mejorado atendiendo a cuestiones de usabilidad, que surgieron a partir de su integración en el recorrido de actividades del Centro de Innovación y Transferencia Tecnológica de la Facultad de Informática (CIyTT). EmpoderAR se utiliza en diferentes talleres que se brindan en este espacio, y se ha trabajado con estudiantes de la escuela media y público en general; lo que permitió obtener feedback y reconocer sus posibilidades y aspectos de mejora. El juego utiliza la realidad aumentada, como resultado de la investigación en esta línea (Lizarralde et al., 2019). Se contó con una becaria UNLP para su implementación (Ibañez y Sanz, 2021). En 2022 también se aprobó una tesis de maestría que analizó la vinculación de experiencias educativas de realidad aumentada y el rendimiento académico. En ese marco se llevó adelante un estudio de caso con un material educativo con RA para la enseñanza de temas relacionados con la programación (Romano, Sanz y Gorga, 2022). También, se profundizó en el diseño de una plantilla de RA para la creación de actividades educativas geolocalizadas por parte de docentes, esto forma parte de una tesis de maestría en realización (Salazar Mesia y Sanz, 2022; Salazar Mesia, Sanz y Gorga 2019 a y b).

En el marco de un proyecto de innovación con estudiantes, se llevó adelante un sistema basado en interacción tangible, vinculado a narrativas digitales. Se trata de un cuadro interactivo que presenta piezas físicas a partir de las cuales se produce la interacción (Fig. 1).

El cuadro presenta a Grace Hopper, una reconocida figura de la historia de la informática. Se integran para la interacción diferentes tipos de sensores, que permiten generar animaciones en el cuadro. Por ejemplo, un interruptor de luz tipo cuerda, al ser jalado, enciende una luz en la imagen digital; mientras que una canilla física, al abrirla provoca que se anime una caída de agua con palabras de lenguajes de programación que se transforman en 0 y 1, para generar una metáfora sobre el concepto de compilador. De esta manera se despierta la curiosidad de quiénes interactúan, y a partir de ese propio interés se dialoga en los talleres con los participantes. Este trabajo se vincula especialmente con el paradigma de interacción tangible (Garay et al., 2022).

Al mismo tiempo, se ha continuado con el desarrollo de Escapados, creado en forma interdisciplinaria utilizando la metodología DIJS. Este juego se orienta a reforzar aprendizajes del área de Química y Matemática, a través de un entorno 3D, y el uso de interacción tangible con objetos del entorno físico que deben ser reconocidos ante la cámara web del jugador para integrarse al mundo del juego. Durante 2022 Escapados se presentó como demo en el marco del congreso TEyET, y fue utilizado en el contexto de actividades del CIyTT (Zeballos et al., 2021, 2022). En 2023 se avanzará en su aplicación en un estudio de caso en el marco de una tesis de maestría, y se realizarán algunas funcionalidades adicionales vinculadas a registrar en un log, las interacciones que realizan los jugadores. También se continuará la modificación del juego Albores para su uso en el CIyTT (Bigurrarena et al., 2021).



Fig. 1 – Cuadro interactivo, basado en interacción tangible

Durante el año, también se avanzó en el estudio de técnicas de mapeo de proyecciones para generar

experiencias educativas inmersivas e interactivas¹. Se llevaron adelante los siguientes proyectos (CIyTT, 2022):

- Flores y movimiento: permite experimentar el uso del cuerpo como medio para la interacción. Se trabaja con proyecciones que son modificadas a partir de la detección del movimiento de las manos de los participantes. Este proyecto se está extendiendo también con el uso de sensores para capturar el ritmo cardíaco. Se realizará un experimento que lo vinculará con el área de salud y bienestar.
- Tiempo y Tecnología (TicTec): el proyecto se basa en técnicas de mapeo de proyecciones y relata la evolución en el tiempo de diferentes tecnologías que han aportado a la disciplina Informática. TicTec se presenta como un ejemplo de narrativas digitales.
- Mapping y Videojuegos: se trata de una experiencia interactiva para revivir reconocidas creaciones de personajes de videojuegos, a partir de la interacción con sensores y mapeo de proyecciones.

Todos estos proyectos se desarrollaron en vinculación con las líneas aquí presentadas y están disponibles para quienes participan de las actividades del CIyTT. Al mismo tiempo, se utilizaron en este espacio los juegos previamente desarrollados como Murales (Sanz et al., 2020).

Respecto de los avances en la investigación de realidad virtual, durante 2022, se estudiaron nuevos trabajos sobre el uso de RV en educación, y aspectos relacionados con el diseño de aplicaciones. Se desarrollaron dos ejes nuevos de HuVi Parque Nacional Iguazú (Chirinos, Sanz y Dapoto., 2020), y se abordó la implementación interdisciplinaria de HuVi Tango y Casa Curutchet. Este desarrollo se enmarca en un proyecto de extensión que se lleva adelante junto a la Facultad de Ciencias Económicas de la UNLP, en el que se propone acercar el patrimonio argentino a través de la realidad virtual, utilizando dinámicas lúdico-educativas. Ya se cuenta con 3 juegos de HuVi (<https://www.huellaspatrimoniales.econo.unlp.edu.ar/>)

Estos fueron utilizados en el marco de diferentes talleres con niños y niñas de diferentes comedores de la región, y con escuelas. En particular, se avanzó en el uso de HuVi Ischigualasto y Talampaya que fue desarrollado en 2021 (Mazza, Rucci & Sanz, 2021).

Otro trabajo realizado en el marco de estas líneas de investigación, a través de una tesina de grado (aprobada en 2022), fue el diseño de un mando para poder sensar movimientos del jugador (correr, caminar, saltar). El mando desarrollado, a manera de prototipo, cuenta con una tobillera y un joystick con botones y joystick. Para poder probarlo, se implementó un minijuego llamado “Capitana Aldana”, y se realizaron sesiones con potenciales usuarios, analizando la usabilidad (Del

Gener, Iglesias y Sanz, 2022). En 2023 se avanzará en estas acciones y el proyecto se dejará disponible para trabajar con estudiantes visitantes del CIyTT.

En el marco de los temas de computación afectiva, se dirigen dos tesis doctorales. Una de ellas se vincula con la recomendación de canciones a partir de la metaanotación de éstas con emociones. Esta tesis está en su última etapa de realización y se espera entregarla en 2023 (Ospitia-Medina et al., 2023; Ospitia-Medina et al., 2021).

Al mismo tiempo, se continuó con la investigación sobre metadatos para caracterizar videos educativos con emociones (Astudillo, Sanz & Baldassarri, 2021a), y sobre bases de datos emocionales, a partir de una revisión sistemática de literatura (Astudillo, Sanz & Baldassarri, 2021b). Durante 2022, se avanzó en la tesis doctoral vinculada a estos temas, y el tesista realizó una estancia de investigación en la Universidad de Zaragoza, donde se está llevando adelante un experimento con usuarios para el análisis de emociones, en particular, el aburrimiento como emoción de interés en el contexto educativo.

En cuanto a los proyectos vinculados y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se cuenta con un acuerdo de colaboración con la Universidad de Zaragoza y la Universidad de Islas Baleares para trabajar en algunas de las temáticas aquí expuestas.
- Se ha obtenido financiamiento en el marco de una convocatoria K107 de Erasmus para estancias de viaje entre la Universidad de Zaragoza y la Universidad Nacional de La Plata. Se llevaron a cabo dos estancias, un docente de Zaragoza estuvo en la UNLP y un estudiante de doctorado viajó a Zaragoza.
- Se participó en el proyecto “Pervasive Gaming Experiences For @ll (Pergamex)”, en particular en el subproyecto (RTI2018-096986-B-C31) de la UZ. Se está avanzando en la presentación de un nuevo proyecto.
- Además, la Dra. Sanz es miembro colaborador del grupo de Investigación en Interfaces Avanzadas (AffectiveLab).
- Se participa en la Red constituida por universidades de Iberoamérica en el marco del programa "Pablo Neruda" dentro del Espacio Iberoamericano del Conocimiento (EIC) y de la Organización de los Estados Iberoamericanos (OEI) orientada a la movilidad de estudiantes y docentes de doctorado.
- Se asesora en el proyecto de la Universidad de Costa Rica: “Juegos serios basados en interfaces tangibles con objetos activos” con código: 510-C2-321
- Se participa en la RedAUTI: Red temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva. En este ámbito se analizan materiales educativos para la TVDI.

¹ En las redes del CIyTT pueden verse acciones al respecto: @infociytt

- Se participa del proyecto TEMOR: Ecosistema tecnológico para el reconocimiento del estado de ánimo de los pacientes durante el proceso de rehabilitación cardíaca. Proyecto de la Universidad de Zaragoza, aprobado en 2022. Duración 2 años.

4. FORMACION DE RECURSOS HUMANOS

Se trabaja en la formación de recursos humanos a través de la dirección de tesis de doctorado, maestría, trabajos finales de especialización y tesinas de grado; también con becas de diferentes organismos de ciencia y técnica. En 2022, se ha aprobado 1 tesis de Maestría dirigida en relación a los temas de RA. Se ha desarrollado un proyecto de innovación con alumnos, con 3 alumnos participantes de Ingeniería en Computación. Se ha finalizado una tesina de grado vinculado a estas temáticas. Se están dirigiendo 2 tesis doctorales en relación a temas de computación afectiva, y 2 orientadas a realidad virtual y juegos serios, respectivamente.

5. BIBLIOGRAFÍA

- Alvarado, M. C., Sanz, C., & Baldassarri, S. (2021). Desarrollo de una experiencia educativa con juguetes activos en una mesa de interacción tangible. *Revista de la Asociación Interacción Persona Ordenador (AIPO)*, 2(2), 74-84.
- Archuby, F., Sanz, C., & Manresa-Yee, C. (2020). Metodologías de diseño y desarrollo para la creación de juegos serios digitales. Tesis de maestría finalizada. Sep. 2020. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/111123>
- Archuby, F., Sanz, C., & Manresa-Yee, C. (2023). DIJS: Methodology for the Design and Development of Digital Educational Serious Games in *IEEE Transactions on Games*, doi: 10.1109/TG.2022.3217737. Early access.
- Artola, V., Sanz, C., & Baldassarri, S. (2022). A Novel Tangible Interaction Authoring Tool for Creating Educational Activities: Analysis of its Acceptance by Educators. *IEEE Transactions on Learning Technologies*, 2022, doi: 10.1109/TLT.2022.3216117.
- Artola, V., Sanz, C., & Pesado, P. (2020). Tesis Doctoral. Interacción tangible en escenarios educativos. Diseño de una herramienta de autor para la creación de Actividades educativas basadas en interacción tangible. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/98135>
- Astudillo, G.J., Sanz, C.V. & Baldassarri, S. (2022). Emotional Meta-annotation of Educational Videos. A Review of Emotional Database Characterization. In: Abásolo, M.J., Olmedo Cifuentes, G.F. (eds) *Applications and Usability of Interactive TV. jAUTI 2021. Communications in Computer and Information Science*, vol 1597. Springer, Cham. https://doi.org/10.1007/978-3-031-22210-8_4
- Astudillo, G. J., Sanz, C. V., & Baldassarri, S. (2021a). Revisión sistemática sobre la meta-anotación de videos educativos con emociones. In XVI Congreso de Tecnología en Educación & Educación en Tecnología-TE&ET 2021 (La Plata, 10 y 11 de junio de 2021).
- Astudillo, G., Sanz, C. & Baldassarri, S. (2021b). Análisis del etiquetado emocional de videos educativos. *JAUTI* 2021. <https://link.springer.com/conference/jatui>
- Astudillo, G.; Sanz, C. & Baldassarri, S. (2020). Definición de un modelo de relaciones entre tipos de videos educativos, perfiles de usuarios y emociones. Propuesta de tesis de doctorado aprobada en 2020.
- Barroso-Osuna, J. & Palacios-Rodríguez, A. (2022). Expanding the virtual universe of university students. Educational use of augmented reality and contributions of Rafodiu Project. *Tecnología, Ciencia y Educación*, 23, 137-154. <https://doi.org/10.51302/tce.2022.874>
- Bigurrarena, N., Ballardini, E., Artola, V., Buffarini, A., Nordio, M., & Sanz, C. V. (2021). Albores: un juego basado en interacción tangible para conocer figuras destacadas de la historia de la informática. In XVI Congreso de Tecnología en Educación & Educación en Tecnología-TE&ET 2021 (La Plata, 10 y 11 de junio de 2021).
- Carrizo, M. A., Barutti, M. E., y Soto, S. B. (2022). Incorporación de realidad aumentada como propuesta didáctica para la enseñanza y el aprendizaje de ciencias. *Educación En La Química*, 28(01), 63–73.
- Centro de Innovación y Transferencia Tecnológica, Facultad de Informática, Universidad Nacional de La Plata <https://ciytt.info.unlp.edu.ar/>
- Chettaoui, N., Atia, A., Bouhrel, M.S. et al. Usability of Montessori tangible user interfaces to support learners' retention skills in preschools. *Pers Ubiquit Comput* (2022). <https://doi.org/10.1007/s00779-022-01706-9>
- Chirinos, Y., Sanz, C. & Dapoto, S. (2020) La realidad virtual como mediadora de aprendizajes. Desarrollo de una aplicación móvil de realidad virtual orientada a niños. Tesis de Maestría en Tecnología Informática Aplicada en Educación finalizada en Diciembre de 2020. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/111879>
- Del Gener, A., Iglesias, L., & Sanz, C. (2022). Exergames: propuesta de un gamepad para sensor movimientos del jugador. Tesina de grado de la Licenciatura en Sistemas. Facultad de Informática, UNLP. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/145940>
- Garay, F., Bruno, L., Lambre, J., Artola, V. & Sanz, C. (2022). Informe final del proyecto de innovación con alumnos. III LIDI. Facultad de Informática. UNLP
- Hornecker E (2011) The role of physicality in tangible and embodied interactions. *Interactions* 18(2):19–23. <https://doi.org/10.1145/1925820.1925826>

- Ibáñez, M. B., & Kloos, C. (2018). Augmented reality for stem learning: A systematic review. *Computers & Education* (123), 109-123.
- Ibáñez, B. R. & Sanz, C. (2021). Informe final de beca. Entregado en octubre de 2021.
- Kiili, K., & Ketamo, H. (2018). Evaluating Cognitive and Affective Outcomes of a Digital Game-Based Math Test. *IEEE Trans. Learn. Technol.*, vol. 11, no. 2, pp. 255–263.
- Laurens-Arredondo, L. (2022). Mobile augmented reality adapted to the ARCS model of motivation: a case study during the COVID-19 pandemic. *Educ Inf Technol*. <https://doi.org/10.1007/s10639-022-10933-9>
- Lizarralde, A.; Sanz, C.; Gorga, G. & Buffarini, A.; Beltrán, E. & Kraselsky, R. (2019). Ruta Darwin: un juego con realidad aumentada para conocer las experiencias de Charles Darwin en su travesía a bordo del Beagle. *TE&ET 2019*. ISBN: 978-987-733-196-7. Pp. 262-264.
- Mazza, M.A., Sanz, C., Artola, V. (2021). Virtual Reality Serious Game to Bring the History of Key Figures of Computer Science Closer to Young People. In: Pesado, P., Eterovic, J. (eds) *Computer Science – CACIC 2020*. CACIC 2020. Communications in Computer and Information Science, vol 1409. Springer, Cham. https://doi.org/10.1007/978-3-030-75836-3_4
- Menjivar Valencia, E., Sánchez Rivas, E., Ruiz Palmero, J., & Linde Valenzuela, T. (2021). Revisión de la producción científica sobre la Realidad Virtual entre 2016 y 2020 a través de Scopus y WOS. *EDMETIC*, 10(2), 26-55. <https://doi.org/10.21071/edmetic.v10i2.13422>
- Ospitia-Medina, Y., Baldassarri, S., Sanz, C., & Beltrán, J. (2021). Music Recommender Systems: A Review Centered on Biases. Aprobado para publicar en “Advances in Speech and Music Technology: Computational Aspects and Applications” of the Springer book series “Signal and Communication Technology”. Esperando su publicación.
- Ospitia-Medina, Y., Baldassarri, S., Sanz, C., Beltrán, J.R. (2023). Music Recommender Systems: A Review Centered on Biases. In: Biswas, A., Wennekes, E., Wiczorkowska, A., Laskar, R.H. (eds) *Advances in Speech and Music Technology. Signals and Communication Technology*. Springer, Cham. https://doi.org/10.1007/978-3-031-18444-4_4
- Rodić, L. D., & Granić, A. (2021). Tangible interfaces in early years’ education: a systematic review. *Personal and Ubiquitous Computing*. doi:10.1007/s00779-021-01556-x
- Romano, L., Sanz, C. V., & Gorga, G. M. (2022). Realidad aumentada en contextos educativos y su relación con el rendimiento académico universitario. Tesis de Maestría en Tecnología Informática Aplicada en Educación. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/145578>
- Salazar, N.; Sanz, C. & Gorga, G. (2019 a). Análisis comparativo de librerías de realidad aumentada. Sus posibilidades para la creación de actividades educativas. *Especialización en TIAE*.
- Salazar Mesía, N.; Sanz, C. & Gorga, G. (2019 b). Diseño de plantillas para la creación de actividades educativas con Realidad Aumentada en AuthorAR. XXV Congreso Argentino de Ciencias de la Computación. Octubre, Córdoba. ISBN: 978-987-688-377-1.
- Salazar Mesía, N. & Sanz, C. (2022). Diseño de una herramienta para la creación de actividades educativas basadas en Realidad Aumentada. XXVIII Congreso Argentino de Ciencias de la Computación. Octubre, La Rioja. ISBN: 978-987-1364-31-2.
- Sanz, C.; Artola, V.; Nordio, M.; Pirondo, F.; Ibáñez, B.; Corro, B. (2020). Murales: creando puentes entre el mundo digital y el físico, XV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2020)
- Tsaramirsis, G., Buhari, S., AL-Shammari, K. O., Ghazi, S., Nazmudeen, M. S., & Tsaramirsis, K. Towards simulation of the classroom learning experience: Virtual Reality approach. (2016) In H. Mn (Ed.), *The 10th Indiacom - 2016 3rd International Conference on Computing for Sustainable Global Development* (pp. 1343–1346). New Delhi, India: IEEE.
- Zeballos, M., Lombardo, S., Fanelli, V., Artola, V., Archuby, F., Ferreyra, P., Gubaro, M., & Sanz, C. (2021). Escapados. Presentación de la Jornada de CyT 2021, FI, UNLP. Disponible en: <https://youtu.be/ffqFTiPult4>
- Zeballos, M., Lombardo, S., Fanelli, V., Gubaro, M., Ferreyra, P., Artola, V., & Sanz, C. V. (2022). Juego educativo basado en realidad virtual e interacción tangible para el aprendizaje de temas de matemática y química. XVII Congreso de Tecnología en Educación & Educación en Tecnología (TE&ET 2022, Entre Ríos).

Realidad aumentada en contextos educativos y su relación con el rendimiento académico universitario.

Tesis de Maestría en Tecnología Informática Aplicada en Educación, Facultad de Informática, Universidad Nacional de La Plata

Tesista: Romano, Lucas

lromano@untdf.edu.ar

Directora: Cecilia Sanz (III LIDI, Facultad de Informática, UNLP). **Codirectora:** Gladys Gorga (III LIDI, Facultad de Informática, UNLP)

Fecha de exposición: julio de 2022

RESUMEN

La Realidad Aumentada (RA), se refiere a la inclusión, en tiempo real, de elementos virtuales dentro de un entorno físico. La combinación del mundo real y virtual a través de RA tiene el potencial de permitir a los individuos descubrir teorías, fenómenos, procesos y comportamientos abstractos, así como características que generalmente no están disponibles en entornos del mundo real. El concepto de Realidad Aumentada, por lo tanto, refiere a un incremento de la información que un ser humano puede obtener por sí mismo al interactuar con el mundo físico.

Las tecnologías de Realidad Aumentada en escenarios educativos posibilitan a los estudiantes el desarrollo de habilidades tales como las espaciales y mejoras en las prácticas con temas abstractos, que en el mundo real son difíciles de implementar. También se han identificado mejoras en varios aspectos, al utilizar estrategias pedagógicas con materiales con Realidad Aumentada, tales como: motivación, interés, atención, satisfacción, interacción, sensación de control, entre otras.

Con esta motivación este trabajo se enfocó en realizar una investigación acerca de las posibilidades de la Realidad Aumentada en diferentes contextos educativos y su vinculación con el rendimiento académico, a través de variables de estudio asociadas a éste.

Para ello, se llevó adelante una revisión acerca del concepto de Realidad Aumentada, y de las posibilidades que ofrece cuando se la aplica a diferentes niveles educativos con una propuesta didáctica que la integre.

Luego, se abordó una investigación sobre el concepto de rendimiento académico y sus variables asociadas que provocan impacto en el desempeño de los estudiantes. Se realizó un análisis de investigaciones que, en mayor medida, concluyen que el rendimiento académico de los estudiantes presenta mejoras al utilizar tecnologías de Realidad Aumentada con una propuesta educativa determinada.

Como parte de la tesis, se desarrolló un material educativo con el uso de Realidad Aumentada y se diseñó un estudio de caso en el que se utilizó dicho material en el nivel universitario. En particular se aplicó a temas de enseñanza de programación. Para el estudio de caso se diseñaron instrumentos de recolección de datos del tipo cuestionario y entrevistas con informantes clave del proceso.

Finalmente, tanto a través de la investigación teórica como de los resultados del estudio de caso realizado, este trabajo permitió observar mejoras en el rendimiento académico de los estudiantes, en especial en las variables motivación, interés, atención, satisfacción y calificación, al aplicar tecnologías de Realidad Aumentada durante los procesos de enseñanza y aprendizaje.

OBJETIVOS

El objetivo general de este trabajo fue investigar acerca de la aplicación de Realidad Aumentada (RA), como tecnología emergente, en diferentes contextos educativos y su vinculación con variables que influyen en el rendimiento académico. Particularmente, se analizaron estudios acerca de la relación que existe entre RA y el rendimiento académico utilizando materiales educativos con tecnologías de Realidad Aumentada en el aula universitaria.

Como objetivos específicos se plantean:

- Investigar acerca de la utilización de RA en los diferentes contextos educativos.
- Indagar acerca de los aspectos relacionados con el rendimiento académico.

- Investigar acerca de la vinculación que existe entre la RA y diferentes variables que influyen en el rendimiento académico.
- Diseñar una experiencia que utilice tecnologías con RA para su aplicación en el ámbito de la enseñanza universitaria.
- Realizar un estudio de caso para visibilizar alcance y potencialidades.

APORTES Y RESULTADOS

Este trabajo aporta una revisión sistemática de literatura sobre el concepto de rendimiento académico, y sobre experiencias educativas que utilizan Realidad Aumentada y han encontrado alguna relación con variables asociadas al rendimiento académico. También se investigaron herramientas de autor para la construcción de materiales educativos con Realidad Aumentada.

A la luz del estudio de estado del arte realizado se diseñó y desarrolló un material educativo con Realidad Aumentada que permite incorporar en una guía de ejercicios prácticos un personaje virtual (un robot), que ofrece feedback afectivo y de procesamiento de la tarea para acompañar el desarrollo de una práctica vinculada a la enseñanza de la programación. El material educativo con RA fue integrado al contexto de una asignatura de la carrera de la Licenciatura en Sistemas, de la Universidad Nacional de Tierra del Fuego. Se trabajó con una metodología de grupos contrastados, uno experimental con 25 estudiantes (utilizó el material con RA diseñado) y otro grupo de control (con 40 estudiantes) que trabajó con materiales impresos y las presentaciones tradicionales de la materia.

Finalizada la experiencia, a partir de la implementación de instrumentos de recolección de datos, se analizaron los resultados obtenidos sobre los grupos contrastados analizando variables asociadas al rendimiento académico.

Los resultados del estudio de caso desarrollado permiten afirmar que los alumnos recibieron y trabajaron con el material con RA motivados, manifestando su curiosidad, y luego interés al trabajar sobre la guía de trabajo práctico aumentada. Calificaron el material con RA como atractivo, al compararlo con otros materiales que venían trabajando durante las unidades anteriores. Muchos de ellos expresaron sentir satisfacción de logro, al haber completado los ejercicios prácticos con el material con RA. Los resultados de los exámenes parciales arrojaron una calificación mayor en el grupo de la experiencia respecto del grupo de control, si bien esto es alentador respecto del uso de tecnologías de RA, no se considera un resultado concluyente debido a la muestra con la que se ha trabajado.

Los docentes involucrados valoraron la experiencia en su totalidad como enriquecedora y positiva, y consideraron que la experiencia tiene un alto potencial, no sólo durante la clase, sino fuera de ella.

La propuesta didáctica presentada, mediada a través de tecnologías con RA, se considera de gran utilidad tanto para los alumnos como para los docentes. El tener un “asistente virtual de cátedra” al cual un alumno puede acudir, de manera sencilla, a través de su dispositivo móvil se presenta como un beneficio innovador para los procesos de enseñanza y aprendizaje.

La ayuda del asistente virtual (un robot 3D animado en este caso), en el material creado con tecnologías con RA, para la resolución de ejercicios fue considerada por los docentes como una herramienta valiosa y eficaz.

LÍNEAS DE TRABAJO FUTURO

Como línea de trabajo a futuro, se considera de interés continuar con la investigación de herramientas de autor emergentes que permitan optimizar la creación de materiales aumentados con fines educativos. De esta manera, los docentes o expertos en contenidos contarán con más opciones de incorporación de contenido virtual, y funcionalidades para las diferentes plataformas de los dispositivos que los alumnos utilicen.

Otra futura línea de investigación de interés consiste en profundizar en las variables asociadas al rendimiento académico y su vinculación con las tecnologías de RA. Esta investigación obtiene resultados similares a los encontrados en la revisión sistemática de la literatura, sin embargo, la muestra no permite ser concluyente. Se aporta a sostener los resultados previos encontrados en otros estudios. Al mismo tiempo, se abre la puerta para investigar en variables de entrada que también pueden incidir en los resultados tales como género, edad y conocimientos previos.

Por último, se espera avanzar en el diseño y desarrollo de materiales educativos con tecnologías con RA en otras cátedras en el ámbito universitario, con el uso de un “asistente virtual”, que acompañe a los alumnos a realizar ejercicios, a fines de obtener resultados con una muestra más amplia y diversa de alumnos de nivel universitario.

Tecnologías digitales aplicadas al escenario educativo. Desarrollos y experiencias

Sanz Cecilia^{1,3} , Artola Verónica¹ , Gonzalez Alejandro¹ , Zangara Alejandra¹ , Iglesias Luciano¹ ,
Ibáñez Eduardo¹ , Violini Lucía^{1,2} , Fachal Adriana¹ , Archuby Federico^{1,2} , Abásolo María José¹ ,
Manresa-Yee Cristina⁴ , Paula Dierse⁵ , Teresa Coma-Roselló⁶ , Pesado Patricia¹ 

¹Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC.

Facultad de Informática – Universidad Nacional de La Plata

² Han contado con beca UNLP

³Investigador Asociado de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

⁴Departamento de Ciencias Matemáticas e Informática. Universidad de las Islas Baleares, España

⁵Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa.

⁶ Departamento de Educación, Universidad de Zaragoza, España

{csanz, bartola, agonzalez, li, eibanez, lviolini, farchuby, mjabasolo, ppesado}@lidi.info.unlp.edu.ar,
alejandra.zangara@gmail.com, afachal@hotmail.com, paula.dierse@gmail.com, cristina.manresa@uib.es,
tcoma@unizar.es

RESUMEN

Este trabajo presenta los resultados correspondientes a un proyecto de investigación, desarrollo e innovación vinculado al área de tecnología informática aplicada a escenarios educativos. En el marco de este proyecto se investiga en un conjunto de líneas que abarcan sistemas colaborativos, juegos serios educativos, entornos virtuales de enseñanza y aprendizaje y aplicaciones móviles que se gestan atendiendo a necesidades específicas del ámbito educativo. Se estudian metodologías y herramientas de diseño y se entran saberes del área de Informática y Educación como parte de las actividades del proyecto. Se realizan desarrollos que forman parte de los resultados y que se integran a experiencias educativas específicas. Además, se participa en la formación de recursos humanos en el área y en la cooperación con otras universidades del país y del exterior.

Palabras clave: materiales educativos digitales, entornos digitales para el aprendizaje, juegos serios, herramientas colaborativas, educación, educación especial

CONTEXTO

El presente trabajo está vinculado al subproyecto “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos”, que pertenece al proyecto titulado: “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” (período 2018-2022), y su continuidad en el proyecto: “Diseño, desarrollo y evaluación de sistemas en escenarios híbridos para áreas clave de la sociedad actual: educación, ciudades inteligentes y gobernanza digital”.

Todos se desarrollan en el Instituto de Investigación en Informática LIDI, de la Facultad de Informática de la Universidad Nacional de La Plata.

1. INTRODUCCIÓN

En el proyecto que aquí se presenta se abordan diferentes líneas de investigación, desarrollo e innovación; en algunas de ellas se viene trabajando desde ya hace varios años y se profundizan de acuerdo a los avances en el área. Tales líneas se vinculan con sistemas colaborativos para el ámbito educativo, diseño y desarrollo de entornos virtuales de enseñanza y aprendizaje, y el desarrollo de materiales educativos digitales. En los últimos años se ha puesto foco en particular, en el estudio de metodologías de diseño de experiencias colaborativas mediadas por tecnologías digitales, y en el desarrollo de herramientas específicas para el seguimiento del proceso colaborativo. El trabajo colaborativo se presenta como un tipo especial de trabajo en grupo que promueve el desarrollo de habilidades de aprendizaje, personales y sociales, en el que cada miembro expresa compromiso tanto con su propio aprendizaje como con el de los demás miembros de su grupo (Zangara y Sanz, 2020). De acuerdo a Dillenbourg (1999) en los procesos colaborativos intervienen componentes tales como la interdependencia positiva, la interacción, la contribución individual y las habilidades personales y de grupo. Al ser mediados por herramientas tecnológicas digitales, existen factores claves a considerar como el establecimiento de protocolos y sesiones de comunicación, tanto sincrónicas como asincrónicas, y el uso de herramientas que facilitan el seguimiento del proceso. Se aplican así técnicas de *mirroring*, que posibilitan reflejar de forma dinámica y mediante gráficos, líneas de tiempo y otros, el desempeño individual y colectivo de un grupo, con la consideración

de métricas e indicadores de interés. Van Leeuwen y Rummel (2020), por ejemplo, utilizan *mirroring* y, con diferentes gráficos, proporcionan información de la interacción de los participantes. Diversos estudios muestran beneficios del *mirroring* para procesos colaborativos (Fidalgo-Blanco et. al, 2013; Ullman et al., 2019; Zangara y Sanz, 2020). En estas líneas se viene avanzando en el marco del proyecto que aquí se presenta, y también se vinculan con investigaciones relacionadas con la autorregulación del aprendizaje en actividades mediadas por tecnologías digitales. El trabajo de Dieser, Sanz y Zangara (2022 y 2020) aportan en este sentido y forman parte de los resultados de este proyecto. Se analizan diversos estudios que abordan la autorregulación, sus dimensiones, e instrumentos y técnicas que permiten dar cuenta de cómo los estudiantes llevan adelante este proceso en actividades educativas mediadas por tecnologías digitales.

Otra línea en la que se está investigando es la de entornos virtuales para la enseñanza y el aprendizaje. A partir de la situación de pandemia se ha visto un incremento del uso de diferentes entornos virtuales para mediar procesos educativos (Spandre, Dieser y Sanz, 2022). El diseño y desarrollo de este tipo de sistemas centrados en la web, puede aportar a la mejora de mediaciones con herramientas que asisten tanto al docente como al estudiante. En relación a esto, se viene trabajando en el estudio de herramientas y metodologías para gamificar dichos entornos (Vera et al. 2023), así como en indicadores de participación que den cuenta de la presencia social, cognitiva y docente, mencionadas en el modelo de Comunidad de Indagación, en el que sustentan estas líneas (Garrison, Cleveland-Innes y Fung, 2010). También se investigan y diseñan entornos 3D para diferentes procesos de aprendizaje, ya que aportan a la exploración, la estimulación audiovisual y a la simulación. Al mismo, tiempo los materiales educativos digitales, y en particular los objetos de aprendizaje forman parte de la agenda de cooperación y estudio actuales (Astudillo, Sanz y Santa Cruz Valencia, 2017; Carvalho y Yeoman, 2021).

Finalmente, los juegos serios educativos (JSE) (Dörner et al., 2016) constituyen otra de las líneas que se viene abordando en el proyecto. Los JSE aportan resultados en cuanto a la motivación de los estudiantes, en el desarrollo de determinadas competencias y habilidades (Baranyi et al., 2020) y también se utilizan en la formación y actualización de docentes (Sandí Delgado et al., 2022). Se estudian en particular metodologías de diseño de JSE, con el fin de atender a los objetivos educativos, de entretenimiento y de co-diseño que se propone alcanzar (Archuby, Sanz y Manresa-Yee, 2023 y 2020; Gros y Durall, 2020).

Las líneas de I+D+I que se abordan se vinculan con las temáticas de la Maestría y Especialización en

Tecnología Informática Aplicada en Educación de la misma Facultad.

2. LÍNEAS DE INVESTIGACIÓN / DESARROLLO/ INNOVACIÓN

Se presentan aquí las principales líneas de investigación, desarrollo e innovación abordadas en el marco del proyecto:

- Entornos digitales para la mediación de procesos educativos: entornos virtuales de enseñanza y aprendizaje, entornos 3D, redes sociales, entornos virtuales gamificados. Funcionalidades y formas de acceso a estos espacios, trazabilidad de las actividades, estándares. Estrategias para el diseño e implementación de estos tipos de entornos.
- Materiales educativos digitales. Metodologías para su diseño y producción. Objetos de aprendizaje. Multimedia e hipermedia en escenarios educativos. Nuevos entramados de medios, soportes y lenguajes.
- Integración de TIC en procesos educativos. Hibridación de las modalidades educativas. Diseño de MOOC. E-actividades. Actividades transmedia.
- Juegos Serios Educativos con diferentes paradigmas de interacción. Metodologías para su creación. Objetos activos.
- Trabajo colaborativo mediado por TICs. Herramientas para promover la autorregulación y el desarrollo de capacidades metacognitivas. Conceptualización, análisis y desarrollo de software y metodologías. Actividades colaborativas aprovechando dispositivos móviles.
- Formación de recursos humanos en el área de Informática. Pensamiento computacional, desarrollo de la empatía, y la resiliencia académica en estudiantes de carreras en Informática.

3. RESULTADOS OBTENIDOS

Se detallan aquí los resultados alcanzados durante el año 2022 e inicios de 2023.

En relación a la línea de entornos virtuales de enseñanza y aprendizaje, se continuó con el desarrollo de una nueva versión del entorno IDEAS¹, que será puesta en producción durante este año. También se avanzó en una tesis doctoral sobre el uso de entornos 3D para procesos educativos con personas con discapacidad auditiva. En el marco de esta tesis se cuenta con diferentes escenarios desarrollados en OpenSim para abordar una secuencia de actividades educativas específicas (Fachal, y Abásolo, 2021a; Fachal, Abásolo y Sanz, 2021b). En 2022, se llevaron adelante sesiones de trabajo con diferentes grupos de personas usando la propuesta de actividades creadas. En primer lugar, se realizó una

¹ Proyecto IDEAS:

<https://proyectoideas.info.unlp.edu.ar/>

prueba piloto con un grupo de 6 participantes con discapacidad auditiva, con un buen nivel de lectura de español. A partir de esta prueba piloto se profundizó el trabajo y se realizó una nueva experiencia, de forma completa, con todos los escenarios preparados en el entorno OpenSim. Se trabajó con 10 personas con discapacidad auditiva. Finalmente, a fines de 2022, se consolida el estudio con un grupo presencial en el contexto del IIIDE, un centro educativo de Nivel Medio y Terciario que trabaja por la Integración de las personas sordas o hipoacúsicas y oyentes (ver Fig.1). Los resultados de estas experiencias se están analizando y se espera finalizar la tesis durante este año. Se cuenta con investigaciones previas relacionadas a estas temáticas y forman parte de la cooperación con otras universidades del país y el exterior (Herrera, Manresa-Yee y Sanz, 2021).



Fig. 1. Experiencia en el marco de la integración de personas sordas o hipoacúsicas y oyentes en entornos virtuales 3D

Además, se realizaron avances en los estudios teóricos de una tesis doctoral sobre gamificación en entornos virtuales de enseñanza y aprendizaje. A partir de una revisión sistemática, se abordó el análisis de estado del arte en la temática y los resultados se publicaron en revista (Vera et al., 2023).

En 2022 se comenzó un trabajo de tesina de grado que implementará un sistema de notificaciones para celulares en Moodle, sobre el desarrollo de AulasWebColegios de la UNLP. Esta propuesta se orienta a facilitar la comunicación entre los usuarios del sistema, quienes pueden recibir, borrar y destacar notificaciones en su celular (Gonzalez, Ungaro, y Romanut, 2022).

En la misma línea, en 2022 se finalizó un trabajo de Especialización en Tecnología Informática Aplicada en Educación vinculado a la mediación, a través de entornos virtuales, de procesos de formación de usuarios de biblioteca (Kessler y Sanz, 2022).

Además, se trabaja en la finalización de una tesis de Maestría Tecnología Informática Aplicada en Educación vinculada con los entornos rurales a distancia donde se busca identificar y describir las estrategias que desarrollan los estudiantes de nivel educativo secundario con modalidad rural, en un contexto de inmersión tecnológica obligatoria.

Se continuó con los avances de una tesis doctoral de uno de los miembros del proyecto con el desarrollo de MarCOA, un framework para la creación de objetos de aprendizaje, que sigue la metodología CROA. Se realizó una nueva publicación sobre este framework que da

cuenta de su propuesta y su evaluación (Violini, Sanz, y Pesado, 2022). También se continuó con la dirección de una práctica profesional vinculada al desarrollo de MarCOA (Maccarini, Violini y Sanz, 2022).

Se avanzó también en el diseño de una plataforma para Esports, en el marco de una tesina de grado (Handula, Sanz e Iglesias, 2022). Esta plataforma brinda estadísticas para el entrenamiento de los jugadores y permite la gestión de torneos de esports.

En 2022 se avanzó con el proyecto Cap4city (Erasmus), en el que se aplicó la metodología de creación de MOOC elaborada en el equipo (Zangara y Sanz, 2020). Se participó de un encuentro de difusión de resultados del proyecto donde se presentó la metodología propuesta y las capacitaciones realizadas durante 2020 y 2021.

En relación a los avances de las líneas de investigación referidas al desarrollo de materiales educativos digitales a fines de 2022 se presentó, para su evaluación y exposición, el trabajo de tesis de Maestría en Tecnología Informática aplicada en Educación en el que se desarrolló un laboratorio virtual y se evaluó la efectividad que poseen las actividades prácticas de electrónica básica al realizarse con laboratorios virtuales dentro de un aula extendida (Navarria, Gonzalez, y Zangara, 2021).

En cuanto a los avances en las líneas vinculadas a trabajo colaborativo, metodologías y herramientas para su seguimiento, se avanzó en una tesis doctoral en la que se desarrolló un *dashboard* para Moodle con el fin de implementar estrategias de *mirroring* (Vazquez-Bermudez et al., 2021). También se trabajó en el diseño de una encuesta para su evaluación durante experiencias de uso concretas de este *dashboard*. Se espera completar los estudios de caso de la tesis durante este año.

Se está finalizando una tesis de maestría relacionada también con la mediación de actividades educativas a través de tecnologías digitales, y el estudio de los procesos de autorregulación que se abordan en estos contextos; en particular en el microcontexto de una actividad de foro. Parte de los estudios de esta tesis de maestría fueron compartidos en una publicación en 2022 (Dieser, Sanz y Zangara, 2022). También se dirige una tesis de maestría en el que se estudian las interacción entre estudiantes y docentes, y estudiantes entre sí, a la luz del modelo de Comunidad de Indagación (CoI). Esta tesis ya está en la fase final, y se está plasmando la escritura de sus resultados (Spandre, Dieser y Sanz, 2022).

En el marco de un proyecto de innovación docente de la Universidad de Zaragoza en el que se participa, se viene trabajando en la creación de un modelo para el desarrollo de la competencia de empatía en estudiantes de carreras en Informática, que ha sido aplicado en diferentes experiencias (Sanz et al., 2022; Aguelo et al., 2022).

En cuanto a la línea de juegos serios educativos, se avanzó en 2022 en la investigación, diseño y desarrollo de juegos. Es de resaltar que a partir de una revisión

sistemática sobre metodologías de diseño de juegos, se abordó el diseño de DIJS que es la metodología aplicada en el marco del proyecto (Archuby et al., 2023). Así se gestaron juegos como EmpoderAR, basado en realidad aumentada y que resalta el aporte de mujeres que han sido innovadoras en la disciplina informática (EmpoderAR, 2022). Este juego está disponible en el Centro de Innovación y Transferencia Tecnológica y se utilizó en visitas educativas con estudiantes de escuela y facultad. Al mismo tiempo, se presentó: “Escapados” en diferentes muestras y espacios educativos. Se trata de un juego desarrollado, como parte del proyecto, en 2021 para el aprendizaje de temas de Matemática y Química. Todos los JSE desarrollados se vienen utilizando como parte de las actividades del Centro de Innovación y Transferencia Tecnológicas de la Facultad de Informática².

En cuanto a los proyectos vinculados con la temática y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se cuenta con un acuerdo de colaboración con la Universidad de Zaragoza y se trabaja en forma conjunta en estas temáticas.
- Se tiene un acuerdo de cooperación con la Universidad de Islas Baleares, en particular se trabaja con el Departamento de Ciencias Matemáticas e Informática. En este contexto se dirigen tesis y se participa de proyectos conjuntos.
- Se participa del Proyecto Nexos de articulación entre la escuela y la universidad, a partir del cual se llevan a cabo acciones con diferentes escuelas de la región de La Plata y se participa de ferias y exposiciones de ciencia, tecnología y educación.
- Se coopera con la Universidad Nacional de Santiago del Estero con asesoría en vinculación algunos de los ejes presentados aquí.
- Se ha participado en el proyecto PERGAMEX y en particular en el subproyecto RTI2018-096986-B-C31: “Design of pervasive gaming experiences for intergenerational social and emotional well-being (PERGAMEX-INTERGEM)”.
- Se participa en proyectos de Innovación Docente de la Universidad de Zaragoza.
- Se viene participando del proyecto Erasmus Cap4city, a través de diferentes actividades en los últimos 3 años.

4. FORMACION DE RRHH

Se aborda la formación de recursos humanos a través de la dirección de becas, tesis de doctorado, maestría y trabajos finales de especialización y de grado.

En 2022, se cuenta como resultado con 1 tesis de maestría entregada para su evaluación, 1 trabajo de especialización finalizado, 1 tesina de grado finalizada, y 3 tesis doctorales en curso en relación a la temática. Se

cuenta con una tesina de grado por finalizar para este año.

5. BIBLIOGRAFÍA

- Aguelo, A., Sanz, C., Coma, T., Álvarez, P., Baldassarri, S. (2022) Application of the SPC model for the development of empathy in university classrooms, *EDULEARN22 Proceedings*, pp. 8651-8654.
- Archuby, F., Sanz, C. y Manresa-Yee, C. (2023) "DIJS: Methodology for the Design and Development of Digital Educational Serious Games," in *IEEE Transactions on Games*, doi: 10.1109/TG.2022.3217737.
- Archuby, F., Sanz, C. y Manresa-Yee, C. (2020). Metodologías de diseño y desarrollo para la creación de juegos serios digitales. Tesis de maestría finalizada. Sep. 2020. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/111123>
- Astudillo, G., Sanz, C. y Santacruz-Valencia, L. (2017). Proceedings of the Twelfth Latin American Conference on Learning Technologies (LACLO), ISBN: 978-1-5386-2376-3, págs. 1-4, doi: 10.1109/LACLO.2017.8120939.
- Carvalho, L.; Yeoman, P. (2021) Performativity of Materials in Learning: The Learning-Whole in Action. *Journal of New Approaches in Educational Research*, [S.l.], v. 10, n. 1, p. 28-42, jan. 2021. ISSN 2254-7339. Disponible en: <https://naerjournal.ua.es/article/view/627>
- Baranyi, R., Czech, P. Hofstätter S, Aigner, C. y Grechenig, T. (2020). “Analysis, design, and prototypical implementation of a serious game reha@ stroke to support rehabilitation of stroke patients with the help of a mobile phone,” *IEEE Transactions on Games*, 2020.
- Dieser, P., Sanz, C., Zangara, A. (2022). Metodologías para la evaluación de la autorregulación del aprendizaje en contextos educativos mediados por tecnología digital: una revisión sistemática. Congreso de Tecnología en Educación y Educación en Tecnología 2022, Disponible en: <http://sedici.unlp.edu.ar/handle/10915/139902>
- Dieser, P., Sanz, C., Zangara, A. (2020). Capítulo: Autorregulación del aprendizaje en contextos educativos mediados por tecnologías digitales. Teoría, investigación y aplicaciones en la educación superior iberoamericana. Libro: Aprendizaje y tecnologías: habilidades del presente, proyecciones de futuro. Editorial Noveduc. 978-987-538-764-5
- Dillenbourg, P. (1999). “What do you mean by collaborative learning?,” in *Collaborative-learning: Cognitive and Computational Approaches*, Ed., Oxford: Elsevier, 1999, ch 1, pp.1-19.

² CIyTT: <https://ciytt.info.unlp.edu.ar/>

- Dörner, R., Effelsberg, W., Göbel, S., Wiemeyer, J. (2016). *Serious Games. Foundations, concepts and practice*. Springer. Alemania.
- EmpoderAR (2022). Divulgación en medios: <https://unlp.edu.ar/investiga/cienciaenaccion/cuando-los-jueguitos-electronicos-son-un-motor-del-aprendizaje-43461/>
- Fachal, A. y Abásolo, M.J. (2021a). Entorno Virtual 3D en OpenSim para el trabajo con estudiantes con discapacidad auditiva. Demo presentada en el Congreso TEyET 2021.
- Fachal, A., Abásolo, M. J., y Sanz, C. (2021b). Dictionary of Computer Terms in LSA with Operational Signs proposed by and for Hearing-Impaired Students. *IEEE Rev. Iberoamericana de Tecnologías del Aprendizaje*.
- Fidalgo-Blanco, A., Leris, D., Sein-Echaluce, F. J. y García-Peñalvo, M.L. (2013). Indicadores para el seguimiento y evaluación de la competencia de trabajo en equipo a través del método CTMTC, [Online] Disponible en: <https://gredos.usal.es/handle/10366/122531>
- Garrison, R., Cleveland-Innes, M., Fung, T. (2010) Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework”, *The Internet and Higher Education*, vol. 13, n.1-2, 31--36. DOI: 10.1016/j.iheduc.2009.10.002
- Gonzalez A. H., Ungaro L., y Romanut L.(2022). Estrategias de comunicación en escenarios educativos híbridos: implementación y mejoras al sistema de notificaciones push de la aplicación de Moodle para AulasWebColegios de la UNLP. CACIC 2022.
- Gros, B., y Durall, E. (2020). Retos y oportunidades del diseño participativo en tecnología educativa. *EduTec. Revista Electrónica De Tecnología Educativa*, (74), 12-24. <https://doi.org/10.21556/edutec.2020.74.1761>
- Herrera, S. I., Manresa-Yee, C., y Sanz, C. V. (2021). Mobile learning for hearing-impaired children: Review and analysis. *Universal Access in the Information Society*, 1-19.
- Kessler, I. y Sanz, C. (2022). Las estrategias de formación de usuarios mediadas por tecnología informática en bibliotecas de instituciones de educación superior. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/135900>
- Handula, L., Sanz, C., y Iglesias, L. (2022). Desarrollo de una Plataforma Web para eSports orientada a la gestión de competencias y al entrenamiento competitivo Propuesta de tesina de grado, en ejecución. Aprobada.
- Maccarini, E., Violini, L., Sanz, C. Informe de avance del Proyecto. (2022)
- Navarria, L., González, A., y Zangara, A. (2021). Laboratorios virtuales de electrónica básica para alumnos universitarios dentro del aula extendida. *Trayectorias Universitarias*, 7(13), 075-075.
- Sandí-Delgado, J. Sanz, C. y Lovos, E. (2022). Acceptance of Serious Games to Develop Digital Competencies in Higher Education. *The Electronic Journal of e-Learning*, 20(3), pp. 351-367, Disponible en: www.ejel.org
- Sanz, C., Coma-Roselló, T. Aguelo, A., Álvarez, P y Baldassarri, S. (2022) "Model and Methodology for Developing Empathy: An Experience in Computer Science Engineering," in *IEEE Trans. on Education*, doi: 10.1109/TE.2022.3231559.
- Spandre, O., Dieser, P., Sanz, C. (2022). Systematic Review of Educational Methodologies Implemented During the COVID-19 Pandemic in Higher Education in Ibero-America. In: Pesado, P., Gil, G. (eds) *Computer Science – CACIC 2021. Communications in Computer and Information Science*, vol 1584. Springer, Cham. https://doi.org/10.1007/978-3-031-05903-2_4
- Ullmann, T. D., De Liddo, A., y Bachler, M. (2019). A Visualisation Dashboard for Contested Collective Intelligence Learning Analytics to Improve Sensemaking of Group Discussion. *RIED. Rev. Iberoamericana de Educación a Distancia*, 22(1), 41. <https://doi.org/10.5944/ried.22.1.22294>
- van Leeuwen, A. y Rummel, N. (2020) Comparing teachers' use of Mirroring and advising dashboards. In: *Proc. of the Tenth International Conference on Learning Analytics & Knowledge*, Frankfurt Germany, pp. 26–34, March 2020. <https://doi.org/10.1145/3375462.3375471>
- Vázquez-Bermúdez, M., Sanz, C., Zangara, M. A., y Hidalgo, J. (2021). Visualization Tools for Collaborative Systems: A Systematic Review. In *International Conf. on Technologies and Innovation* (pp. 107-122). Springer, Cham.
- Vera Mora, G. R., Sanz, C., Baldassarri, S. y Coma, T. (2023). Entornos virtuales de enseñanza y aprendizaje gamificados a la luz del concepto de presencia: Revisión sistemática de literatura. *Rev. Iberoamericana TEyET*, (33), e3. <https://doi.org/10.24215/18509959.33.e3>
- Violini, M. L., Sanz, C. V., y Pesado, P. M. (2022). Prototipo de un framework para la creación de Objetos de Aprendizaje y su evaluación mediante juicio de expertos. *EduTec. Rev. Electrónica De Tecnología Educativa*, (81), 137-154. <https://doi.org/10.21556/edutec.2022.81.2529>
- Zangara, M. A., y Sanz, C. V. (2020). Trabajo colaborativo mediado por tecnología informática en espacios educativos: Metodología de seguimiento y su validación. *Revista Iberoamericana De Tecnología En Educación Y Educación En Tecnología*, (25), e1. <https://doi.org/10.24215/18509959.2>

TECNOLOGÍAS DISRUPTIVAS Y AULAS EMERGENTES EN EL NIVEL MEDIO Y SUPERIOR EN EL CONTEXTO DE LA UNRN

Edith Lovos¹; Martin Goin², Alejandra Marin¹, Noelia Verdun², Ivan Basciano, Milagros Salas

¹ Universidad Nacional de Río Negro, Sede Atlántica, Viedma, Río Negro

² Universidad Nacional de Río Negro, Sede Andina, San Carlos de Bariloche, Río Negro
Centro Interdisciplinario en Derechos, Inclusión y Sociedad (CIEDIS)
{elovos,mgoin,mmarin,nverdun}@unrn.edu.ar - {ivaanbas19,salasmilagros}@gmail.com

RESUMEN

A partir del contexto de pandemia, los docentes de los diferentes niveles educativos nos encontramos frente a la necesidad de rediseñar espacios, procesos, y formas de interacción involucradas en la enseñanza y aprendizaje. Así también, fue necesario revisar y repensar aspectos vinculados al diseño y producción de los materiales educativos digitales, atendiendo a las modalidades de las prácticas áulicas emergentes, a la vez que avanzar en la intención de promover espacios de enseñanza y aprendizaje más cercanos a las prácticas culturales de los estudiantes. En este sentido, en el proyecto que se presenta continuación, se buscó estudiar un conjunto de tecnologías que son consideradas disruptivas como la realidad aumentada, la realidad virtual y/o los juegos serios, partiendo de la premisa que no son estas tecnologías en sí mismas disruptivas sino las ideas/oportunidades a las que abren paso (Litwin; 2004).

Palabras clave: Materiales Educativos Digitales, Prácticas Áulicas Emergentes, Tecnologías Disruptivas.

CONTEXTO

Este trabajo presenta a modo de resumen, algunos resultados alcanzados en el marco del proyecto de investigación denominado “Materiales Educativos Digitales y

Tecnologías Disruptivas. Diseño, Desarrollo y Aplicación en los Contextos Educativos Emergentes”, aprobado y financiado por la UNRN. El mismo es de ejecución bianual y se inició en mayo de 2021.

El proyecto se lleva adelante en el marco de la unidad ejecutora: Centro Interdisciplinario de Estudios sobre Derechos Inclusión y Sociedad (CIEDIS) perteneciente a la Sede Atlántica de UNRN, específicamente en la línea de investigación sobre epistemología, investigación e innovación docente.

1. INTRODUCCIÓN

Diversos autores (Flavin, 2012, Cabero y Fernández, 2018, Coicaud, 2019, Zuñiga et al.,2021), coinciden en describir a las tecnologías disruptivas como aquellas que permiten reconfigurar, transformar y/o mejorar las prácticas establecidas en los ámbitos donde se las aplique, aun cuando no hayan sido diseñadas para los mismos. Entre estas tecnologías, es posible mencionar la realidad aumentada, la realidad virtual y/o los juegos serios entre otras. Sobre estas, diversos estudios y experiencias dan cuenta de los aportes que las mismas pueden hacer a las prácticas pedagógicas en sus diferentes modalidades, posibilitando el diseño de experiencias de enseñanza y aprendizaje más dinámicas, interactivas e inmersivas, y a la vez más afines a la cultura digital de los estudiantes

(Mazza, 2020, Ávila y Crespo, 2021, Castro et al., 2021, Mansilla et al., 2021, Romano y Moyano, 2021). Sin embargo, como cualquier otra tecnología su inclusión en la práctica docente, no está libre de obstáculos y desafíos, y uno de ellos es la necesidad de contar con Materiales Educativos Digitales (MED) que las incluyan y se ajusten al contexto específico de aplicación. Así, la producción de estos MED, demanda a los docentes de competencias digitales que les permitan avanzar en la tarea (López y Bernal, 2018), combinando el diseño instruccional con el diseño tecnológico, teniendo en cuenta que *“el potencial de las TIC para la promoción de las oportunidades de aprendizaje depende de las habilidades utilizadas para diseñar actividades de aprendizaje que se alineen con la pedagogía y la tecnología para el beneficio de los alumnos”* (Begonia Gross 2015, pp. 2). En este contexto, las herramientas de autor, adquieren un valor significativo, dado que ponen a disposición de los usuarios (docentes y/o estudiantes en este caso), la posibilidad de avanzar en la construcción de sus propios MED, ya sea desde cero y/o a partir de otros producidos por terceros sin necesidad de ser programadores (Lytridis, et al., 2018, Salazar Mesias et al., 2019, Vidal et al., 2019). Asimismo, y como señala Rodríguez (2021), poner el potencial de estas tecnologías al servicio del bien común, en este caso en el ámbito académico y en particular en contextos de desigualdad, requiere del compromiso de los estados para revisar y adecuar los marcos jurídicos y legales, y el diseño y aplicación de las políticas públicas que garanticen oportunidades, derechos e intereses de la comunidad en su conjunto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo principal del proyecto está puesto en generar conocimiento sobre el

diseño, desarrollo y aplicación de Materiales Educativos Digitales (MED) que incluyan tecnologías consideradas disruptivas entre ellas, la realidad aumentada (RA), la realidad virtual (RV) y/o los juegos serios, en escenarios educativos de nivel medio y superior en el contexto de la UNRN y afines, a las disciplinas a las que está vinculado el equipo de investigación. Esto último, con la intención de avanzar en la transferencia de resultados en forma concreta y focalizada. A partir del proyecto será posible conocer la influencia de las tecnologías antes mencionadas, en el diseño de propuestas didácticas, así como también el impacto en las experiencias áulicas. Así los interrogantes de la investigación los siguientes:

- ¿De qué manera el diseño de MED que incluyen tecnologías consideradas disruptivas puede apoyar los procesos de aprendizaje que se dan en los formatos presenciales, virtuales o híbridos?
- ¿Qué estrategias metodológicas deberían ser tenidas en cuenta en el diseño, la producción y aplicación de MED que incluyan tecnologías disruptivas en particular RA, RV y/o juegos serios?
- ¿Qué oportunidades y qué obstáculos se presentan en la integración de MED que incluyan tecnologías disruptivas en escenarios educativos específicos del contexto de UNRN?

El proyecto sigue la metodología de investigación-acción participativa (Boggiano y Rosekrnas, 2004). La misma se ajusta a la problemática abordada, a los objetivos de trabajo propuestos y a la necesidad de realizar aportes y transferencias concretas de la investigación.

3. RESULTADOS OBTENIDOS/ESPERADOS

En forma resumida se presentan los avances obtenidos hasta el momento:

- Revisión de antecedentes, abordando temas como: metodologías de diseño de MED que incluyan tecnologías disruptivas, prácticas de enseñanza con tecnologías disruptivas, focalizando en el contexto de nivel medio y superior, y en la formación técnica en particular.
- Diseño de un instrumento para recolección de datos, que busca recuperar información en relación a las prácticas pedagógicas con tecnologías y en particular con aquellas consideradas disruptivas en el contexto de la UNRN. El mismo se aplicó a través de una encuesta virtual durante 2022, con la participaron de 62 docentes de las 3 sedes académicas de la institución (Andina, Valle Medio y Alto Valle, y Atlántica). Actualmente se continúa trabajando en el informe de análisis de los datos recolectados.
- Exploración de aplicaciones móviles con posibilidades para ser utilizadas en prácticas kinésicas de promoción de la higiene postural. Esta actividad se realizó en el marco de una beca CIN, cuyo objetivo consiste en identificar las posibilidades, limitaciones y contraindicaciones que los docentes y estudiantes avanzados de la Lic. en Kinesiología y Fisiatría de la UNRN, encuentran en la experimentación con aplicaciones móviles destinadas al área, como recursos complementarios en actividades de promoción y educación postural.
- Diseño e implementación de 2 experiencias formativas una en modalidad virtual y otra en formato presencial destinada a los docentes de nivel medio de escuelas técnicas de la región norpatagónica. En ambas propuestas, se buscó generar un acercamiento experimental al concepto de tecnologías disruptivas en particular a la realidad aumentada, la realidad virtual y los juegos serios. A la vez que

generar un espacio de vinculación entre actores del sistema académico y científico de la UNRN y el colectivo de docentes de escuelas secundarias técnicas de la zona de influencia de la Sede Atlántica. La primera, se llevó adelante en modalidad virtual entre Octubre y Diciembre de 2021 con el aval de la UNRN y del Consejo de Educación de la provincia de Río Negro, y la segunda se llevó adelante entre Agosto y Noviembre de 2022 con docentes de una escuela técnica agropecuaria del partido de Patagones (Buenos Aires). En relación a la primera, los resultados se han presentado y discutido en el Congreso TEYET (Lovos, Marin y Basciano, 2022). Las experiencias dan cuenta del potencial como complemento educativo que los docentes advierten en este tipo de tecnologías en relación al abordaje de los contenidos, su explicación y el sentido lúdico del aprendizaje que promueven. Aunque también señalan diferentes desafíos a tener en cuenta, como los provenientes de las culturas institucionales y/o de la construcción de equipos de trabajo, y otros asociados a la exploración de los MED existentes y a las prácticas de enseñanza vigentes.

Por último, en el marco del V Congreso de la Asociación Argentina de Humanidades Digitales¹ llevado adelante el pasado noviembre, se dictó un taller sobre posibilidades de las tecnologías disruptivas para comunicar y divulgar la ciencia. Así como también, se presentaron algunas notas de la investigación en el 9º Seminario Internacional de Educación a Distancia.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores del proyecto se conformó inicialmente por tres docentes de la

¹ https://aahd.net.ar/wp-content/uploads/2022/11/PROGRAMA_Final.pdf

Sede Atlántica (Viedma y Sierra Grande) y cuatro docentes de la Sede Andina (San Carlos de Bariloche) de la UNRN. Luego se incorporaron 3 estudiantes de la Lic. en Sistemas, en dos de los casos, a través de becas de formación en investigación propias de UNRN, y en otro como estudiante avanzando en el marco del desarrollo del Trabajo Social Obligatorio que dispone la UNRN como requisito para obtener los títulos de grado. Así también, en 2022 se sumó una estudiante de la Lic. en Kinesiología y Fisiatría de la UNRN como becaria CIN. Por otra parte, una de las investigadoras co-dirige un Trabajo Final de Carrera de la Lic. en Arte y Sociedad de la Universidad Nacional Del Comahue (CURZA) que aborda el uso de la realidad aumentada para la producción de una experiencia artística sobre relatos sobrenaturales.

Es importante señalar que los investigadores llevan adelante su práctica docente en asignaturas del área de matemáticas, literatura y programación de computadoras a través de las carreras vinculadas a ingenierías, sistemas, administración y educación tanto en la modalidad presencial como virtual que ofrece la institución.

Asimismo, una de las integrantes del proyecto, ha finalizado y aprobado el trabajo final de la Especialización en Docencia Universitaria de la UNRN, en el que aborda el desarrollo de competencias blandas en carreras vinculadas a la ciencia informática.

Como actividades de extensión de este proyecto, se desprenden:

- el diseño y desarrollo del juego de mesa JAM en el marco de un proyecto de extensión, financiado por la Secretaría de Políticas Universitarias (SPU REC32-22UCT- SPU-14444), y destinado a promover el desarrollo del pensamiento computacional.
- el diseño e implementación de una propuesta de formación en algoritmia destinada a docentes de nivel primario y secundario de la provincia de Río Negro a través del Programa “Nuestra Escuela” del Consejo Federal de Educación y coordinado por el Instituto Nacional de Formación Docente

(INFoD). La primera implementación comenzó en noviembre de 2022 y se extenderá a mayo de 2023

En relación al juego JAM, en el mes de Octubre de 2022, se llevó adelante un taller sobre el uso y aplicación del mismo en el marco de las XI Jornadas Cordilleranas de Enseñanza de la Matemática realizadas en la ciudad de Esquel, destinado a docentes de nivel primario, primeros años de nivel medio y estudiantes de Educación.

5. REFERENCIAS

- Ávila, M. U., & Crespo, L. F. (2021). Herramientas de Realidad Aumentada para la conceptualización del límite de una función en un punto. In XVI Congreso de Tecnología en Educación & Educación en Tecnología-TE&ET 2021 (La Plata, 10 y 11 de junio de 2021).
- Boggino, N., y Rosekrans, K. (2004). Investigación-acción: reflexión crítica sobre la práctica educativa. Buenos Aires: HomoSapiens.
- Castro, Y. G., Duran, O. M., & Zamudio, M. T. (2021). Tecnologías disruptivas en educación virtual. Boletín Redipe, 10(7), 185-200.
- Coicaud, Silvia. (2019). Potencialidades didácticas de la inteligencia artificial: mediaciones tecnológicas para una enseñanza disruptiva (1.^a ed. papel). Noveduc, 128 pp.
- Flavin, M. (2012). Disruptive technologies in higher education. Research in Learning Technology, 20.
- Gros Begonia. (2015). The Dialogue Between Emerging Pedagogies and Emerging Technologies. Cap. 1. en Gros, B., Kinshuk, K., & Maina, M. The future of ubiquitous learning : learning designs for emerging pedagogies. Springer Publishing Company, Incorporated ©2015
- Litwin, E. (2004) Prácticas con tecnologías. Praxis Educativa (Arg), núm. 8, pp. 10-17. Universidad Nacional de La Pampa.
- Lytridis, C., Tsinakos, A., & Kazanidis, I. (2018). ARTutor—an augmented reality platform for interactive distance learning. Education Sciences, 8(1), 6

- López, M., y Bernal, C. (2018). El perfil del profesorado en la Sociedad Red: reflexiones sobre las competencias digitales de los y las estudiantes en Educación de la Universidad de Cádiz. *International Journal of Educational Research and Innovation. IJERI*, (11), 83-100.
- Lovos, E., Marin, A., & Basciano, I. (2022). Tecnologías disruptivas en el nivel medio emergente: una experiencia de formación con docentes de escuelas técnicas. In XVII Congreso de Tecnología en Educación & Educación en Tecnología-TE&ET 2022 (Entre Ríos, 15 y 16 de junio de 2022).
- Mansilla, C. M. ., Becchio, R. M. ., Schaspchuck, P. ., Ordoñez, V. ., Brizi, M. C. ., Peralta, G. ., Paye, I. ., Regis, A. ., Blatter, M. C. ., & Cámara, C. . (2021). Realidad aumentada y blog en el ciclo básico de carreras ciclo básico de carreras ingeniería *Revista De Enseñanza De La Física*, 33, 421–429. Recuperado a partir de <https://revistas.unc.edu.ar/index.php/revistaEF/article/view/35535>
- Mazza, M. A. (2020). Juego serio con realidad virtual para jóvenes orientado a conocer hitos de la historia de la Informática (Doctoral dissertation, Universidad Nacional de La Plata).
- Rodríguez, Gladys Stella. (2021). Tecnologías disruptivas: Contexto Político-Jurídico, Desafíos y Oportunidades en Latinoamérica. *Lex-Revista De La Facultad De Derecho y Ciencias Políticas*, 19(28), 41-74.
- Romano, L., & Moyano, E. (2021). Aplicación de herramienta de realidad aumentada para la enseñanza de programación en el nivel superior. In XXVII Congreso Argentino de Ciencias de la Computación (CACIC)(Modalidad virtual, 4 al 8 de octubre de 2021).
- Salazar Mesía, N. A., Sanz, C. V., & Gorga, G. M. (2019). Diseño de plantillas para la creación de actividades educativas con Realidad Aumentada en AuthorAR. In XXV Congreso Argentino de Ciencias de la Computación (CACIC)(Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019)
- Vidal Jr, E. C. E., Ty, J. F., Caluya, N. R., & Rodrigo, M. M. T. (2019). MAGIS: mobile augmented-reality games for instructional support. *Interactive Learning Environments*, 27(7), 895-907.
- Zúñiga, K. M., Rodríguez, A., & Velázquez, R. V. (2021). Tecnologías e innovación disruptiva en la educación superior. *Serie Científica de la Universidad de las Ciencias Informáticas*, 14(3), 177-186.

DESARROLLO Y APLICACIÓN DE LABORATORIOS VIRTUALES DESTINADOS A EDUCACIÓN DE GRADO Y PREGRADO

Godoy Pablo Daniel^{1,2,3,4}, Marianetti Osvaldo Lucio^{1,4}, Fontana Daniel^{1,4}, Manganelli Silvina²

¹Universidad Nacional de Cuyo, Facultad de Ingeniería

²Universidad Nacional de Cuyo, Facultad de Ciencias Exactas y Naturales

³Universidad Nacional de Cuyo, ITIC

⁴Universidad de Mendoza

pablo.godoy@ingenieria.uncuyo.edu.ar

RESUMEN

El presente artículo describe un proyecto aprobado y financiado por la Universidad Nacional de Cuyo, cuyo plazo de ejecución es entre 2022 y 2024.

El proyecto tiene como objetivos el desarrollo de laboratorios virtuales para ser utilizados en tareas docentes, realizar aportes en el campo de los laboratorios virtuales o su aplicación en tareas de enseñanza, y poner a prueba la hipótesis de que las prácticas educativas STEAM y las herramientas digitales puede servir para mejorar las competencias científicas, tecnológicas y digitales de los estudiantes.

El proyecto se encuentra en su etapa inicial de ejecución, estando en desarrollo dos laboratorios virtuales, uno de efecto Doppler y otro de ventanas deslizantes para la enseñanza de redes de computadoras. Se espera que el proyecto finalice con la implementación de un conjunto de laboratorios virtuales que serán de gran utilidad en las tareas de enseñanza de los docentes investigadores que forman parte del proyecto u otros docentes o estudiantes, con experiencias de aplicación que serán compartidas en congresos o revistas, y avances en una tesis doctoral.

Área: TIAE (Tecnología Informática Aplicada en Educación)

El presente trabajo describe un proyecto de investigación y desarrollo financiado por la Universidad Nacional de Cuyo a través de los proyectos "SIIP B008: Desarrollo y aplicación de laboratorios virtuales destinados a educación de grado y pregrado" y "SIIP B036-T1: La seguridad de los dispositivos utilizados en IoT mediante la utilización de Blockchain, NFT y/o VPN". El plazo de ejecución es entre 2022 y 2024.

El proyecto surge como una evolución de dos proyectos de investigación y la integración de un trabajo de tesis doctoral. Dichos proyectos se denominan "Implementación de laboratorios remotos basados en cloud computing" y "Laboratorio remoto y nómada de arquitectura de computadoras destinado a enseñanza e investigación". El primero de dichos proyectos tuvo como objetivo investigar e implementar un laboratorio remoto destinado específicamente a que puedan realizar sus trabajos prácticos estudiantes de la carrera Licenciatura en Ciencias de la Computación, dictada en la Facultad de Ingeniería de la Universidad Nacional de Cuyo. El segundo proyecto tuvo como objetivo ampliar las capacidades del primer laboratorio remoto en cuanto a equipamiento y prestaciones. Se adicionó nuevo equipamiento, se incluyeron las características de nómada (que pueda trasladarse físicamente al aula) y máquinas virtuales con la finalidad de permitir la interacción con computadoras de un

CONTEXTO

laboratorio típico de computación. Estos laboratorios remotos y nómadas fueron utilizados para que estudiantes de distintas carreras realicen sus trabajos prácticos, entre estas: Licenciatura en Ciencias de la Computación, Universidad Nacional de Cuyo, Instituto Tecnológico Universitario (ITU), Universidad Nacional de Cuyo y la Universidad de Mendoza. Pueden consultarse más detalles de dichos laboratorios remotos y experiencias de uso por parte de estudiantes de las carreras mencionadas en [1,2,3,4,5,6,7].

1. INTRODUCCIÓN

1.1 Objetivos

En este proyecto se propone, como objetivos principales, adicionar algunos experimentos virtuales (o simulados) a las plataformas ya disponibles. El fundamento es que existen experimentos importantes para algunas carreras dictadas en la Universidad Nacional de Cuyo para los cuales no se han creado laboratorios virtuales o simulaciones, o los existentes no poseen las características necesarias.

Como segundo objetivo principal, se espera colaborar con el avance de una tesis doctoral que tiene como título: “Desarrollo de competencias científico-tecnológicas en alumnos universitarios de carreras del área de las ciencias exactas mediante el método STEAM y el pensamiento computacional, necesarias para afrontar los desafíos del siglo XXI”. Se espera que las herramientas a desarrollar sean de utilidad para el desarrollo de la tesis mencionada. También se espera que dicha tesis pueda aportar una visión desde el punto de vista de las ciencias de la educación sobre la mejor manera de utilizar las herramientas a desarrollar.

Como objetivo secundario se propone mantener en funcionamiento el laboratorio remoto/nómada ya construido y que será

utilizado para el dictado de varias materias en las carreras mencionadas anteriormente.

1.2 Trabajos Relacionados

Existen un número importante de simulaciones y laboratorios virtuales accesibles a través de Internet. Se mencionan a continuación 3 ejemplos. Existen varios laboratorios virtuales de física, uno de los más importantes son las Simulaciones Interactivas de la Universidad de Colorado [8], en la cual se proveen laboratorios virtuales de varios tipos, por ejemplo, laboratorios de ondas, laboratorios de termodinámica y laboratorios de óptica física. La Universidad Técnica de Berlín provee un simulador de ventanas deslizantes de TCP ampliamente utilizado en cursos de redes de computadoras [9], [10]. El Instituto Tecnológico de Massachusetts provee un simulador de lenguaje ensamblador de 8 bits [11]. El mismo permite introducir a los estudiantes en el empleo de dicho lenguaje y de la arquitectura de procesadores.

1.3 Estado actual de avance del proyecto

El proyecto se encuentra en sus etapas iniciales de ejecución. Actualmente se trabaja en el desarrollo de dos laboratorios virtuales:

1) Laboratorio virtual de Efecto Doppler y/o Laboratorio virtual de ondas estacionarias. Existen en la bibliografía algunos laboratorios virtuales de física muy completos, pero carecen de algunas experiencias de laboratorio necesarias por estudiantes que cursan Física 2 en algunas carreras de la Universidad Nacional de Cuyo. Se propone implementar algunos experimentos virtuales que complementen los ya existentes. Se menciona que uno de los integrantes del grupo de trabajo del proyecto es profesor de Física 2 en la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de Cuyo.

2) Laboratorio virtual de ventanas deslizantes de redes de computadoras. Este es un tema

visto en la materia Redes de Computadoras de la carrera Licenciatura en Ciencias de la Computación, como también en otras carreras de redes o computación. Existen en la bibliografía algunos simuladores, pero carecen de detalles profundos del comportamiento del mecanismo, siendo estos detalles importantes para los estudiantes de la carrera mencionada. Se propone implementar un simulador o laboratorio virtual que muestre el comportamiento de este sistema de ventanas deslizantes. Se menciona que uno de los integrantes del grupo de trabajo del proyecto es profesor de la asignatura Redes de Computadoras, en la Facultad de Ingeniería de la Universidad Nacional de Cuyo.

Actualmente se avanza en la construcción de las interfaces de usuario de ambos laboratorios virtuales. En las figuras 1 y 2 se muestran capturas de pantalla de ambos laboratorios virtuales.

Los laboratorios virtuales a implementar estarán basados en tecnología web, de manera de poder ser accedidos a través de navegadores web típicos, desde cualquier computadora o teléfono celular.

Figura 1: Interfaz de usuario del laboratorio virtual de efecto Doppler

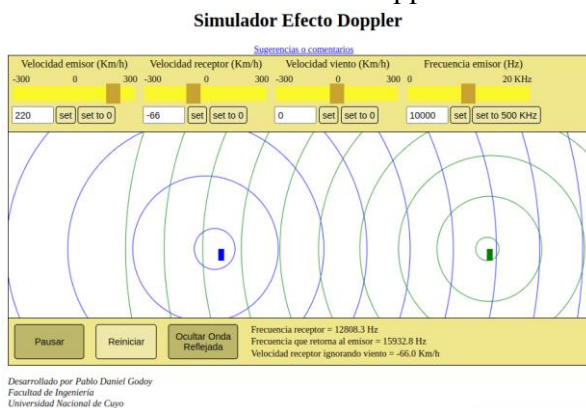
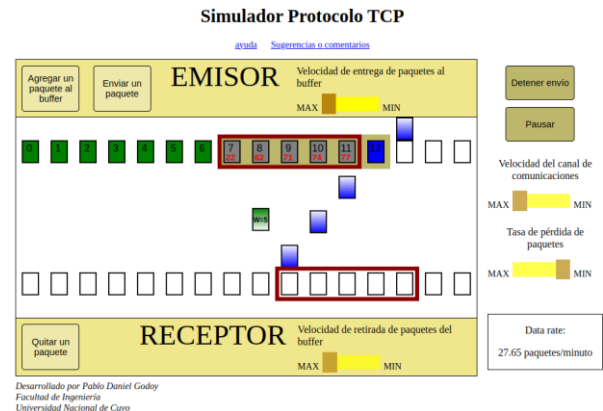


Figura 2: Interfaz de usuario del laboratorio virtual de ventanas deslizantes



2. LINEAS DE INVESTIGACION y DESARROLLO

2.1 Laboratorios virtuales

Estos se basan en un modelo físico o de funcionamiento de diferentes tipos de sistemas implementado mediante software [11], [12]. Tienen como ventaja que pueden escalar más fácilmente y no hay limitación sobre la cantidad de equipamiento disponible, sobre todo si se implementan en el cloud [13]. Además, si están disponibles a través de Internet, pueden estar disponibles todos los días del año, las 24 horas. Tienen como desventajas que los modelos pueden no ser exactos y los estudiantes no trabajan sobre equipos ni sistemas reales. Sin embargo, su utilización ha demostrado a los investigadores que forman parte de este proyecto que pueden ayudar en gran medida a que sus estudiantes comprendan fácilmente conceptos de difícil entendimiento.

2.2 Prácticas educativas STEAM

Actualmente existe un amplio abanico de herramientas digitales que pueden usarse en la enseñanza de la ciencia, la tecnología, la ingeniería y la matemática (educación STEAM). Como parte del proyecto, una tesis doctoral en curso tiene como objetivo probar que una adecuada relación entre las prácticas educativas STEAM y las herramientas digitales puede servir para mejorar tanto las

competencias científicas tecnológicas de los estudiantes como sus competencias digitales, necesarias para el desarrollo personal y profesional en la era digital. Para comprobar esta hipótesis, se han diseñado instrumentos con pautas para evaluar el desarrollo de competencias científicas, tecnológicas, digitales y las necesarias para afrontar los desafíos del siglo XXI, en alumnos pertenecientes a carreras del área de las ciencias exactas e ingeniería [14]. Por otro lado, este trabajo de tesis doctoral requiere de las herramientas para llevar adelante prácticas STEAM, entre estas, los simuladores.

2.3 Laboratorios remotos y nómadas

Permiten a los estudiantes acceder a equipamiento real a través de Internet o una red LAN. Tienen como ventaja que los estudiantes trabajan con sistemas reales, usualmente el tiempo durante el cual los equipos están disponibles es mayor que en una experiencia de laboratorio presencial, y los estudiantes pueden estar en ubicaciones alejadas, incluso en otros países. La característica de nómada permite que el equipamiento pueda trasladarse al aula o laboratorio para que los estudiantes puedan interactuar físicamente con el mismo, combinando acceso remoto y presencial. Tienen como desventaja que el costo del equipamiento y el acceso al mismo a través de Internet tiene un costo económico, por lo que la cantidad de equipamiento y el acceso al mismo es limitado.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1 Contar con herramientas que ayuden en el proceso de enseñanza

Los investigadores que forman parte de este proyecto son docentes de varias asignaturas: Arquitectura de Computadoras, Arquitecturas Distribuidas, Redes de Computadoras,

Sistemas Embebidos, Física 2, etc. en varias facultades de la Universidad Nacional de Cuyo y de la Universidad de Mendoza. Mediante un proceso exhaustivo de búsqueda de herramientas que asistan en el proceso de enseñanza, se ha detectado que no hay disponibilidad de los laboratorios virtuales propuestos. Por lo que el principal resultado esperado es contar con estas herramientas para ser utilizadas durante el dictado de clases.

3.2 Mejorar competencias científicas, tecnológicas y digitales de los estudiantes

Se espera probar la hipótesis de que la adecuada relación entre las prácticas educativas STEAM y las herramientas digitales puede servir para mejorar tanto las competencias científicas tecnológicas de los estudiantes como sus competencias digitales.

3.3 Realizar aportes en las áreas de los laboratorios virtuales y enseñanza

Se espera que el desarrollo de los laboratorios virtuales propuestos y su aplicación en tareas de enseñanza genere conocimientos y resultados útiles para la comunidad científica. Por tal motivo, los conocimientos generados, experiencias de aplicación en tareas docentes y resultados obtenidos serán publicados en congresos y revistas. Por otro lado, los propios laboratorios virtuales a desarrollar estarán disponibles libremente para su utilización por parte de profesores o estudiantes que deseen utilizarlos.

AGRADECIMIENTOS

Los autores agradecen el financiamiento recibido de la Secretaría de Investigación, Internacionales y Posgrado (SIIP), Universidad Nacional de Cuyo, a través de los proyectos "SIIP B008: Desarrollo y aplicación de laboratorios virtuales destinados a educación de grado y pregrado" y "SIIP B036-T1: La seguridad de los dispositivos utilizados

en IoT mediante la utilización de Blockchain, NFT y/o VPN”.

5. BIBLIOGRAFIA

[1] Godoy, Pablo; Marianetti, Osvaldo; García Garino, Carlos; "Experiences With Computer Architecture Remote Laboratories"; en Handbook of Research on Software Quality Innovation in Interactive Systems”, Chapter 5; IGI Global, Hershey, Pennsylvania, USA, 2020.

[2] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "A WSN testbed for teaching purposes"; IEEE Latin America Transactions. 14(7),3351–3357 (2016).

[3] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "Communication channel occupation and congestion in wireless sensor networks"; Computers & Electrical Engineering. 72, 846–858 (2018).

[4] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "Laboratorio remoto para la formación de usuarios basado en el cloud"; Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación 18, 7–18 (2016).

[5] Godoy, Pablo; García Garino, Carlos; Cayssials, Ricardo; "A nomadic testbed for teaching computer architecture"; In: XVII Workshop Tecnología Informática Aplicada en Educación (WTIAE), XXIV, CACIC, 2018.

[6] Godoy, Pablo; Cayssials, Ricardo; García Garino, Carlos; "Zigbee WSN Round Trip Latency in Function of Channel Occupation and Nodes Configuration"; In IEEE ArgenCon 2016.

[7] Godoy, Pablo Daniel; Marianetti, Osvaldo; Mansilla Roberto Alejandro; García Garino, Carlos; "Estudio de herramientas para la implementación de laboratorios remotos"; en XXIII Workshop de Investigadores en Ciencias de la Computación (WICC) 2021, Chilecito, La Rioja.

[8] Universidad de Colorado, PhET Interactive Simulations website. Accedido en febrero de 2022.

[9] Telecommunication Networks Group, Universidad Técnica de Berlín, simulador de ventanas deslizantes de TCP:

[10] López-Pedroza I.A., de Asís López-Fuentes F. (2021) Teaching of TCP Fundamental Operations Using a Digital Tool. In: Computer Science and Health Engineering in Health Services. COMPSE 2020. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 359. Springer, Cham.

[11] Darrah, M., Humbert, R., Finstein, J. et al. Are Virtual Labs as Effective as Hands-on Labs for Undergraduate Physics? A Comparative Study at Two Major Universities. J Sci Educ Technol 23, 803–814 (2014).

[12] Dongfeng Liu, Priscila Valdiviezo-Díaz, Guido Riofrio, Yi-Meng Sun, Rodrigo Barba, Integration of Virtual Labs into Science E-learning, Procedia Computer Science, Volume 75, 2015, Pages 95-102, ISSN 1877-0509.

[13] José A. González-Martínez, Miguel L. Bote-Lorenzo, Eduardo Gómez-Sánchez, Rafael Cano-Parra, Cloud computing and education: A state-of-the-art survey, Computers & Education, Volume 80, 2015.

[14] Manganelli, Silvina. Aportes de las herramientas digitales a STEM, durante la investigación científica, para fomentar el desarrollo de competencias científicas, tecnológicas y digitales. Congreso Argentino de Ciencias de la Computación (CACIC). Tecnología Informática Aplicada en Educación (WTIAE). 2021.

LA MEJORA DEL APRENDIZAJE EN AVAs APLICANDO ANALÍTICA DEL APRENDIZAJE

Delia Esther Benchoff, Ivonne Gellon, Carlos Guillermo Lazurri, Constanza R. Huapaya, Erik Borgnia Giannini y Matias H. Gutierrez

Grupo de Investigación Inteligencia Artificial aplicada a Ingeniería / Departamento de Informática/
Facultad de Ingeniería/ Universidad Nacional de Mar del Plata

Juan B. Justo 4302, Mar del Plata, Buenos Aires, Argentina.

{ebenchoff, guillermol, ivonne}@fi.mdp.edu.ar

constanza.huapaya@gmail.com, {eborgnia, mhgutierrez}@alumnos.fi.mdp.edu.ar

RESUMEN

Los ambientes virtuales de aprendizaje (AVAs), reúnen gran cantidad de datos sobre los estudiantes, los cuales son difíciles de analizar manualmente, por lo que se necesitan herramientas y métodos que contribuyan a explorar y aprovecharlos en pos de mejorar el aprendizaje de los usuarios. La Analítica del Aprendizaje (AA) se enfoca en la toma de decisión apoyada en datos, e integra las dimensiones técnicas y pedagógicas del aprendizaje mediante la aplicación de modelos predictivos. Con los avances de AA, especialmente de las técnicas de Machine Learning, puede mejorarse la personalización y adaptación del aprendizaje en AVAs. Los aspectos a enriquecer en la adaptación son: estilos de aprendizaje, estados cognitivos, evaluaciones de proceso y finales, retroalimentación y materiales instruccionales. Las universidades buscan predecir el rendimiento de los estudiantes para minimizar el abandono de sus estudios. La exploración de la personalización avanzada será usada para indagar y corregir las fallas académicas.

Palabras claves: Aprendizaje Personalizado, Analítica del Aprendizaje, Ambiente Virtual de Aprendizaje, Ingeniería Informática.

CONTEXTO

Nuestra línea de I/D está inserta en el proyecto *Mejora del aprendizaje personalizado aplicando analítica del aprendizaje* (ING656/22), el cual continua y profundiza las indagaciones ya iniciadas en los proyectos anteriores denominados: *Adaptación en un*

ambiente virtual de aprendizaje: pruebas y materiales personalizados (ING464/16) y *Personalización en un ambiente virtual de aprendizaje basada en estilos, conocimiento previo y errores frecuentes* (15H27/PSI/19). El primero contribuyó con el diseño y desarrollo de materiales de estudio y pruebas personalizadas para estudiantes universitarios, y en el segundo se emplearon los recursos adaptativos que provee el LMS Moodle, para diseñar itinerarios de aprendizaje, materiales de estudio y actividades personalizadas, que respondan a las características diagnosticadas de los estudiantes.

1. INTRODUCCIÓN

El aprendizaje personalizado explora itinerarios de aprendizaje a la luz de las diferencias personales de los estudiantes considerando su conocimiento y habilidades [1] [2]. Se entiende por itinerarios (o trayectorias) de aprendizaje a cualquier secuencia de contenidos pedagógicos contruidos con la finalidad de alcanzar los objetivos de aprendizaje de los estudiantes en un tema o área específica.

En este contexto, los parámetros de la personalización son las variables utilizadas para determinar las características y necesidades de los estudiantes. Los parámetros más usados son objetivos de aprendizaje, conocimiento previo, estilo de aprendizaje, habilidades adquiridas y el tiempo requerido para alcanzar cada objetivo.

Desde el inicio de la carrera Ingeniería Informática, se ha estudiado y desarrollado,

con un esfuerzo sostenido, la personalización de diseños instruccionales, y su adaptación progresiva en AVAs. Durante los primeros años, como complemento de las clases presenciales y de acceso opcional para los estudiantes, a partir de marzo de 2020, a raíz del aislamiento ocasionado por el SAR-COV2, se estableció la modalidad virtual con acceso obligatorio, en 2022 como complemento de las clases presenciales y con acceso obligatorio para todos los estudiantes. Fundamentos de la Informática, ha sido la materia sobre la cual se logró un mayor avance en la personalización y adaptación en el AVA (materiales de estudio, actividades y pruebas formativas), reportando los resultados obtenidos en congresos vinculados a la temática [3] [4]. La gran cantidad de datos existentes actualmente, generados por la actividad en los AVAs (p.e. resultados de encuestas sobre conocimiento previo y aspectos de la cursada, resultados de calificaciones formativas y finales, tiempo de realización de actividades y evaluaciones, estadísticas sobre la dificultad o no de los temas evaluados), particularmente a partir de 2020, demanda la aplicación de métodos y técnicas que brinden conocimiento automático para el diseño de nuevas recomendaciones, las cuales impactarían en la diversificación de los itinerarios de aprendizaje. En otras palabras, la aplicación de algoritmos de software para predecir y/o detectar aspectos desconocidos del proceso de aprendizaje, basados en datos históricos y el comportamiento actual de los estudiantes dentro del entorno en el cual sucede, posibilitarían el progreso de la personalización de los diseños instruccionales y su adaptación en AVAs.

La experiencia acumulada a lo largo de 8 años de investigación en el tema se ha trasladado a 3 materias específicas en la formación del Ingeniero Informático: Programación I. Taller de programación I y Programación III [6].

El aprendizaje personalizado [1] [5] [7] responde al modelo de enseñanza centrado en los estudiantes.

Los ambientes computacionales de aprendizaje tienen que contemplar la adaptación entendida como la adecuación de los contenidos o visualización del sistema a las preferencias del usuario. El proceso de adaptación está basado en las preferencias y objetivos del estudiante. Estas propiedades están almacenadas en el modelo del usuario. Estos modelos brindan la posibilidad de distinguir la intervención de los diferentes actores de la comunidad educativa, y proveer al sistema de la habilidad de adecuar su reacción con datos del modelo del usuario [9]. En el contexto del e-learning, los sistemas adaptativos se enfocan en la presentación de estos contenidos, ajustándose a diferentes condiciones. En los ambientes de aprendizaje se puede distinguir dos tipos de ajustes. En primer lugar, hay sistemas adaptables que pueden cambiar sus parámetros cuando el usuario modifica al sistema de acuerdo a sus necesidades. En segundo lugar, los sistemas adaptativos implican adecuar automáticamente el sistema al usuario. Las necesidades del usuario son inferidas por el sistema, a través de la observación de su comportamiento, cuando éste lo cambia de acuerdo a sus necesidades [8].

Si bien la personalización ha avanzado en estos años [1] [10] [11], el aprendizaje personalizado y su adaptación en AVAs todavía necesita progresar en diversos aspectos como recomendaciones individualizadas, filtrado de contenidos, fidelidad de los estilos de aprendizaje, evaluaciones personalizadas y trabajo colaborativo.

Una posible respuesta en pos de la mejora del aprendizaje personalizado, se basa en la aplicación de Analítica de Datos (AD) [12]. AD comprende una colección de datos valiosos, el análisis inteligente de la información y el hallazgo de conocimiento útil para la toma de decisiones. i.e., una gran cantidad de datos, técnicas estadísticas y modelos predictivos. La AD es un campo donde interactúan diferentes disciplinas, como estadística, reconocimiento de patrones, teoría

de sistemas, investigación operativa e Inteligencia Artificial.

En el campo de la educación, la Analítica de Datos dio lugar a la Analítica del Aprendizaje (AA). Desde que AA fue mencionada por primera vez en el reporte de New Media Consortium (NMC) Horizon Report 2012 [13] ha ganado una relevancia creciente. AA se define como la medición, colección, análisis y reporte de datos sobre estudiantes y sus contextos para entender y optimizar el aprendizaje y los ambientes donde ocurre.

Siguiendo a Campbell y Oblinger [14], AA consta de un proceso de 5 pasos: *Captura*, *Reporte*, *Predicción*, *Actuación* y *Refinamiento*.

AA posee cierta semejanza con EDM (Educational Data Mining). Ambas comunidades de investigación analizan información educacional, típicamente recopilada en plataformas en línea. La primera se centra en proveer herramientas para la intervención humana mientras que la segunda considera una adaptación automática.

El propósito de AA es crear un ambiente de aprendizaje amigable y personalizado para los estudiantes y facilitar recomendaciones a fin de alcanzar sus logros. Uno de los beneficios más importantes es el aporte sobre el desempeño de los actores de la comunidad educativa no solo en un momento específico, sino también sobre sus futuros logros o fallas p.e., provee a los estudiantes, docentes y autoridades información sobre su desempeño con relación a otros, identifica estudiantes que necesiten ayuda extra, ayuda a docentes con su planificación de sus cursos, entre otras. Estas metas necesitan un modelo de clasificación y predicción a fin de actuar con acciones remediabiles. Es aquí donde las técnicas de Machine Learning (ML) han mostrado su utilidad [15] [16].

ML es una convergencia de la estadística y las ciencias de la computación donde las máquinas

aprenden a mejorar el rendimiento desde sus experiencias previas, de una manera similar a los humanos. La única diferencia es que las computadoras aprenden desde los datos y las personas aprenden desde sus experiencias. [17].

Algunos autores han establecido que los métodos de aprendizaje de ML (p.e., ANNs (redes neuronales artificiales), SVM (support vector machines) y árboles de decisión pueden ser superiores a sus contrapartes estadísticas (p.e., regresión logística y análisis discriminante) en términos de ser menos restrictivos en sus asunciones y la producción de mejores resultados en sus predicciones. [18].

Las ANNs [19] [20] son herramientas usadas para análisis de datos cuando existen relaciones funcionales entre variables que luego son modeladas, su ventaja es que se pueden aplicar a estudios en los cuales los datos de entrada son incompletos o ambiguos por naturaleza. El estudio del desempeño académico ha ganado gran significancia no sólo para alcanzar un alto nivel, sino en la búsqueda de factores que influyen el proceso educativo [21] [22]. Por ejemplo, para detectar estudiantes en riesgo de abandono de sus estudios, o para reconocer estilos de aprendizaje de estudiantes individualmente de acuerdo a las acciones que ha ejecutado en un AVA [23].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- Personalización y adaptación del aprendizaje en AVAs como complemento de clases presenciales y propuestas de modalidad mixta (presencial y remota)
- Identificación de estilos de aprendizaje como aportes al diseño y desarrollo de propuestas formativas personalizadas.
- Analítica del Aprendizaje, exploración e implementación de modelos para la detección de estudiantes en riesgo de abandono, o para reconocer estilos de

aprendizaje de estudiantes individualmente de acuerdo a las acciones que ha ejecutado en un AVA.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Aplicar la experiencia acumulada por el grupo de investigación en personalización y adaptación en AVAs a 4 asignaturas de la carrera Ingeniería Informática.
- Ampliar la base de estilos de aprendizaje para evaluar la permanencia y/o cambio en los estilos y mejorar la adaptación de AVAs.
- Implementar Analítica del Aprendizaje en AVAs para mejorar la personalización y adaptación del aprendizaje.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está compuesto por cuatro docentes investigadores especialistas en educación e informática, y por dos estudiantes avanzados de la carrera Ingeniería Informática en etapa de formación como investigadores.

5. BIBLIOGRAFÍA

- [1] Nabizadeh A. H., Leal J.P., Rafsanjani H.N. y Shah R.R. Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Expert Systems with Applications*. vol. 159. 2020. 113596. ISSN: 0957-4174. <https://doi.org/10.1016/j.eswa.2020.113596>.
- [2] Zhang L., Basham J. D. y Yang S. Understanding the implementation of personalized learning: A research synthesis. *Educational Research Review*. vol. 31. 2020. 100339. ISSN: 1747-938X. <https://doi.org/10.1016/j.edurev.2020.100339>
- [3] González M. P., Benchoff D. E., Huapaya C. R., Remón, C., Lazurri, G., Guccione L., Lizarralde F. A. J. Avances en la

Personalización y Adaptación de Pruebas en un Ambiente Virtual de Aprendizaje. XXIV Congreso Argentino de Ciencias de la Computación. Octubre 2018. pp. 193-202. ISBN: 978-950-658472-6. <http://sedici.unlp.edu.ar/handle/10915/73088>

[4] Benchoff D. E., Lizarralde F. A. J., Huapaya C. R., Aguiar A. L., González M. P. Impacto del COVID19 en la enseñanza personalizada. XVI Congreso de Tecnología en Educación & Educación en Tecnología (TE&ET 2021): libro de actas, Argentina: Facultad de Informática (UNLP). 2021. pp. 61-69. ISBN 978-950-34-2014-0. <http://sedici.unlp.edu.ar/handle/10915/121569>

[5] Mansur A. B. F., Yusof N. and Basori A. H. Personalized Learning Model Behaviour based on Deep Learning Algorithm for Student Behaviour Analytic. *Procedia Computer Science* 16. 2019. pp. 125–133. ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2019.12.094>

[6] Benchoff, D., Lazzurri, C., Gellon, I., Borgnia Gianini, E., González, M., Huapaya, C., and Dai Pra, A. (2022). Personalización y adaptación del aprendizaje en Ingeniería Informática. *Memorias De Las JAIIO*, 8(8), 145-157. Recuperado a partir de <https://publicaciones.sadio.org.ar/index.php/JAIIO/article/view/414>

[7] Chrysafiadi K., Virvou M. *Advances in Personalized Web-Based Education*. e-Book. Springer Cham Heidelberg. 201. ISBN 978-3-319-12895-5

[8] Froschl C. *User Modeling and User Profiling in Adaptive E-learning Systems*. Master's Thesis de la Graz University of Technology. (2005)

[9] Brusilovsky P. y Maybury M.T. From adaptive hypermedia to the adaptive web. *Communications of the ACM*. 2002. e-Book, vol. 45, no. 5 pp. 30–33. DOI: 10.1145/506218.506239

- [10] Tian F., Zheng Q., Gong Z., Du J. and Li R. Personalized Learning Strategies in an intelligent e-Learning Environment. 2007 11th International Conference on Computer Supported Cooperative Work in Design, 2007, pp. 973-978, doi: 10.1109/CSCWD.2007.4281570
- [11] Benchoff D. E., González M. P. and Huapaya C. R. Personalization of Tests for Formative SelfAssessment. in IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, vol. 13, no. 2, pp. 70-74, May 2018, doi: 10.1109/RITA.2018.2831759
- [12] Wibawa B., Siregar J. S., Asrorie D. A. and Syakdiyah H. Learning Analytic and Educational Data Mining for Learning Science and Technology. The 2nd Science and Mathematics International Conference (SMIC 2020) AIP Conf. Proc. 2331, 060001-1–060001-7; <https://doi.org/10.1063/5.0041844> Published by AIP Publishing. 978-0-7354-4075-3.
- [13] Johnson L., Adams S., Cummins M., The NMC horizon report: 2012 higher education-edition. The New Media Consortium, Austin. 2012
- [14] Campbell J.P., DeBlois PB, Oblinger DG Academic analytics: a new tool for a new era. EDUCAUSE Review 42 Vol. 4. 2007. pp. 40–57. <https://er.educause.edu/articles/2007/7/academic-analytics-a-newtool-for-a-new-era>
- [15] Mitchell T. Machine Learning. New York, McGraw Hill. 1997. ISBN OCLC 36417892.
- [16] Mitchell T. Cap. 14 Key Ideas in Machine Learning. DRAFT OF December 4, 2017
- [17] Bishop, C. M., Pattern recognition and machine learning. Book. Springer. 2006. ISBN 978-0-387-310732
- [18] Delen D. A. Comparative analysis of machine learning techniques for student retention management. Decision Support Systems 49 (2010) 498–506.
- [19] Rumelhart, D.E.; James McClelland. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Cambridge: MIT Press. Book. 1987. <https://ieeexplore.ieee.org/servlet/opac?bknumber=627682>
- [20] Bouwmans, T., Javed, S., Sultana, M., and Jung, S. K. Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. Neural Networks, Vol. 117. 2019. pp. 866. ISSN 0893-6080. <https://doi.org/10.1016/j.neunet.2019.04.024>
- [21] Mason, C., Twomey, J., Wright, D., and Whitman, L. Predicting Engineering Student Attrition Risk Using a Probabilistic Neural Network and Comparing Results with a Backpropagation Neural Network and Logistic Regression. Res. High. Educ. Vol. 59, pp. 382–400. 2018. doi: 10.1007/s11162-017-9473-z
- [22] Figueiredo, J., Lopes, N., and García-Pébalvo, F. J. Predicting student failure in an introductory programming course with multiple back-propagation. ACM Int. Conf. Proc. Ser. 2019, pp. 44–49. 2019. doi: 10.1145/3362789.3362925
- [23] Villaverde J. E., D. Godoy and Amandi A. Learning styles' recognition in e-learning environments with feed-forward neural networks. June 2006. Journal of Computer Assisted Learning 22(3):197-206. doi: 10.1111/j.1365-2729.2006.00169.x

La incorporación de la Ciencia de Datos como estrategia transversal para refozar la enseñanza de la programación en el aula

Sofía Martín¹, Claudia Banchoff¹, Paula Venosa¹, Liliana Hurtado²

¹LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

Facultad de Informática. Universidad Nacional de La Plata

Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528

{smartin, cbanchoff, pvenosa}@info.unlp.edu.ar}

²Facultad de Ciencias Económicas, Jurídicas y Sociales. Universidad Nacional de Salta

lilianahur@hotmail.com

RESUMEN

La incorporación de contenidos de la Ciencia de la Computación en la educación de nivel obligatorio ha avanzado en los últimos años a través de diferentes experiencias y propuestas. Los conceptos más abordados de esta ciencia, se centran mayormente en el pensamiento computacional y la programación, a través de propuestas relacionadas a contenidos de la caja curricular o bien contenidos específicos del área de Informática.

La Ciencia de Datos, parte de la Ciencia de la Computación, ha incrementado su importancia en áreas tales como la industria, ciencia y trabajo, debido a que se utiliza como base al momento de analizar situaciones y definir medidas a tomar. Es por esto que la comprensión al momento de trabajar con un conjunto de datos cobra importancia para luego realizar un correcto análisis.

Luego de haber analizado algunas de las experiencias implementadas en distintos niveles de enseñanza secundaria de nuestra región se pudo relevar que la Ciencia de Datos aún no se ha incluido como contenido de las propuestas pedagógicas en el sistema educativo de nivel secundario argentino. Por lo tanto, esta línea de investigación se centra en el análisis de experiencias pedagógicas sobre la enseñanza de la Ciencia de Datos en ámbitos educativos de nivel secundario, así como las herramientas y metodologías utilizadas, con el objetivo de generar propuestas transversales de enseñanza de esta área en las escuelas secundarias de la Provincia de Buenos Aires. Se trabaja

especialmente con herramientas basadas en el lenguaje de programación Python.

Palabras clave: Ciencias de la Computación, Ciencia de Datos, nivel secundario, Python, Jupyter notebook

CONTEXTO

En el Laboratorio de Investigación en Nuevas Tecnologías Informáticas (LINTI) se viene trabajando en diferentes proyectos relacionados a la incorporación de la Ciencia de la Computación en la escuela desde el año 2006. Algunos de los proyectos de extensión como “Programando con Robots y Software Libre” [1][2] y los proyectos que involucran contenidos de la Ciencia de la Computación nos han permitido estar en continuo contacto con escuelas para el armado y realización de distintas propuestas pedagógicas.

Las propuestas llevadas a cabo son variadas en cuanto a contenidos abordados. Generalmente se realizan en función de la posibilidad de acceso a los recursos con que cuenta cada institución y la heterogeneidad de las formaciones docentes. En particular, en el nivel secundario, su adaptación e incorporación depende de la modalidad de orientación del establecimiento.

La línea de investigación en este artículo se focaliza en un contenido que aún no es ampliamente abordado en las escuelas secundarias, especialmente en nuestra región,

como la Ciencia de Datos, que cada día cobra más relevancia para su enseñanza.

Esta línea de investigación se encuentra enmarcada en el proyecto: “De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región”, del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI.

1. INTRODUCCIÓN

La incorporación de los contenidos de la Ciencia de la Computación en las instituciones educativas de nivel secundario ha sido muy diferente en cada una, debido a los recursos con que cuenta cada establecimiento. El avance de este proceso se ve muy relacionado con los planes gubernamentales, nacionales y provinciales, los cuales llevan a cabo la entrega de diferentes tipos de recursos tecnológicos y propuestas pedagógicas para su utilización.

En el nivel secundario, estos contenidos que han sido incorporados no incluyen conceptos relacionados a la Ciencia de Datos, a diferencia de otros niveles educativos como el universitario, donde se introducen cada vez más carreras, especializaciones y cursos de perfeccionamiento en esta área.

El crecimiento de los datos en los últimos años, debido al amplio uso de dispositivos digitales e Internet, ha puesto en evidencia la importancia del análisis y los diferentes usos que pueden realizarse de los mismos. Los objetivos de uso de los datos es muy amplio, dado que los datos surgidos de las interacciones individuales de los dispositivos digitales generan posibilidades de adaptación y orientación de contenido, mientras que en el ámbito de la industria, ciencia y educación los tipos de análisis se enfocan de diferente manera[3]. Un estudio realizado desde

el sitio Statista¹ estima que durante el año 2020 se han generado 64.2 zettabytes de datos y este número sigue en crecimiento en forma exponencial cada año, se estima un crecimiento hasta los 181 zettabytes para el año 2025.

Este incremento de datos potenció la búsqueda de trabajos orientados al análisis de los mismos, y pone en juego habilidades que se pueden trabajar desde los niveles educativos previos a la universidad.

Algunos ámbitos de uso que se pueden nombrar son el análisis de cambios climáticos a lo largo de los últimos años, datos de enfermedades contrarrestados con los síntomas, en las administraciones de los gobiernos para generar sus políticas públicas a implementar y como forma de transparencia al compartir datos no sensibles en el manejo presupuestario, entre otras. El acceso a los datos generados por las políticas gubernamentales también brindan la posibilidad de abordar temas relacionados a los Objetivos de Desarrollo Sostenibles (ODS) propuestos por la ONU[4]. Desde un punto de vista similar podemos encontrar organizaciones como el sitio World Bank², que publican datos para su análisis o bien como fuente para generar desarrollos informáticos privados que permiten brindar distintos servicios como ser aplicaciones de tránsito, de clima, entre otras. Podemos encontrar otros proyectos en que se utilizan y publican datos abiertos como es el caso de DataGénero³, que visualizan situaciones problemáticas sobre perspectiva de género y a través del análisis y una buena comunicación de los mismos permiten poner en discusión dichas situaciones.

Asimismo como los contenidos de la Ciencia de la Computación ponen en juego habilidades importantes como el pensamiento crítico y la resolución de problemas, el proceso que implica el análisis de datos, facilita trabajar estas

¹ Link del sitio Statista:

<https://www.statista.com/statistics/871513/worldwide-data-created/>. Último acceso 07-03-2022

² Sitio con datos abiertos :

<https://databank.worldbank.org/home>. Último acceso 22-09-2021.

³ Sitio de DataGénero <https://linktr.ee/datagenero>.

Último acceso 22-09-2021.

habilidades en la escuela secundaria, además de generar una motivación adicional por tratarse de problemas del mundo real que permiten tomar decisiones críticas y pensar en formas de agrupar y visualizar los datos[5].

La experiencia a lo largo de los años de trabajar en proyectos en conjunto con escuelas nos permitió visualizar el desafío que conlleva la incorporación de contenidos que no son propios de la caja curricular, para lo cual es importante tener en cuenta varios aspectos, tales como la adaptación del contenido al nivel educativo y la adecuación de la herramienta a trabajar. Una de las aplicaciones informáticas que permite trabajar con análisis de datos que abordaremos es Jupyter Notebook, la cual facilita el trabajo de forma simplificada e interactiva y es una herramienta ampliamente utilizada[6]. En una primera instancia de trabajo se propone el uso de herramientas de escritorio visuales más simples como PandasGUI⁴ y PandaStable⁵ que simplifican el manejo de las operaciones de análisis de datos, para luego incorporar en las actividades el lenguaje de programación Python a través del uso de Jupyter notebook y de librerías tales como pandas y matplotlib.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta línea de investigación se centra en los siguientes temas:

- Análisis sobre el impacto de la incorporación de conceptos de la Ciencia de Datos en la enseñanza transversal en la escuela.
- Relevamiento de distintas experiencias de enseñanza de la Ciencia de Datos en escuelas en otros países, incluyendo metodologías y herramientas de licencia libre utilizadas. En este sentido, se

analizarán las posibles adecuaciones de estas propuestas teniendo en cuenta las diferencias tanto de la caja curricular propia de cada jurisdicción como del nivel de contenidos específicos de Ciencia de la Computación ya abordados anteriormente.

- Diseño, elaboración y evaluación de materiales didácticos para la enseñanza de la Ciencia de Datos en el nivel secundario.
- Diseño, ejecución y evaluación de actividades específicas con estudiantes y docentes de nivel secundario.

Es importante destacar que este equipo de investigación viene interviniendo en distintas actividades en las escuelas de la región de La Plata, Berisso, Ensenada y Magdalena. También ha integrado el equipo que trabajó en el diseño curricular de la carrera “Especialización docente en didáctica de las Ciencias de la Computación” de la provincia de Buenos Aires y su implementación, y en la elaboración del “Primer manual de didáctica de las Ciencias de la Computación”⁶ ha generado experiencia en cuanto a poder intervenir en formación docente de nuestra provincia [8].

3. RESULTADOS Y OBJETIVOS

El objetivo principal de esta línea de trabajo es evaluar y adecuar herramientas de software que permitan elaborar propuestas de actividades que posibiliten introducir aspectos de la Ciencia de Datos, conforme a la caja curricular de las escuelas de nivel secundario pertenecientes a la Provincia de Buenos Aires, utilizando opciones de licencia libre.

Para alcanzar este objetivo, se proponen los siguientes objetivos específicos:

⁴ Disponible en :

<https://github.com/adameroose/pandasgui>

⁵ Documentación de PandaStable:

<https://pandastable.readthedocs.io/en/latest/description.html>

⁶ Este manual está disponible en

https://program.ar/descargas/cc_para_el_aula-2do_ciclo_secundaria.pdf . Último acceso: marzo de 2022

- Releva las diferentes experiencias pedagógicas desarrolladas en la región para la enseñanza de contenidos de la Ciencia de la Computación, área que integra la Ciencia de Datos, de manera tal de identificar los diferentes planes y programas que fueron acompañando la realización de las propuestas pedagógico-didácticas en la región.
- Releva las metodologías utilizadas y las herramientas propuestas para la enseñanza de contenidos específicos de Ciencias de la Computación y Ciencia de Datos, en las escuelas de nivel secundario.
- Investigar las diferentes opciones de las herramientas disponibles de licencia libre para el análisis y visualización de datos e identificar sus características y comparar uso, funcionalidad y posibles adecuaciones para escuelas de la región.
- Desarrollar propuestas pedagógicas para la enseñanza de la Ciencia de Datos en la escuela secundaria.
- Diseñar, implementar y evaluar las intervenciones con docentes y estudiantes de nivel secundario.
- Habilitar un espacio de aprendizaje formativo destinado a los docentes de los establecimientos del nivel secundario, con el fin de que utilicen y transfieran los resultados de la investigación en sus propuestas pedagógicas.
- Generar un sitio web público de referencia con ejemplos de análisis de datos adaptados a la escuela secundaria y brindar tutoriales de uso de las diferentes herramientas.

Durante el año 2022 se comenzó a trabajar con la escuela EET 9 de la ciudad de La Plata. Estas actividades permitieron definir una serie de acciones y ajustes a la propuesta inicial. Se analizaron herramientas visuales alternativas a Jupyter Notebook de manera de poder introducir los conceptos básicos de una manera más

simple para luego realizar actividades con el lenguaje de programación Python y librerías como pandas y matplotlib⁷.

La fig. 1 muestra capturas de PandaStable y PandasGUI. Estas herramientas fueron probadas en el sistema operativo Huayra Linux, distribución provista por las netbooks disponibles en la mayoría de los establecimientos educativos de la provincia de Buenos Aires.

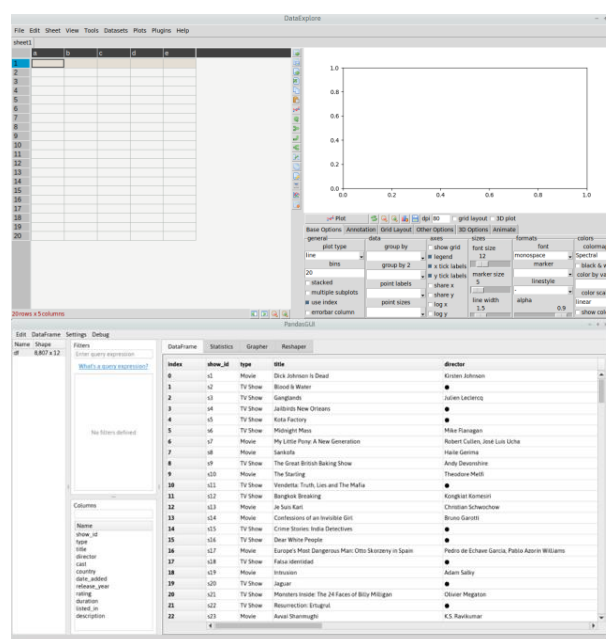


Figura 1: imagen de PandaStable y PandasGUI. Una primera propuesta se presentó como demo educativa en TEyET 2022[9] la cual tuvo una gran aceptación por parte de los docentes participantes.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I+D+i presentada en este artículo está formado por docentes investigadores categorizados del LINTI. Si bien el enfoque hacia la Ciencia de Datos es una línea que recién inicia, tiene como antecedentes años de investigación en trabajo

⁷ <https://matplotlib.org/stable/index.html>

con escuelas presentados en diversos ámbitos [10][11].

Esta línea de investigación dió lugar al desarrollo de una tesis de postgrado y actividades de cátedras relacionadas con esta temática, como así también el proyecto de extensión denominado “Ciencia de Datos en la escuela” que permitió obtener evidencias con base empírica para continuar indagando sobre las formas de incorporar esta disciplina en la educación media.

5. BIBLIOGRAFÍA

[1] Claudia Queiruga, Claudia Banchoff Tzancoff, Paula Venosa, Sofia Martin, Vanessa Aybar Rosales, Soledad Gomez, Isabel Kimura. *EscuelasTIC: estrategias para trabajar el pensamiento*. XIII Edición del Workshop de Investigadores en Ciencias de la Computación. Chilecito : UNDeC, 2021. ISBN 978-987-24611-3-3.

[2] Banchoff Tzancoff, C. M., Martin, E. S., Gómez, N. S., & López, F. E. M. (2019). Experiencias en robótica educativa: Diez años trabajando con escuelas argentinas. XIV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2019) (Universidad Nacional de San Luis, 1 y 2 de julio de 2019). ISBN 978-987-733-196-7, pág. 84-93. <http://sedici.unlp.edu.ar/handle/10915/90729>.

[3] Rao, A. R., Desai, Y., & Mishra, K. (2019). Data science education through education data: An end-to-end perspective. *2019 IEEE Integrated STEM Education Conference (ISEC)*, 300-307. <https://doi.org/10.1109/ISECon.2019.8881970>

[4] Departamento de Asuntos Económicos y Sociales (DESA). Naciones Unidas. Informe de los Objetivos de Desarrollo Sostenible 2020. e-ISBN: 978-92-1-004963-4 ISSN: 2521-6899 e-ISSN: 2521-6902. Disponible en <https://unstats.un.org/sdgs/report/2020/>

[5] De Veaux R, Agarwal M, Averett M, et al. Curriculum Guidelines for Undergraduate Programs in Data Science. *Annu Rev Stat Appl*. 2017;4(1):15-30. doi:10.1146/annurev-statistics-060116-053930

[6] Randles BM, Pasquetto IV, Golshan MS, Borgman CL. Using the Jupyter Notebook as a Tool for Open Science: An Empirical Study. In: *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*; 2017:1-2. doi:10.1109/JCDL.2017.7991618

[7] Díaz Barriga F. (2003). *Cognición situada y estrategias para el aprendizaje significativo*. Revista. Electrónica de Investigación Educativa, 5(2). Disponible en: <https://bit.ly/1JSJZpG>. Último acceso marzo de 2021.

[8] Martin, S., Banchoff, C., Venosa, P., & Hurtado, L. (2022). *Ciencias de datos en escuelas secundarias*. XVII Congreso de Tecnología en Educación & Educación en Tecnología - TE&ET 2022 (Entre Ríos, 15 y 16 de junio de 2022). <http://sedici.unlp.edu.ar/handle/10915/139959>

[9] Queiruga C., Banchoff Tzancoff C., Gómez S., Venosa P. (2020). *Enseñar a Enseñar Ciencias de la Computación. Una experiencia sobre políticas educativas y contenidos de Ciencias de la Computación*. En Actas II Jornadas Argentinas de Didáctica de la Programación / Alejandro Iglesias... [et al.] ; editado por Araceli Acosta... [et al.]. - 1a ed.- Córdoba: Universidad Nacional de Córdoba. Facultad de Filosofía y Humanidades, 2020. ISBN 978-950-33-1600-9

[10] Queiruga, C. A., Banchoff Tzancoff, C. M., Venosa, P., Martín, S. S., Aybar Rosales, V. del C., & Soledad Gomez, I. K. (2021). *EscuelasTIC: Estrategias para trabajar el pensamiento computacional en la escuela argentina*. XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja). ISBN 978-987-24611-3-3. <http://sedici.unlp.edu.ar/handle/10915/120472>

[11] Queiruga, C. A., Banchoff Tzancoff, C. M., Venosa, P., Martin, E. S., Aybar Rosales, V. del C., Gómez, N. S., & Kimura, I. *EscuelasTIC. El pensamiento computacional en la escuela*. XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan).

Realidad Virtual y Educación

Claudia Banchoff, Laura Fava, Facundo Diaz Gira, Sofía Martin

LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata
Calle 50 esq. 120, 2^{do} Piso. Tel: +54 221 4223528
{cbanchoff, lfava}@info.unlp.edu.ar, facundodiazgira@gmail.com,
smartin@info.unlp.edu.ar

RESUMEN

El desarrollo de aplicaciones educativas, enriquecidas con una amplia gama de técnicas ha sido un tema de investigación y desarrollo abordado en el LINTI, Laboratorio de Investigación en Nuevas Tecnologías Informáticas de la UNLP (Argentina), desde hace más de 15 años. En este artículo se presenta el estado de avance del desarrollo de aplicaciones que utilizan Realidad Virtual (RV) para potenciar aplicaciones tradicionales (web, de escritorio y móviles). Dentro de esta línea se abordan aspectos relacionados al desarrollo de aplicaciones educativas interactivas y materiales didácticos, así como también una nueva línea relacionada a la evaluación de usabilidad y formas de interacción en aplicaciones que incorporan RV.

En este artículo, se continúa con las líneas de I+D presentadas en WICC 2022[1], donde se analizaron y evaluaron entornos de desarrollo para aplicaciones de gamificación usando Realidad Aumentada (RA) y RV y se presentan los avances en los desarrollos de aplicaciones usando RA y RV.

Palabras clave: juegos serios, realidad virtual, gamificación, usabilidad en realidad virtual, enseñanza de programación, programación basada en bloques.

CONTEXTO

Esta línea de investigación incluye el desarrollo de aplicaciones interactivas vinculadas con la educación. Se focaliza en el uso de tecnologías de RV como elemento motivador para complementar las actividades que se trabajan en distintos contextos.

En el LINTI se viene trabajando, desde hace más de 15 años, en proyectos focalizados en educación y en la enseñanza de programación [2][3]. En los últimos años se finalizaron dos desarrollos que utilizan Realidad Aumentada los cuales dejaron una base de experiencia para abordar el desarrollo de otras aplicaciones. Actualmente se están finalizando desarrollos con Realidad Virtual que se describirán en detalle en las siguientes secciones. Esta línea de investigación se encuentra enmarcada en el proyecto: "De la Sociedad del Conocimiento a la Sociedad 5.0: un abordaje tecnológico y ético en nuestra región", del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI.

1. INTRODUCCIÓN

La realidad aumentada y la realidad virtual se han utilizado en múltiples áreas. El foco de este trabajo se centra en el ámbito educativo y, en este contexto, en nuestra región, no ha sido sencilla su incorporación.

En el caso de la RV, la adopción y aplicación es más compleja debido al requerimiento adicional de los dispositivos de hardware necesarios. Si bien existen experiencias con cascos de bajo

costo como los Google Cardboard, las mismas suelen estar asociadas a entornos de educación no formal [4].

Aún con las limitaciones de hardware que hacen menos populares este tipo de aplicaciones, se ha introducido un nuevo desafío: *¿qué aportes introduce la RV sobre las aplicaciones tradicionales?*

De la misma forma cuando se trasladan los test de usabilidad a un entorno de RV, para la mayoría de los usuarios se trata de un entorno nuevo, en el que se dejan los límites de una pantalla de 2 dimensiones para adentrarnos en un espacio de 3 dimensiones. Por este motivo, asoma un nuevo desafío: definir parámetros para evaluar nuevas formas de interacción que todavía no han sido definidas en ningún estándar de interacción.

Los ámbitos de implementación de aplicaciones en RV se han ido ampliando a lo largo de los años. Se pueden encontrar experiencias de aplicación en medicina, telepresencia, empatía para tomar conciencia sobre situaciones críticas o situaciones que suceden en comunidades lejanas [5][6]. Se utiliza RV en el desarrollo de simuladores de vuelo, manejo de automóviles y otros vehículos, cuidado de personas con problemas médicos, además de los ámbitos más conocidos como educación y videojuegos.

Si bien el hardware y software es importante, no deja de ser menos relevante el estudio de la percepción humana para generar entornos realmente inmersivos [7].

En este contexto, además de avanzar con el desarrollo de una aplicación interactivas con RV aplicada a diferentes áreas de arquitectura que incluyen la revisión de diseños y la simulación de operaciones dinámicas, se comenzó a desarrollar una aplicación, que utiliza tecnologías de RV para la enseñanza de programación mediante bloques. Esto incluye, además del análisis de herramientas tradicionales para programación basadas en bloques, que puedan ser reproducidas en entornos de RV, la definición de pruebas de usuario en entornos RV para lograr una comparativa con las herramientas tradicionales.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Como se ha mencionado, en el último año se avanzó con el desarrollo de una aplicación con RV para profesionales del área de arquitectura, a fin de crear entornos que faciliten el uso de modelos arquitectónicos para su visualización y manipulación. A partir de este trabajo surge un nuevo desarrollo con RV orientado a la enseñanza de programación con bloques. Si bien hay algunas iniciativas similares tales como VCoder[8] y Cubely[9], se diseñó una herramienta más simple que permitirá realizar un análisis comparativo con herramientas tradicionales, tales como Scratch¹ y Blockly². Este desarrollo también permitirá sumar una nueva herramienta que podrá utilizarse en otra área de investigación del LINTI en la cual se aborda la enseñanza de programación en las escuelas [10].

A continuación se detallan los ejes centrales de investigación para alcanzar los objetivos planteados:

- Relevamiento de herramientas para diseño 3D y renderización usadas por arquitectos.
- Evaluación de nuevos dispositivos disponibles dedicados para RV, analizando ventajas y desventajas de cada uno de ellos para determinar su adopción en el proyecto presentado.
- Análisis del estado de herramientas tradicionales basadas en bloques para la enseñanza de programación (Blockly, Scratch, etc.) que permitan el desarrollo de escenarios que puedan reproducirse o utilizarse en entorno de RV.
- Revisión de test de usuarios, extensión y adaptación para ambientes con RV.
- Elaboración de actividades que permitan generar una comparativa entre herramientas para programar con bloques con dispositivos de interacción tradicionales y la desarrollada con RV.

¹ Scratch: <https://scratch.mit.edu/>

² Blockly: <https://blockly.games>

3. RESULTADOS Y OBJETIVOS

El objetivo general de la línea de trabajo presentada es continuar con el uso de técnicas de RV en contextos educativos con especial énfasis en el aporte que la RV introduce en la enseñanza de programación. Para cumplir con el objetivo general, se proponen los siguientes objetivos específicos:

- Implementar prototipos que incluyan tecnologías de RV para introducir la enseñanza de programación a través de la manipulación de bloques 3D.
- Elaborar pautas de evaluación que permitan comprobar la adecuación y usabilidad de las aplicaciones con RV realizada
- Realizar una comparativa entre las herramientas para la enseñanza de programación basadas en bloques tradicionales y la que introduce RV.
- Evaluar el aporte de la tecnología de RV en la presentación de diseños arquitectónicos durante las etapas de revisión del mismo con el cliente de un proyecto de construcción. Continuar promoviendo esta temática en otros escenarios y dentro del marco del desarrollo de las tesinas de grado.

Como resultado se cuenta con dos prototipos de aplicaciones que utilizan RV.

Uno de ellos, denominado *CEIT (Creador de Entornos Inmersivos Transitables)*, que permite generar un entorno de RV a partir de un modelo 3D ingresado por el usuario, en el cual es posible también simular variaciones del ambiente relacionadas con cambios de horario, efectos climáticos y distintos tipos de iluminación. Esto posibilita al usuario final recorrer e interactuar con ese espacio a través de la inmersión provista por la realidad virtual.

La Figura 1 presenta una captura del tutorial interno provisto por CEIT donde se muestra un entorno que permite agregar y quitar objetos o experimentar con texturas, a fin de que el usuario se familiarice con los controles del sistema.



Figura 1. Captura de un entorno manipulable

CEIT se encuentra en su etapa final de desarrollo, donde será testeado por estudiantes de la Facultad de Arquitectura y Urbanismo.

El segundo prototipo con RV, denominado *BlocklyVR*, está destinado a la enseñanza de programación con bloques. Esta aplicación es un juego educativo de codificación, enseña a programar directamente dentro de la realidad virtual. El enfoque ofrece beneficios significativos sobre las plataformas tradicionales, ya que transforma el lenguaje de codificación en algo físicamente interactivo, en donde se espera que esto mejore la velocidad de aprendizaje.

BlocklyVR incluye un lenguaje de codificación interactivo basado en bloques estilo LEGO o RASTI, una ciudad que debe recorrerse a medida que se alcanzan los objetivos y controles 3D para las interacciones físicas. Los desafíos propuestos plantean recorridos en la ciudad como cruzar un puente, llegar a un punto determinado dentro de la ciudad, etc., entre otros.

La Figura 2 muestra el escenario propuesto para los desafíos del juego, desde la perspectiva del casco y controles de RV -captura de la ejecución del juego usando un Oculus, tomada desde la computadora.

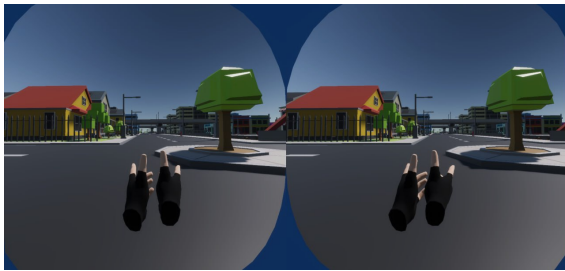


Figura 2. Escenario propuesto para BlocklyVR

El recorrido que debe hacer el jugador para avanzar en el juego cuenta con estaciones de interacción (Ver Figura 3), donde se puede manipular los bloques, primera versión de los modelos, para generar el programa que permite realizar los distintos recorridos.



Figura 3. Estación de interacción

Una vez finalizado el programa, la ejecución del mismo inicia una simulación inmersiva, en donde el jugador puede validar su programa. El desarrollo de BlocklyVR permitirá establecer algunos parámetros y guías para evaluar el impacto del uso de tecnologías de RV en distintos contextos educativos y comparar la experiencia con herramientas tradicionales. Algunos estudios de investigación están demostrando que la RV puede hacer que la educación sea más efectiva y tenga beneficios sobre el aprendizaje tradicional. *"En comparación con la educación tradicional, la educación basada en RV tiene una ventaja obvia en la enseñanza de conocimientos teóricos, así como en la formación de habilidades prácticas. En el entrenamiento de habilidades prácticas, ayuda a mejorar las habilidades operativas de los estudiantes, proporciona una experiencia de aprendizaje*

inmersiva y mejora el sentido de participación, haciendo que el aprendizaje sea más divertido, más seguro y más activo" [11].

Es de nuestro interés poder hacer pruebas de campo con la herramienta desarrollada y concluir sobre la eficiencia, la retención de conocimientos, la concentración y el grado de aceptación de BlocklyVR respecto de las herramientas basadas en bloques tradicionales.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I+D+i presentada en este artículo está formado por docentes investigadores categorizados del LINTI y estudiantes de la Facultad de Informática. Los trabajos mencionados corresponden a dos tesinas de grado en desarrollo y esta línea de trabajo ha dado lugar a otros proyectos y tesinas relacionadas a la temática planteada.

A través de la generación permanente de conocimiento por medio de líneas de investigación y desarrollo de aplicaciones vinculadas al sector productivo y su aplicación en el ámbito social, el LINTI promueve el uso innovador de las tecnologías informáticas en la región.

5. BIBLIOGRAFÍA

- [1] Banchoff C, Fava, L., Martin S., Díaz Gira, F. (2020). Explorando con Realidad Virtual Interactiva. XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022). Mendoza. Universidad Champagnat. ISBN: 978-987-48222-3-9.
- [2] Díaz J., Queiruga C., Banchoff. Tzancoff C., Fava L. y Harari V., "Educational robotics and videogames in the classroom," 2015 10th Iberian Conference on Information Systems and Technologies (CISTI), Aveiro, Portugal, 2015, pp. 1-6, doi: 10.1109/CISTI.2015.7170616.
- [3] Díaz J., Queiruga C., Banchoff. Tzancoff C., Fava L, *Juegos serios y aplicaciones interactivas*

usando realidad aumentada y realidad virtual, LACLO 2017:12th Latin-American Conference on Learning Technologies, La Plata, Argentina, 207.

[4] Chirinos Delfino, Y. (2020). La realidad virtual como mediadora de aprendizajes: desarrollo de una aplicación móvil de realidad virtual orientada a niños. Revista TE&ET no. 27. ISSN: 1850-9959. Páginas: 98-99.

[5] Craig, A, Sherman, B., William R., Jeffrey D. (2009). *Introduction to Virtual Reality, Developing Virtual Reality Applications*, Morgan Kaufmann, Boston, Pages 1-32, ISBN 978-0-12-374943-7, DOI: 0.1016/B978-0-12-374943-7.0000 1-X.

[6] Das, P.; Zhu, M.; McLaughlin, L.; Bilgrami, Z.; Milanaik, R.L. (2017). *Augmented reality video games: new possibilities and implications for children and adolescents*. Multimodal Technologies and Interaction, 2017; 1:8.

[7] M. Lavalley, M., *Virtual Reality*, 1st ed. Cambridge: Cambridge University Press, 2019. Disponible en: <http://vr.cs.uiuc.edu/vrbook.pdf>.

[8] Sitio oficial de vCoder, aplicación con realidad virtual para enseñar a programar, <https://www.vcoder.org/about/index.html>, Último acceso febrero de 2023.

[9] Cubely: Virtual Reality Block-based Programming Environment. 2018. Disponible en: <https://hoanghog.github.io/Cubely/dl/cubely.pdf>

[10] Queiruga C., Banchoff TzancoffC. , Venosa P., Martin S., Aybar Rosales V., Gomez S. y Kimura I.. *EscuelasTIC: estrategias para trabajar el pensamiento. XIII Edición del Workshop de Investigadores en Ciencias de la Computación*. Chilecito: UNdeC, 2021. ISBN 978-987-24611-3-3.

[11] BlueFocus, *A Case Study The Impact of VR on Academic Performance*, 2016, https://uploadvr.com/wp-content/uploads/2016/11/A-Case-Study-The-Impact-of-VR-on-Academic-Performance_20161125.pdf

DIAGNÓSTICO DE COMPETENCIAS GENÉRICAS UNIVERSITARIAS, A TRAVÉS DE SOFTWARE BUSINESS INTELLIGENCE

Silvia Poncio, Luciana Semino, Cintia Cuña, Nicolás Vozella, Ana Trottni,
Pablo Contesti, Yanina Montella

Universidad Abierta Interamericana. Facultad de Tecnología Informática.
Centro de Altos Estudios en Tecnología Informática. Rosario, Argentina.
{Silvia.Poncio, Luciana.Semino}@uai.edu.ar,
{Cintianatalia.Cuna,Nicolasalejandro.Vozella}@alumnos.uai.edu.ar,
anatottni@hotmail.com, pablo.contesti@outlook.com, yaninamontella@gmail.com

RESUMEN

La realidad laboral demanda profesionales capaces de gestionar situaciones que requieran de competencias específicas técnicas y genéricas que le permitan desarrollarse en el entorno donde interactúan, de forma satisfactoria, eficiente y constructiva, con proactividad, resiliencia y trabajo en equipo. En el área de Ingeniería en Sistemas representa para la sociedad un espacio de suma importancia en su desarrollo y en la mejora constante de los sistemas de información. Esto plantea un desafío para la educación superior porque es preciso garantizar un perfil de egresado del estudiante de ingeniería con las competencias específicas necesarias para el cumplimiento de las actividades reservadas de su área, pero también con competencias *genéricas sociales, políticas y actitudinales*, que vayan al encuentro de esa demanda que el entorno laboral requiere, aprendizaje basado en competencias y centrado en el alumno, como futuro profesional. Se propone un modelo diagnóstico de competencias universitarias apoyado en una herramienta de análisis de datos que brinda un espacio de utilidad para permitir observar cómo se encuentra un grupo de alumnos en relación con las competencias genéricas.

Palabras claves: diagnóstico, competencias genéricas sociales, políticas y actitudinales, capacidades, business intelligence.

CONTEXTO

El proyecto de investigación se desarrolla en la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI) sede Rosario, Santa Fe, República Argentina y se contextualiza dentro de la definición de las competencias genéricas definidas en la “Propuesta de Estándares de Segunda Generación para la Acreditación de Carreras de Ingeniería en la República Argentina” publicado por el Consejo Federal de Decanos de Ingeniería (CONFEDI) [1], máximo representante de la educación en Ingeniería a nivel nacional, una asociación civil sin fines de lucro que reúne a decanos y decanas de más de 120 facultades públicas y privadas de todo el país preocupados por la demanda social y la evolución de los estándares del aprendizaje/enseñanza de las ingenierías. El documento, también denominado como “Libro Rojo” [2], representa un compilado de competencias y habilidades de cada una de las ingenierías, que pretende garantizar la formación específica para las actividades reservadas de cada una de ellas. Es en dicha publicación que se definen, como condiciones curriculares comunes para las carreras, las competencias genéricas que debe tener el ingeniero en su perfil de egreso.

Cabe destacar que no sólo constituyen las competencias genéricas del argentino, sino que también sirvió como referencia utilizada para contribuir a la caracterización del Ingeniero Iberoamericano, así como soporte orientativo de la región. Ejemplo de este hecho es la

“Declaración de Valparaíso” firmada en noviembre de 2013, cuando la Asamblea General de la ASIBEI [3] adoptó como propia la síntesis de competencias genéricas de egreso acordadas por el CONFEDI, otro ejemplo similar es el Proyecto Visión 2025 de Formación Científica e Ingenierías de las Américas, en que la Organización de Estados Americanos (OEA) [4] adopta como síntesis las competencias del CONFEDI para la meta esperada.

1. INTRODUCCIÓN

El objetivo general de este proyecto consiste en la utilización de software de Business Intelligence (BI), para analizar las capacidades de las competencias blandas desarrolladas por el alumno.

Los objetivos específicos fueron: evaluar las capacidades alcanzadas; identificar habilidades que formen parte de éstas para desarrollar el proceso de formación y, a partir de ellas, relevar actitudes que evidencien el estado de situación del alumno.

BI es un término paraguas que abarca los procesos, las herramientas y las tecnologías para convertir datos en información, conocimiento y planes para conducir de forma eficaz las actividades de los negocios. El datawarehousing incluye dentro del negocio, consultas, informes, análisis, herramientas [5]. El ETL (Extract, Transform and Load), permite a los integrantes de una organización empresarial, la mejora en la dirección en forma efectiva, integrando y alineando la estrategia del negocio a las actividades diarias y a las mediciones de desempeño, dando como resultado, el análisis de la información gerencial para diagnosticar, controlar, planificar y dirigir el rumbo empresarial. “*En forma análoga*”, la propuesta en este proyecto es demostrar que es posible diagnosticar capacidades, habilidades y actitudes que forman parte de una competencia blanda universitaria, a través del análisis de BI.

La carrera Ingeniería en Sistemas se encuentra organizada en tres áreas de competencias, bajo

la estructura de Ejes socio-profesionales: “Los desarrollos Físico-tecnológicos de Sistemas Computacionales, la Administración de los Recursos Humanos e Informáticos como Proceso Estratégico y Los Modelos como Representación Abstracta de la Realidad. Se seleccionaron los objetivos planteados en este último, como variables intangibles. Partimos del concepto de las competencias universitarias expresadas en el Proyecto Tuning América Latina (2004-2007) [6]. Éste surge en un contexto de intensa reflexión sobre Educación Superior, tanto a nivel regional como internacional. Según la definición que da Tuning en el Informe Final de la Fase 2 del proyecto europeo entendemos que en inglés, “tune” significa sintonizar una frecuencia predeterminada en la radio; también se utiliza para describir la “afinación” de los distintos instrumentos de una orquesta, de modo que los músicos puedan interpretar los sonidos sin disonancias.

El Proyecto Tuning quiere reflejar la idea de búsqueda de puntos de acuerdo, de convergencia y entendimiento mutuo para facilitar la comprensión de estructuras educativas. Se enfoca como una metodología que procede de una perspectiva, cuya finalidad es incorporar los diferentes aspectos de la diversidad, de los países que en él intervienen e interactúan. Es una herramienta construida por las universidades para las universidades. El inicio del proyecto está dado por la búsqueda de puntos comunes de referencia, centrados en las competencias.

En el marco del proyecto, en el año 2015 analizamos y evaluamos las características del proyecto, principalmente sus atributos, a partir de la aplicación de un test de personalidad llamado Myers-Briggs [7] basado en la teoría de Carl Jung. Seleccionamos preguntas que nos permitieron aproximarnos a la realidad buscada y describir algunas de las competencias sociales de la Categoría 2 de “El Proyecto de Definición y Selección de Competencias (Deseco, 1997) [8], extraída del Resumen Ejecutivo del Programa para la Evaluación Internacional para Estudiantes (PISA) [9].

Posteriormente fue presentada: “La evaluación de competencias universitarias”, como tesis del Máster en Innovación y Emprendimiento cursado en la Universidad de Salamanca, España [10], que profundiza tres habilidades de la competencia “Interactuar en grupos heterogéneos”: manejar y resolver conflictos, cooperación y relación con otros.

En 2016, procesamos los resultados en el BI, arrojando un análisis comparativo entre dos de las categorías Deseco, *Interactuar en grupos heterogéneos* y *Actuar de manera autónoma*, para finalmente dar a conocer las conclusiones obtenidas en el marco de la Responsabilidad Social Universitaria [11].

Reconociendo que el profesional que toma decisiones en las organizaciones es nuestro egresado, es responsabilidad social ineludible de la universidad su preparación.

Una verdadera investigación científica debe expandir los límites del conocimiento y convocar nuevos desafíos. Si una investigación no abre interrogantes mayores, entonces solamente corrobora prejuicios que anquilosan el saber [12].

En este marco, los fines de la Universidad son la formación humana y profesional y la gestión de nuevos conocimientos, en base a ellos construimos los contenidos académicos que transferimos a los estudiantes para su formación.

Es por ello que en 2017 construimos un Modelo de Evaluación de Competencias Universitarias [13], a partir de éstas dos competencias claves, teniendo en cuenta las necesidades en el desarrollo de habilidades para cooperar o trabajar en equipo, gestionar y resolver conflictos, ser capaz de desenvolverse en sociedades cada vez más diversas y pluralistas, de empatizar y ponerse en el lugar de los demás, de manejar las propias emociones y de promover el capital social; así como también la autonomía para actuar dentro del contexto más amplio, formar y poner en práctica planes de vida y proyectos personales, defender y afirmar los propios derechos, intereses, límites y necesidades.

La propuesta consistió en diseñar un Modelo de Evaluación de Competencias Universitarias (MECU) [13], como indicador para evaluar las dos competencias sociales requeridas en los ámbitos sociales, tanto del trabajo como profesionales.

2. MODELO DIAGNÓSTICO COMPETENCIAS

Observamos que estas categorías de competencias sociales tomaban cada vez más importancia en el transcurso de los años, ya que son altamente requeridas en el perfil del alumno universitario, para el futuro laboral y su inserción social, por tal motivo avanzamos en una redefinición del modelo considerando que en realidad es posible diagnosticar ya que la categoría evaluación es un concepto altamente controvertido en el ámbito educativo y como nuestra intencionalidad trasciende las fronteras evaluativas, lo denominamos: “*Modelo Diagnóstico Competencias Universitarias*” que permitiera visibilizar el estado de grupos de alumnos para poder realizar ajustes a través de la formación continua.

Presentamos el Modelo Diagnóstico de Competencias Universitarias (MDCU) [14], resultante de un proceso de investigación iniciado en 2011, basado en BI, con el fin de obtener aproximaciones del desarrollo de las capacidades de las competencias genéricas del alumno universitario, de manera visual.

3. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente trabajo está identificado como proyecto de investigación y desarrollo en el Centro de Altos Estudios en Tecnología Informática (CAETI), dentro de la línea de investigación Sociedad del Conocimiento y Tecnologías aplicadas a la educación. Además, se ramifica en diferentes áreas temáticas específicas, a saber: Niveles de Business Intelligence, procesos de extracción, transformación y carga, datawarehouse, almacén multidimensional.

4. RESULTADOS OBTENIDOS y ESPERADOS

i. Resultados obtenidos:

El Management 3.0 [15] como modelo de gestión de los procesos colaborativos, entre alumnos universitarios, puede fortalecer el desarrollo de éstas. En relación con los trabajos anteriores realizados en esta misma Institución, permite avanzar en el diagnóstico de las competencias para poder plantear una acción concreta y así generar actividades que promuevan su desarrollo lo que será beneficioso para la trayectoria de la experiencia educativa universitaria. Utilizar prácticas de referencia como el Management 3.0 ayuda a elegir actividades para aplicar durante el proceso de enseñanza/aprendizaje en el marco de la carrera consiguiendo un plano pragmático pasando del saber, al saber hacer, permitiendo generar herramientas con las que los docentes (desde una visión más operativa en el campo) y la propia institución (desde una visión más estratégica), se pueden valer para conseguir los objetivos educativos de formación propuestos.

Según Adecco [16] una de las mayores compañías del mundo de RRHH, las necesidades de reclutamiento en la actualidad están cada vez más enfocadas en conseguir identificar y contratar talento con capacidades para el trabajo en equipo, el liderazgo, la comunicación efectiva, o sea con habilidades blandas, ya que si bien las duras son muy importantes, es más viable, hacer que una persona las pueda alcanzar con un costo bastante menor que el desarrollar una habilidad blanda, la misma (en caso de no poseerla o no estar suficientemente desarrollada) requiere un tiempo mayor, por tanto resulta más costoso.

Sumado a esto, en el contexto actual a nivel global y fruto de los constantes cambios, el elevado nivel de incertidumbre y las necesidades de adaptación a dichas condiciones, la persona representa un capital fundamental en la organización, siendo las habilidades blandas un aspecto muy importante y valioso que contribuye a la concretización de los objetivos de esta.

ii. Resultados esperados:

Este análisis de datos llevado a cabo en esta experiencia educativa realiza aportes para la formación y desarrollo internalización de las competencias genéricas, a través del modelo diagnóstico de competencias universitarias, la capacidad de tomar buenas decisiones para así obtener mejores resultados, que visibiliza el BI, nos desafió a la búsqueda de herramientas que permitan fortalecerlas. Se espera la implementación del diseño de la herramienta para diagnosticar el *aprender en forma continua y autónoma*, partiendo de capacidades específicas.

El uso de este modelo que integra la tecnología de BI podría ser contemplado como una posibilidad para el análisis de datos que sirva como soporte para la toma de decisiones en el ámbito de la educación universitaria innovadora ya que permite realizar un diagnóstico acerca de cómo se encuentra un grupo de alumnos respecto de una competencia genérica, convirtiéndose en un punto de partida para proseguir con futuras acciones.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado por los Ingenieros, Luciana Semino y Matías Banega, docentes UAI, Rosario; graduados Pablo Contesti, Verónica Castañeira, Lihuel Almada, Yanina Montella, Ingenieros en Sistemas Informáticos; Alejandro Vozella y Evangelina Marino, alumnos de la UAI, Rosario.

Además, se enriquece con la colaboración permanente del Dr. José Carlos Sánchez y la Dra. Brizeida Hernández, de la Universidad de Salamanca (España); con la Magister en Gestión de la Educación Superior Ana María Trotini, de la UNR, Rosario; Magister Cintia Cuña en Gestión de Recursos Empresariales, Universidad Politécnica de Catalunya, Barcelona (España).

Otros alumnos durante el trayecto de la investigación aportaron sus colaboraciones.

6. BIBLIOGRAFÍA

- [1] CONFEDI, Consejo General de Decanos de Ingeniería: <https://confedi.org.ar/quienes-somos/>
- [2] Libro Rojo de Confedi. Disponible en la web: <https://confedi.org.ar/publicaciones-confedi/libros-2/>
- [3] ASIBEI. Disponible en la web: <https://asibei.net>
- [4] OEA. <https://oas.org>
- [5] W. W. Eckerson y C. Howson. (2005, October 13). Enterprise Business Intelligence: Strategies and Technologies for Deploying BI on an Enterprise Scale. [Online]. Disponible en la web: <https://tdwi.org/articles/2005/10/13/enterprise-business-intelligence-strategies-and-technologies-for-deploying-bi-on-an-enterprise-scale.aspx>
- [6] Informe Final Proyecto Tuning América Latina (2007). Reflexiones y perspectivas de la Educación Superior en América Latina. [Online]. Disponible en la web: http://tuning.unideusto.org/tuningal/index.php?option=com_docman&task=view_category&catid=22&Itemid=191&order=dmdate_published&ascdesc=DESC
- [7] N. L. Quenk, Essentials of Myers-Briggs Type Indicator Assessment. 2nd edition. United states: John Wiley & Amp; Sons INC 2009
- [8] Resumen Ejecutivo de "Definición y selección de Competencias Clave" DeSeco OCDE. [Online]. Disponible en la web: <https://es.slideshare.net/primariaraceli85/competencias-clave-deseco>
- [9] PISA 2018 Resumen Ejecutivo. [Online]. Disponible en la web: <https://www.educacionyfp.gob.es/dam/jcr:03be5461-2c86-4663-a052-c224c3f8b617/pisa2018-oe-resumenejecutivo.pdf>
- [10] Poncio, S. (2016). Evaluación de competencias universitarias en estudiantes de la carrera Ingeniería en Sistemas, Universidad Abierta Interamericana, Facultad Tecnología Informática. Montevideo, Uruguay. Trabajo de Máster: Universidad de Salamanca, Salamanca, España.
- [11] Poncio, S. y Hernández, B. (2017). Competencias emprendedoras en la universidad entorno a la formación de responsabilidad social universitaria. V Congreso Internacional de Emprendimiento, AFIDE- 2017. Panamá. Panamá.
- [12] De Vincenzi, E. (2017). *Hacia una reingeniería de la vida*. Buenos Aires, Argentina: Editorial Universidad Abierta Interamericana.
- [13] Poncio, S. y Trottini A. M. (2017). Modelo de Evaluación de Competencias Universitarias (MECU). Sexto Congreso Argentino de la Interacción Persona Computador@, Telecomunicaciones, Informática e Información Científica. Huerta Grande, Córdoba. Argentina. [Online]. Disponible en la web: <http://www.alaipo.com/HCITISI-2017/IPCTIIC-HCITISI.html>
- [14] Poncio, S. y Trottini A. M. (2018). Modelo Diagnóstico de Competencias Universitarias (MDCU): Aprender a Emprender. VI Congreso Internacional de Emprendimiento, AFIDE- 2018. Boca del Río, Veracruz. México.
- [15] Management. Disponible en la web: https://en.wikipedia.org/wiki/Management#21st_century
- [16] Adecco Argentina: "Webinar sobre Habilidades Blandas". [Online]. Disponible en la web: <https://www.instagram.com/p/CdG4m4qJwjP/>

REVISIÓN DE ALTERNATIVAS MOODLE PARA LA EXTRACCIÓN DE INDICADORES DE AUTORREGULACIÓN DEL APRENDIZAJE

**RIVERO, Julieta; ISTVAN, Romina; LASAGNA, Valeria;
BACIGALUPE, María de los Ángeles**

*Universidad Tecnológica Nacional, Facultad Regional La Plata
Grupo de I&D Aplicado a Sistemas Informáticos, GIDAS UTN FRLP*

*Av. 60 s/n° esquina 124, CP 1900, La Plata, Buenos Aires, Argentina.
{ristvan; valerial; mabacigalupe}@frlp.utn.edu.ar; julietaa.rivero@gmail.com*

RESUMEN

La autorregulación del aprendizaje desarrolla la competencia de aprender a aprender, permitiendo a los estudiantes transformar sus aptitudes mentales en competencias académicas, convirtiéndose en un importante factor predictor del rendimiento académico y por consiguiente del logro estudiantil.

En concordancia con este marco, el objetivo final del proyecto es proponer un informe con indicadores de autorregulación del aprendizaje obtenidos de entornos virtuales de enseñanza-aprendizaje (EVEAS), particularmente del Campus Virtual Global (CVG) basado en Moodle que presenta UTN como entorno virtual oficial de enseñanza-aprendizaje.

Para lograr este objetivo se plantea primeramente la revisión de las diferentes alternativas que ofrece el entorno Moodle para obtener en forma automatizada los indicadores de la actividad estudiantil. Una vez elegida la opción más conveniente, confecciona un informe de seguimiento para cada alumno con

el fin de ser incorporado al actual Sistema de gestión tutorial de la UTN La Plata. Se espera que dicho informe pueda ser utilizado por docentes y tutores al momento de evaluar el perfil de riesgo de deserción de cada estudiante.

Palabras clave: Autorregulación del Aprendizaje, Entornos Virtuales de Enseñanza Aprendizaje, EVEAS, Indicadores de Autorregulación del Aprendizaje, Deserción Estudiantil.

CONTEXTO

Para afrontar el fenómeno de deserción estudiantil, la UTN La Plata en el año 2017 comenzó a trabajar en el diseño y desarrollo del Sistema de gestión tutorial ESDEU dentro del marco del Proyecto de Investigación y Desarrollo (PID): «Estudio Sistemático de Deserción Estudiantil Universitaria» homologado por Rectorado UTN con código TEINNLP0003786, el cual continúa su línea de trabajo a partir del año 2022 con un nuevo PID: «Sistema informático de predicción de

deserción estudiantil universitaria basado en un modelo de indicadores de regulación del aprendizaje, en entornos educativos mediados por TIC» homologado con código SIPPBLP008384.

El Sistema ESDEU es una herramienta informática que permite identificar el riesgo de deserción asociado a cada estudiante mediante alertas tempranas y gestiona de manera eficiente el seguimiento de los estudiantes facilitando el conjunto de tareas de gestión tutorial [1].

En respuesta a las nuevas modalidades educativas por la pandemia y teniendo en cuenta que el sistema comenzó su desarrollo cuando las clases presenciales eran parte de la única modalidad brindada por la institución; se plantea en este nuevo contexto la necesidad de incorporar indicadores obtenidos de entornos virtuales de enseñanza-aprendizaje (EVEAS), particularmente del Campus Virtual Global (CVG) basado en Moodle que presenta UTN como entorno virtual oficial de enseñanza-aprendizaje.

1. INTRODUCCIÓN

Moodle es una plataforma utilizada por gran cantidad de universidades y, particularmente, la Universidad Tecnológica Nacional lo ha adoptado como su Campus Virtual Global (CVG).

Dentro de las funciones que presenta el entorno, posibilita la obtención de indicadores de la actividad estudiantil, los cuales permiten identificar el comportamiento y el nivel de esfuerzo que llevan a cabo los estudiantes en el

proceso de aprendizaje.

Con el fin de integrar estos indicadores al sistema ESDEU se analizan en una primera etapa cuatro opciones para la obtención de los mismos, las cuales son provistas por Moodle:

1. Analytics de Moodle
2. Web Services de Moodle
3. Informes generales de seguimiento estudiantil
4. Base de Datos de Moodle

1. Analytics de Moodle

Moodle cuenta con varias funcionalidades que extienden el uso básico de creación y uso de cursos, recursos y actividades. Entre ellas es posible encontrar una aplicación denominada Analytics de Moodle, que permite obtener información del desarrollo de los cursos existentes y del comportamiento de los participantes con el rol de 'alumno' asignado.

La Analytics de Moodle permite a los administradores de los sitios de Moodle definir modelos de predicción, al combinar indicadores para predecir un objetivo particular.

El objetivo es el evento a predecir. Los indicadores son las condiciones que se cree que conducirán a una predicción precisa del objetivo.

Como ejemplo de un modelo de predicción se presenta la generación de alertas de estudiantes en riesgo de abandono, donde la falta de participación y las malas calificaciones en las actividades son consideradas indicadores. De esta manera, en este ejemplo, el objetivo consiste en determinar o predecir si el estudiante puede completar el curso o no.

Para ello Moodle realiza un entrenamiento para calcular estos indicadores y el objetivo para

cada estudiante en un curso ya finalizado. De esta manera predice qué estudiantes corren el riesgo de abandonar los cursos que se están desarrollando [2].

Básicamente Moodle provee tres modelos:

(i) El modelo “Estudiantes con riesgo de deserción” interpreta la definición “deserción” como “sin actividad del estudiante en el último cuarto del curso”. Es un modelo de machine learning que nos permite obtener predicciones utilizando los indicadores de compromiso del estudiante, profundidad cognitiva y amplitud social [3].

(ii) El modelo “Actividades próximas pendientes” indica si un usuario con rol estudiante completará o no las actividades nuevas que surjan en el curso utilizando la información obtenida en el indicador ‘Actividades Pendientes’, es decir que predice tomando el estado de las tareas pasadas del alumno.

(iii) El modelo “Sin enseñanza”, también denominado “Cursos en riesgo de no empezar” permite obtener una predicción sobre el estado del curso teniendo en cuenta dos indicadores: ‘No hay profesores’ y ‘Estudiantes inscriptos’. Además de estos tres modelos, es posible agregar otros por medio de plugins.

Debido a su potencial utilidad en lo que respecta a la detección a tiempo de la deserción estudiantil, se plantea analizar en profundidad el modelo ‘Estudiantes con riesgo de deserción’, llevando a cabo pruebas para determinar su funcionamiento y el formato de los datos resultado (insights).

2. Web Services de Moodle

Moodle ofrece un conjunto de funciones habilitadas particularmente para su uso en un sistema externo mediante sus Web Services. En la actualidad, las funciones básicas integradas que provee son más de 580, entre las que se incluyen muchas relacionadas con usuarios, cursos, chats, glosarios, foros, chats, wikis, y actividades en general, entre otras.

La investigación sobre los Web Services comienza analizando la documentación que Moodle ofrece, a la cual se puede acceder siendo usuario administrador. Utilizando dicho rol desde la interfaz de Moodle, se genera un token que permite acceder a los datos y realizar pruebas mediante la herramienta Postman.

Se plantea como trabajo configurar el servicio y explorar las funciones y los parámetros que necesitan las funciones para proveer los datos.

3. Informes generales de seguimiento estudiantil

Existen diferentes reportes generados por la plataforma que brindan información sobre la participación de los usuarios en los cursos y actividades, pensados en principio para el seguimiento de los cursos por parte del profesor a cargo.

Entre ellos se destacan:

(i) Los registros o bitácoras de *actividad* de curso: los cuales sólo se encuentran disponibles para los usuarios con rol administrador o profesor, y proveen en forma de tabla de doble entrada todas las acciones realizadas en un curso determinado. Éstos registros se pueden filtrar por participante, fecha, tipo de actividad y tipo de acción, por ejemplo, creado, visto o editado.

(ii) Los registros o bitácoras de *participación* en el curso: en éste informe se indica si hubo participación de un usuario con rol alumno en alguna de las actividades pudiendo seleccionar a todas o filtrar por alguna en particular. Si bien éste reporte no puede ser exportado o guardado, permite al usuario con rol profesor seleccionar, según su criterio, a los alumnos y mandarles un mensaje para realizar un seguimiento.

En éste trabajo se plantea focalizar en aquellos reportes que, además de facilitar su visualización desde la página, permitan exportarlos en diferentes formatos como ser .csv o .json.

4. Modelo de Base de Datos relacionado con el seguimiento del estudiante y el riesgo de deserción

Al momento de buscar alternativas para obtener los datos de la plataforma de Moodle, se considera la opción de la base de datos del Moodle de la UTN FRLP (CVG), realizando consultas directamente en la misma.

La exploración es la última opción que se presenta como alternativa a analizar.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto aborda el área de Tecnología Informática Aplicada en Educación y dentro de ella focaliza principalmente en: Educación a Distancia y Psicología Cognitiva.

Como líneas de investigación centrales aborda la Regulación del Aprendizaje, al profundizar en los constructos, dimensiones y métricas y sus indicadores y relaciones en los Entornos Virtuales de Enseñanza-Aprendizaje (EVEA).

De estos ejes centrales se realiza la revisión del estado del arte sobre regulación del aprendizaje y la revisión de las diferentes alternativas que ofrece el entorno Moodle para obtener en forma automatizada los indicadores de la actividad estudiantil.

En la siguiente sección se detallan los avances y resultados alcanzados.

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto aborda el análisis de las características y evaluación de las diferentes opciones que ofrece Moodle para la obtención de métricas e indicadores relacionados con la actividad estudiantil:

(i) la Analytics de Moodle,

(ii) los informes de los alumnos proporcionados en la GUI (Graphic User Interface) de la plataforma,

(iii) el Web Services de Moodle,

y (iv) el modelo de la base de datos de Moodle.

De este estudio se selecciona la herramienta que permita extraer de manera automática indicadores sobre los estudiantes y los cursos con un amplio grado de personalización.

Una vez elegida la herramienta se evalúan las métricas ofrecidas por ella y se selecciona un subconjunto de ellas para la confección de un informe de actividad estudiantil que será incorporado al actual sistema de gestión tutorial ESDEU.

Se espera que el informe propuesto pueda ser utilizado por docentes y tutores al momento de evaluar el perfil de riesgo de deserción de cada estudiante.

De esta manera, se espera como resultado final del proyecto incorporar al actual sistema ESDEU indicadores del desempeño académico obtenidos de entornos virtuales de enseñanza-aprendizaje (EVEAS), permitiendo así, que la herramienta integre información de los estudiantes que cursan tanto en la modalidad presencial, virtual o mixta.

4. FORMACIÓN DE RECURSOS HUMANOS

La UTN-FRLP se encuentra trabajando en Informática Aplicada a Educación desde el año 2017 a través de una de las líneas del Grupo de Investigación y Desarrollo de la especialidad Sistemas de Información, ‘GIDAS: Grupo de Investigación & Desarrollo Aplicado a Sistemas Informáticos y Computacionales’.

El equipo de trabajo actualmente presenta la siguiente estructura: director (1), co-director (1), docentes investigadores (2), tesistas de postgrado (1), tesistas de grado (2), becarios de investigación (2).

Cuenta con una tesis de Maestría aprobada y una de Especialización en curso en la carrera de Tecnología Informática Aplicada en Educación, y cuatro desarrollos de Prácticas Supervisadas (PS) de los estudiantes, necesarias para la obtención del título de grado de la carrera de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

[1] Istvan, R. M. (2023). Sistema informático para la detección temprana de deserción estudiantil universitaria. Estudio sobre ingresantes de la UTN Regional La Plata Deserción estudiantil en la UTN-FRLP

(Magister dissertation, Universidad Nacional de La Plata).

[2] Documentación Moodle. Guía rápida de analítica del aprendizaje. Recuperado el 10 de septiembre de 2022 de: https://docs.moodle.org/all/es/Guía_rápida_de_analítica_del_aprendizaje

[3] Documentación Moodle. Uso de Analítica. Recuperado el 10 de septiembre de 2022 de: https://docs.moodle.org/all/es/Uso_de_analítica

[4] Isohätälä, J., Järvenoja, H., & Järvelä, S. (2017). Socially shared regulation of learning and participation in social interaction in collaborative learning. *International Journal of Educational Research*, 81, 11-24.

[5] Wandler, J. B., & Imbriale, W. J. (2017). Promoting undergraduate student self-regulation in online learning environments. *Online Learning*, 21(2), n2.

[6] Morgan, T. (2020). Exploring socially shared regulated learning in PBL using group digital learning diaries: A study design. In *DS 104: Proceedings of the 22nd International Conference on Engineering and Product Design Education (E&PDE 2020)*, VIA Design, VIA University in Herning, Denmark. 10th-11th September 2020.

[7] Winne, P. H. (2017). Learning analytics for self-regulated learning. *Handbook of learning analytics*, 241-249.

[8] Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology*, 8, 422.

Realidad virtual para la enseñanza de portugués: Una guía para la implementación de contenido educativo en estudiantes de nivel medio

¹ Juan Ruíz Díaz, ² Lucas Kucuk, ³ Natalia Verónica Gallo

¹ Centro de Estudios Avanzados, Facultad de Ciencias Sociales, Universidad Nacional de Córdoba,

² Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.)

Departamento de Ingeniería y Ciencias de la Producción, Universidad Gastón Dachary,

³ UNER Concordia, Facultad de Administración

lukucuk@gmail.com, ruizdiazjuane@gmail.com, nataliagallo133@gmail.com

RESUMEN

Este artículo presenta una guía para la implementación de realidad virtual para la enseñanza del idioma portugués, utilizando VR Box, que son dispositivos para tecnología inmersiva. Al mismo tiempo, se realiza una propuesta de análisis de usabilidad del prototipo obtenido enfocado a una rúbrica denominada COdA (Calidad de Objetos de Aprendizajes). Se realizó un estudio de caso en el que estudiantes de quinto año de secundaria utilizaron los VR Box para explorar escenarios en 3D y realizar actividades interactivas relacionadas con la gramática y el vocabulario en portugués. Los resultados indicaron que los estudiantes se mostraron más motivados y comprometidos con el aprendizaje del idioma cuando utilizaron los dispositivos de realidad virtual, y también informaron sentirse más inmersos en el entorno de aprendizaje. Este artículo proporciona información valiosa para aquellos interesados en implementar tecnología de realidad virtual en la enseñanza de idiomas

Palabras clave: visores 3D, material inmersivo, realidad virtual, gafas de realidad virtual, tecnología educativa.

CONTEXTO

El trabajo se desarrolló en el marco de una investigación para la realización de tesis de grado de la Maestría en Procesos Educativos Mediados por Tecnologías en el Centro de Estudios Avanzados de la Universidad Nacional de Córdoba que investiga la generación y aplicación de materiales inmersivos y dan una aproximación al Metaverso, aplicando la experiencia piloto en la

escuela secundaria N°31 Benito Juárez de la ciudad de Concordia, Entre Ríos, en la materia portugués del 5to año, entre los meses de Octubre y Noviembre del año 2022.

I. INTRODUCCIÓN

La tecnología de realidad virtual ha evolucionado rápidamente, permitiendo experiencias inmersivas impensables antes. Los dispositivos de realidad virtual de bajo coste han hecho que esta tecnología sea más accesible. La tecnología sigue avanzando y se están explorando nuevas formas de utilizarla en diferentes campos, incluyendo la educación. [1] En este artículo, se presenta un estudio sobre el uso de dispositivos de realidad virtual en un aula de escuela secundaria, tiene como objetivo principal indagar si la tecnología de realidad virtual puede mejorar la experiencia educativa. Para ello se despliega un marco metodológico definido por etapas que contemplarán la adecuación y producción de un contenido específico para la materia portuguesa, reflexionando sobre lo experimentado por el docente y los estudiantes.

En esta línea se evalúan beneficios y desafíos al utilizar dispositivos de realidad virtual, presentando los resultados del estudio y sus implicaciones para la educación y la tecnología de realidad virtual en general.

II. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para entender el objeto del análisis de los dispositivos VR Box [2], qué función tienen en el aprendizaje y como los materiales educativos inmersivos deben interactuar, se tuvo presente lo dicho por Sabulsky [3] sobre el

modo en que los materiales tienen efectos significativos porque en su dimensión comunicativa organizan las acciones mediadas entre personas y grupos y son el principal portador de un mensaje, partiendo de la idea de que una información que se transforma en contenido en función de un tratamiento, sumado al recorte y la modificación que realiza el docente para su enseñanza, debe estar ensamblado con unas herramientas tecnológicas que permitan representar las decisiones e improntas innovadoras que se proponen al seleccionar el recorte a presentar.

Las gafas VR Box junto a los contenidos a través de videos e imágenes que se seleccionan en formato de 360°, también denominados panorámicos [4], son el soporte físico, no el hardware en sí, para lo cual se utilizan los smartphones de los estudiantes. Se decide centrar la experiencia en puntos concretos del mapa de Brasil, que volcado en imágenes panorámicas [5] permiten la inmersión en los escenarios que despliega el desarrollo de tópicos y conceptos propios de la materia portugués.

IV. DESARROLLO E INNOVACIÓN

Como se expone en otros trabajos presentados en el Workshop de Investigadores en Ciencias de la Computación (WICC) 2019 [6], los recursos desarrollados en realidad aumentada, y por extensión en la realidad virtual, deben basarse en planteos educativos precisos con el objetivo de ser efectivos, relacionando así las habilidades cognitivas del alumno. También se sugiere que la producción de los recursos educativos digitales sean eficientes para ser empleados en las actividades del docente, con posibilidad de reusabilidad, abiertos y de calidad [7].

Para mantener un criterio y poder ser objetivo a la hora de evaluar la experiencia se utilizará una herramienta en base a rúbrica denominada COdA (Calidad de Objetos de Aprendizajes)[8]. Esta herramienta fue diseñada para evaluar la calidad de un objeto de aprendizaje desarrollado con realidad aumentada utilizando un marco de trabajo.

Para el proceso de evaluación, se adaptará a la experiencia presentada, usando los criterios de calidad y su escala notarial de 1 a 5, implementando en el cuestionario las categorías: Objetivos y coherencia didáctica, Calidad de los contenidos, Capacidad de generar reflexión, crítica e innovación, Interactividad y adaptabilidad, Motivación, Formato y Diseño, Usabilidad, Accesibilidad y Reusabilidad.

La propuesta metodológica para llevar adelante la implementación de realidad virtual para la enseñanza del portugués se basa en tres etapas definidas como se puede observar en la Figura 1.

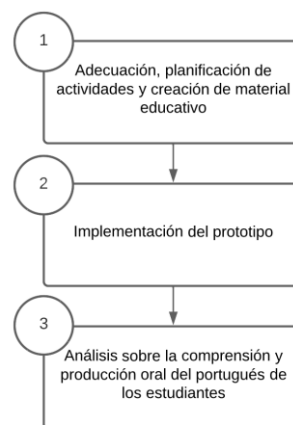


Figura 1. Etapas de la metodología.

La primera etapa consiste en la adecuación, planificación de actividades y creación de material educativo compatible con los dispositivos en cuanto a su forma de uso y requisitos. La segunda etapa contempla el diseño de la implementación del prototipo, esto incluye la utilización de las gafas de los VR Box, luego, en un tercer momento, se realiza un análisis sobre la comprensión y producción oral del portugués de los estudiantes. Al mismo tiempo se contempla la utilización de una rúbrica para que puedan evaluar aspectos de usabilidad del prototipo.

En cuanto a las actividades dentro del marco metodológico se establecen las siguientes:

Primera Etapa:

1) Introducción a los dispositivos VR Box: El profesor podría comenzar presentando los dispositivos VR Box a los estudiantes, explicando su funcionamiento y cómo se utilizarán en la clase de portugués.

2) Creación de material educativo compatible con los dispositivos VR Box: El profesor deberá diseñar material educativo, como videos, audios, textos y juegos interactivos, que sean compatibles con los dispositivos VR Box y que ayuden a mejorar la comprensión y producción oral del portugués.

3) Planificación de actividades de enseñanza: El profesor debe planificar actividades de enseñanza que involucren el uso de los dispositivos VR Box y que permitan a los estudiantes interactuar con el contenido y el entorno virtual.

Segunda Etapa:

1) Se diseña la web con los archivos necesarios para el funcionamiento del visor, las imágenes seleccionadas y las actividades propuestas.

2) Se instruye a tres estudiantes sobre el uso de los visores, los alumnos deben poseer smartphone compatibles.

3) Se accede a la web desde una televisor smart o un proyector conectado a un PC.

4) La página posee los detalles de las actividades que deben ir realizando.

5) La profesora examina, corrige y acompaña en la fonética y sintaxis de las actividades.

Tercera Etapa:

La etapa tres tiene por objetivo analizar cómo ha sido la experiencia del alumnado con el prototipo y verificar la calidad en términos de usabilidad del mismo.

Dicho esto, esta etapa comprende las siguientes actividades, de los cuales se obtiene un informe por cada una:

1) Evaluación cuantitativa: Mediante encuestas o cuestionarios para recopilar información sobre la satisfacción y efectividad de los estudiantes con respecto al material inmersivo. Se pregunta: "¿Cómo te sentiste durante la visualización de las imágenes y el video de 360 grados?" o "¿Crees que el material inmersivo te ayudó a comprender mejor el tema de la materia portugués?".

2) Evaluación cualitativa: A través de entrevistas individuales o discusiones en grupo

con los estudiantes, para obtener una comprensión más profunda de sus experiencias y opiniones sobre el material inmersivo. Para esta instancia tomamos en cuenta la rúbrica.

3) Evaluación del desempeño: Medir el rendimiento de los estudiantes en tareas relacionadas con el tema de portugués que se trató en la clase y compararlos con su rendimiento en clases anteriores sin el material inmersivo. Este punto estuvo a cargo de la profesora de la cátedra.

4) Evaluación del interés: Observar la participación y entusiasmo de los estudiantes durante las clases en la que se utilizó el material inmersivo, y compararlos con su participación y entusiasmo en clases anteriores.

La implementación de la rúbrica COdA se hará al finalizar las actividades listadas. En esta última instancia se proyecta que los alumnos y docentes usuarios utilicen la herramienta, en un formato de formularios virtuales, en donde por cada ítem a evaluar coloquen su apreciación de 1 a 5 siendo 1 malo hasta 5 que corresponde a muy bueno. Una vez obtenida las respuestas se realizará un análisis cuantitativo para determinar qué puntuaciones obtuvo el prototipo implementado, como se puede visualizar en [10].

V. TECNOLOGÍAS Y HERRAMIENTAS

Para implementar realidad virtual de inmersión total una opción es utilizar equipos basados principalmente en cascos que contienen un hardware avanzado y que las escuelas distan mucho de poder adquirir en lo inmediato, ejemplos equipos como los Oculus Rift o los HR Valve. Los VR Box utilizados necesitan como hardware principal a los smartphone de los estudiantes o docentes, estos dispositivos VR son similares a los Google Cardboard, con la ventaja de contar con un control o mando remoto para el movimiento, que permite el desplazamiento e interacción sobre los comandos o acciones que se muestran en los materiales.

El uso de las gafas VR Box es relativamente sencillo. El estudiante coloca las gafas en su cabeza, ajusta las correas o elásticos para que queden firmes y enfoca las lentes de acuerdo

con sus necesidades, pudiendo utilizar lentes o anteojos si es el caso (ver Figura 2). Algunos modelos pueden requerir que se ajusten las lentes manualmente, mientras que otros ajustan automáticamente la distancia interpupilar. Una vez que se han ajustado las gafas, el usuario puede acceder a los contenidos de realidad virtual a través de una aplicación o plataforma compatible, accesible desde la tienda del sistema operativo del smartphone que se utilice.



Figura 2. Gafas VR Box

Para la experiencia se diseñó una web [9] que contiene un mapa de Brasilia, interactivo mediante códigos QR y que cuenta con los software necesarios para el funcionamiento de la inmersión en las panorámicas seleccionadas por la cátedra. Esta experiencia contó con el apoyo de app gratuitas como VR Media Player [10] para poder soportar e interactuar con los materiales inmersivos y con el software Street View Download 360 [11] que permite la descarga y posterior manipulación de las panorámicas extraídas desde Google Street View [12].

VI. RESULTADOS OBTENIDOS

A partir del marco de trabajo planteado y la tecnología vinculada, surge la implementación preliminar en un caso de estudio como primera experiencia de trabajo.

En el año 2022 se pudo aproximar a la utilización del marco de trabajo propuesto con 15 estudiantes a cargo de 1 docente. La experiencia comprendió la aplicación de la guía, en donde se adaptó un contenido didáctico para la enseñanza del portugués. Dicho contenido se basó en videos 360°, embebidos en

una página web, relacionado a la enseñanza del idioma.

Siguiendo las actividades de la segunda etapa, los alumnos utilizaron los VR Box utilizando sus smartphones vinculados a una televisión. Los mismos desarrollaron las actividades enmarcadas en la página, que eran básicamente los mostrados en la figura 3.

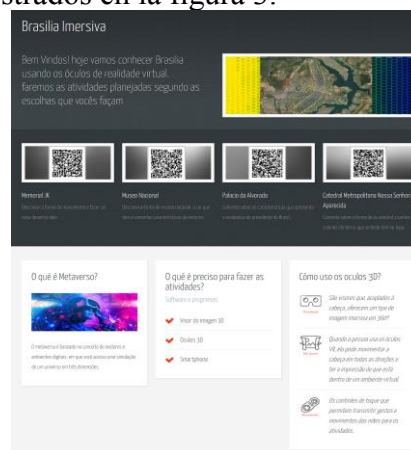


Figura 3. Web con actividades y material educativo.

Para la etapa 3 se generaron las preguntas de los formularios que consultan sobre el uso del dispositivo y los materiales en cuanto apreciaciones sobre las formas de interacción, individualmente se pregunta por los contenidos, posteriormente se mide el desempeño y finalmente se realizan observaciones para dar una conclusión a la experiencia.

Implementada la tercera etapa, se verificó que hubo una notable mejora en la comprensión oral y escrita de los estudiantes ante las actividades propuestas y la factibilidad del uso de los visores VR Box en el aula. No obstante aún no se ha implementado la rúbrica para evaluación de usabilidad, en posteriores presentaciones se hará la implementación de la rúbrica COdA.

Al concluir la experiencia los estudiantes y el profesor informaron que el uso de los dispositivos pueden ser una herramienta eficaz y motivadora, como vemos en la fig.4, para mejorar la comprensión y producción de las actividades propuestas.



Figura 4. Experiencia en aula con VRBox

Se concluye que se presenta una guía viable para la utilización de Realidad Virtual en la enseñanza del portugués, sumado a la evaluación de la calidad con respecto a los materiales diseñados.

VI. LÍNEAS DE DESARROLLO FUTURO

Las líneas futuras de desarrollo se centran en la flexibilización de la implementación de la rúbrica a través de entornos virtuales, y también en la optimización de recursos para la construcción de una mayor cantidad gafas.

VII. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo participan un tesista de la Maestría en Procesos Educativos Mediados por Tecnologías de Universidad Nacional de Córdoba, un Magíster en Tecnologías de la información del Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.) de la Universidad Gastón Dachary con antecedentes de investigación en los temas que se abordan y una Profesora Superior en Portugués de la Universidad Nacional de Entre Ríos.

VIII. REFERENCIAS

[1] Cobo, C. (2016) La innovación pendiente. Reflexiones (y Provocaciones) sobre educación, tecnología y conocimiento. Colección Fundación Ceibal/Debate: Montevideo.

[2] Abásolo M.J., De Giusti A., Naiouf M., Pesado P., Sanz C., Barbieri S., Boza R., Gavilanes W., Mitaritonna A., Prinsich N., Vincenzi, M., Montero F., Perales López F. Aplicaciones de realidad virtual, realidad aumentada e interfaces multimodales. XXI Workshop de Investigadores en Ciencias de la

Computación WICC 2019, Universidad Nacional de San Juan, RedUNCI, ISBN 978-987-3619-27

[3] Sabulsky, G. (2009). Materiales educativos que recuperen el hacer y el pensar del profesor, en Comunicación y educación en entornos virtuales de aprendizaje. Perspectivas teórico-metodológicas. Sara Pérez y Adriana Imperatore son compiladoras. Universidad Nacional de Quilmes. (Pág. 344)

[4] Perez Morente, Luis Miguel, José Jesús Guerrero Campo, and Jesús Bermúdez Cameo. "Reconocimiento y localización de objetos en imágenes de 360 grados."

[5] Brasília, Brasil. Google Maps, 2022

[6] XXI Workshop de Investigadores en Ciencias de la Computación -WICC 2019: libro de actas compilado por Nelson Rodríguez, María Murazzo, Manuel Ortega, María I. Lund. - 1a ed. - San Juan: Editorial UNSJ, 2019.

[7] Schwartzman, G. y Odetti, V. (2013) "Remix como estrategia para el diseño de Materiales Didácticos Hipermediales".

[8] Kucuk, Lucas, and Jorge Salvador Ierache. "Aplicación de rúbrica COdA para evaluación de calidad objetos de aprendizajes basados en realidad aumentada." XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz). 2020.

[9] "materialesinmersivos/settings/pages "

.GitHub. (accedido 26/2/2023)

<https://jruizdiaz.github.io/materialesinmersivos/>

[10] VR Media Player (2022) [Aplicación móvil]. PlayStore. https://play.google.com/store/apps/details?id=com.xojot.vrplayer&hl=es_AR&gl=US

[11] Street View Download 360. [software]. <https://svd360.istreetview.com/>

[12] Google Street View. Recuperado de: <https://www.google.com/intl/es/streetview/>

Definición de Lenguajes Específicos del Dominio para la Generación de Sitios Web de Instituciones Escolares

Analia Magdalena Zaldua, Mario Berón, Germán Montejano, Daniel Riesco
Área de Programación y Metodologías de Desarrollo de Software
Universidad Nacional de San Luis
Facultad de Ciencias Físico Matemáticas y Naturales
anazaldua@gmail.com, {mberon,gmonte,diresco}@unsl.edu.ar

Resumen

Antes de la revolución tecnológica actual, los avisos institucionales se debían realizar con la suficiente antelación para que toda la comunidad educativa estuviera debidamente informada, aunque no se disponía del medio ni tiempo suficiente para hacerlo. En la actualidad esto cambió debido al avance de las Tecnologías de la Información y de la Comunicación. Se puede observar a nivel global que cada institución tiene presencia en internet a través de un sitio web que la representa y donde realiza la comunicación de los eventos más importantes. La realidad mencionada no sucede con todas las instituciones educativas dado que no cuentan con los recursos necesarios para construir su propio sitio web. En este artículo se presenta una línea de investigación que se enfoca en el diseño y construcción de lenguajes específicos del dominio y herramientas asociadas, que permitan especificar espacios web institucionales los cuales posibiliten la generación automática del entorno web institucional.

Palabras clave

LED – sitio web – generación automática - herramienta

Contexto

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de Software (LaCIS) de la

Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto: “Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube”, perteneciente a la misma institución. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional. Además, forma parte de las actividades de extensión-investigación del Proyecto de Extensión de Interés Social *Resolución Creativa de Problemas: Métodos, Técnicas y Herramientas para el Desarrollo de Soluciones Creativas y Tecnológicas de Problemas del Contexto Académico y Laboral*.

1. Introducción

Las instituciones educativas regionales de gestión pública brindan un servicio muy loable a la sociedad como lo es la educación gratuita y de calidad. Para poder llevar adelante dicha tarea, todo el personal realiza mucho esfuerzo para cumplir con los objetivos que propenden a que los estudiantes aprendan significativamente. Entre estas tareas se encuentra la de comunicar información relevante que la institución necesita notificar a su comunidad educativa para estar al tanto de los hechos que acontecen en la misma. Tradicionalmente, este tipo de actividades se lleva a cabo mediante notas en el cuaderno de comunicaciones que los estudiantes hacen firmar por sus tutores, de forma tal, de que conozcan las actividades institucionales. Si bien esta forma de llevar adelante la tarea ha funcionado, muchas noticias e información no

llegan a conocerse hasta tanto el interesado se acerque a la institución o bien lea documentos asociados a la misma. A modo de ejemplo se pueden mencionar las noticias de último momento, las orientaciones disponibles en la institución, las fechas de inscripciones que se conocen cerca del inicio de clases, entre otras tantas posibilidades.

Con el propósito de resolver este inconveniente, las instituciones han evolucionado incorporando tecnologías de punta, las cuales han resuelto en gran medida los problemas antes mencionados. En la actualidad muchas instituciones tienen su propio sitio web en donde, cada tanto, van actualizando las noticias que el equipo de gestión desea que la comunidad educativa conozca [16,17]. Esta aproximación ha resultado un adecuado método de comunicación, sin embargo, aún presenta un inconveniente de difícil solución: las instituciones que poseen dicho avance tecnológico son, en general, instituciones privadas. Las mismas tienen sustento económico como para pagar a un profesional para que desarrolle el sitio y, si es necesario, lo actualice. Las instituciones públicas adolecen de esta peculiaridad con lo cual el diseño de un sitio web es una tarea que puede realizar, de buena voluntad, un docente de la misma, hecho poco probable dado que los mismos están dedicados tiempo completo a la tarea de educar.

Si, por el contrario, un integrante de la comunidad educativa desarrolla el sitio web, la tarea de actualización estaría a cargo de la misma persona dado que, por lo general, será un profesional o idóneo de la informática.

Claramente, ninguno de los enfoques mencionados en los párrafos precedentes es viable en las instituciones educativas públicas porque no se disponen de costos ni de especialistas en informática. Es por este motivo que se consideró importante llevar adelante investigaciones que tengan como finalidad proporcionar herramientas que faciliten la construcción de espacios web institucionales particularmente aquellas que se

basan en lenguajes específicos del dominio dado que en este contexto las soluciones pueden ser planteadas usando el vocabulario del dominio del problema que el usuario conoce cabalmente [5, 6, 8, 12].

2. Líneas de investigación y desarrollo

Un lenguaje específico del dominio es un lenguaje de programación/especificación que posee un nivel superior de abstracción el cual está orientado a resolver problemas de cierto tipo.

Los lenguajes específicos del dominio han sido ampliamente utilizados en diferentes contextos, por ejemplo, para la especificación de circuitos digitales como es el caso de VHDL, Verilog, en donde su manejo simplifica notablemente la implementación de circuitos complejos [4]. También es posible encontrar lenguajes específicos del dominio para modelación y especificación por ejemplo R es un lenguaje específico del dominio que permite realizar estadísticas fácilmente, WSDL (Web Service Description Language) es un dialecto de XML que tiene como finalidad especificar las entradas, salidas y operaciones que tiene un servicio web [1]. Además, se han usado en el contexto de las bases de datos como por ejemplo SQL un lenguaje que facilita la realización de consultas [7]. Cabe destacar que los lenguajes específicos del dominio pueden ser simples o complejos y su implementación puede implicar seguir todos los pasos para la construcción de un compilador, parte de él, o llevar adelante esta tarea mediante el desarrollo de rutinas específicas de software [14,15].

Sea cual fuere el mecanismo de construcción de un lenguaje específico del dominio, la característica que sobresale, en este contexto es: “Las/os especificaciones/programas escritas/os en un lenguaje específico del dominio pueden ser hechas/os mediante la utilización del vocabulario de los usuarios y del dominio de la solución” [2, 9]. Esta peculiaridad hace pensar que el desarrollo de

un lenguaje simple, que utilice la terminología del dominio del problema, que posea construcciones fáciles de entender, sea pequeño de forma tal que la curva de aprendizaje sea breve y permita generar sitios web institucionales sencillos, proporcione una solución adecuada al problema que se describió previamente [18,19].

Existen numerosas investigaciones referidas al uso, desarrollo e implementación de lenguajes específicos del dominio. Varias investigaciones están orientadas a los lenguajes de programación más antiguos (Cobol, Fortran, Lisp) que surgieron como lenguajes específicos para resolver problemas en un área determinada (procesamiento comercial, computación numérica y procesamiento simbólico, respectivamente) [11]. Poco a poco, ha surgido la necesidad de un soporte de lenguajes más especializados para resolver problemas en dominios de aplicación bien definidos.

Se han podido observar varios intentos de construcción de sitios web, no obstante, en todos los casos, se implementan herramientas informáticas que solo pueden ser utilizadas por el programador y que el usuario, miembro de la institución, difícilmente puede modificar.

Para elaborar lenguajes específicos del dominio con las características antes mencionadas se debe seguir un proceso que consta de tres pasos: i) análisis del dominio, ii) modelación e iii) implementación. La actividad realizada en el primer paso se lleva adelante a través de entrevistas con el personal de diferentes instituciones educativas de la región [3]. A partir de ella es posible identificar los conceptos principales que un lenguaje específico del dominio debe contener y de las relaciones existentes entre dichos conceptos, a modo de ejemplo se pueden mencionar los conceptos de: director, docente, materia noticia, etc. En lo que respecta a la segunda etapa, se puede decir que a partir de los conceptos del dominio del problema se estudian diferentes formas de volcarlos en una gramática la cual pueda ser procesada por herramientas especializadas que permitan

generar el código correspondiente. Para llevar adelante el tercer paso, en primer lugar, se debe focalizar el estudio en implementación de lenguajes textuales, con lo cual para realizar esta tarea se deben construir analizadores lexicográficos, sintácticos y semánticos lo que permitirá generar el código ejecutable que producirá el resultado deseado. Para llevar adelante las tareas antes mencionadas, es necesario investigar cuáles de las herramientas de soporte al proceso de compilación se adapta más a las necesidades del proyecto. En este sentido, se analizaron: lex, yacc, javacc, bison, antlr4, entre otras tantas herramientas y se llegó a la conclusión que antlr4 es la que más se adapta a las tareas que se deben realizar en la investigación porque implementa fielmente el concepto de gramática de atributos y porque también posibilita la utilización de tecnologías modernas de compilación como lo son los listeners y visitors los cuales permiten extraer información y generar código con o sin construcción explícita del árbol de parse. Una vez realizado el análisis del dominio, la modelación correspondiente y la selección de las herramientas para la implementación se debe proceder a desarrollar el lenguaje. En este sentido se puede decir que se implementó, con antlr4, un prototipo de lenguaje específico del dominio denominado LAZ (Lenguaje de Analía Zaldúa), cuya gramática se puede ver en la figura 1. Hasta el momento LAZ se ha usado en ejemplos de laboratorio, para verificar que el funcionamiento sea correcto y no presente problemas técnicos a la hora de usar el prototipo en situaciones reales una vez que se superen las pruebas de laboratorio se intentará aplicar la solución con diferentes instituciones educativas.

Es importante destacar que, un lenguaje de estas características propenderá a la independencia de las instituciones para llevar adelante los cambios en su sitio y proporcionará agilidad en las tareas (las novedades se pueden comunicar con mayor facilidad) al usar terminología propia y construcciones simples. En otras palabras, un enfoque como el descrito puede proporcionar

una solución interesante a un problema de la vida real que incumbe a muchas instituciones del contexto regional y nacional y cuyo desarrollo implica la aplicación de diferentes temáticas como, por ejemplo: Modelación, Programación, Desarrollo de Lenguajes, entre otras tantas posibilidades.

```

institución→ <COMIENZO-INSTITUCIÓN>
               quienesSomos informacionInsitucional
               personal
               <FIN-INSTITUCIÓN>
informaciónInstitucional→ STRING
personal→ directivos docentes administrativos
directivos→ <COMIENZO-DIRECTIVOS> director+ <FIN-DIRECTIVOS>
director→ <COMIENZO-DIRECTOR>
               nombre cargo contacto
               <FIN-DIRECTOR>
docentes→ <COMIENZO-DOCENTES> docente+ <FIN-DOCENTES>
docente→ <COMIENZO-DOCENTE>
               nombre materia contacto
               <FIN-DOCENTE>
administrativos→ <COMIENZO-ADMINISTRATIVOS>
               administrativo+
               <COMIENZO-ADMINISTRATIVOS>
administrativo→ <COMIENZO-ADMINISTRATIVO>
               nombre contacto
               <FIN-ADMINISTRATIVO>
quienesSomos→ STRING
nombre→ STRING
materia→ STRING
contacto→ STRING

```

Figura 1: Lenguaje Específico del Dominio LAZ

3. Resultados Obtenidos/esperados

Como ha sido mencionado con anterioridad hasta el momento se ha logrado:

- Definir un lenguaje específico del dominio de experimentación. Dicho lenguaje fue elaborado a partir del estudio de campo realizado por integrantes del equipo de investigación.
- Implementar el lenguaje utilizando tecnologías y herramientas de compilación que soportan los principales conceptos de construcción

de lenguajes. Esta particularidad ha facilitado la implementación de partes de un compilador que por lo general son complejas de construir.

- Se ha utilizado el prototipo en ejemplos de laboratorio con el propósito de eliminar detalles técnicos que pueden dificultar su uso en una situación real.

Como resultados esperados se pueden mencionar:

- Evolucionar el prototipo para que se transforme en una herramienta que pueda ser usada en el día a día en las instituciones. Esta tarea implica mejorar el lenguaje específico del dominio a partir de las opiniones de los usuarios e incorporar esos detalles en el sistema que implementa el lenguaje.
- Construir herramientas de soporte al lenguaje específico del dominio como por ejemplo un editor que ayude al usuario a escribir un programa en el lenguaje.
- Analizar la posibilidad de construir una versión visual del lenguaje y comparar qué es más aceptado por los usuarios si una aproximación textual o una visual.

4. Formación de recursos humanos

Las tareas llevadas a cabo en esta línea de investigación están siendo realizadas en diferentes tesis correspondientes a la Licenciatura en Ciencias de la Computación. Se proyecta a corto plazo la continuación de esta investigación con el desarrollo de tesis de maestría y doctorado que permitan profundizar y continuar la investigación en esta área.

5. Bibliografía

1. Daniel Alberto Giulianelli, Claudia Pons, Rocío Andrea Rodríguez, Pablo Martín Vera, and Víctor Fernández. Integrando uml y dsl en el enfoque mda. In XVI Congreso Argentino de Ciencias de la Computación, 2010.
2. Daniel C. Wang, Andrew W. Appel, Jeff L. Korn, and Christopher S. Serra. The zephyr abstract syntax description language. In Proceedings of The Conference on Domain Specific Languages, pages 213–227, 1997.
3. David A Ladd and J Christopher Ramming. Application languages in software production. In USENIX 1994 Very High Level Languages Symposium (USENIX 1994 Very High Level Languages Symposium), 1994.
4. Deursen, A., Klint, P., and Visser, J. Domain-specific languages: An annotated bibliography. SIGPLAN Notices 35(01 2000), 26–36.
5. Jan A. Bergstra and Paul Klint. The discrete time toolbus, a software coordination architecture. *Science of Computer programming*, 31(2-3):205–229, 1998.
6. Julieta Gatica and Camila Olguín. Vinculación Educativa 3.0: Una herramienta para disminuir las brechas educativas regionales. Proyecto final integrador de ingeniería en informática, 2021.
7. Luca Cardelli and Rowan Davies. Service combinators for web computing. *IEEE Transactions on Software Engineering*, 25(3):309–316, 1999.
8. Mariano Luzzi, Mario Marcelo Beron, and Pedro Rangel Henriques. Ph-helper-a syntax-directed editor for hoshimi programming language, hl. In 1st Symposium on Languages, Applications and Technologies. Schloss Dagstuhl-Leibniz-Zentrum for Informatik, 2012.
9. Mioduser, D., Nachmias, R., Tubin, D. and Forkosh-Baruch, A. Analysis schema for the study of domains and levels of pedagogical innovation. In schools using ICT, Education and Information Technologies (2003), pp. 23–36
10. Mora-Murguía, L.-P., Alor-Hernández, G., Olivares-Zepahua, B. A., Ángel Reyes-Hernández, L., and Chávez-Trejo, A. M. Best practices for web development using rails and Django, 2014.
11. Pressman, R. S. *Software Engineering*, 7 ed. McGraw-Hill, USA, 2010.
12. Visser, E. Web dsl: A case study in domain-specific language engineering. vol. 5235/2008, pp. 291–373.
13. Fielding, R. T., and Taylor, R. N. Principled design of the modern web architecture. *ACM Trans. Internet Techn* (2002), 115–150.
14. Herrera-Izquierdo, L., Quiñonez-Ku, X., and Cavada, J. Generador Automático de Aplicaciones Web e Interfaces de Usuarios con funcionalidad Responsiva en el Lenguaje Python. 01 2018, pp. 37–42.
15. Holkner, A., and Harland, J. Evaluating the dynamic behaviour of python applications. In Australasian Computer Science Conference ACSC (2009).18
16. D. Tubin, S. Klein. “Designing a School Website: Contents, Structure, and Responsiveness”. *Planning and Changing* Vol. 38, No. 3&4, 2007, pp. 191–207.
17. Cogent Computer Solutions. “Structure and function of school websites”. Summary report. Junio 2015.
18. J. O. Meiert. “The Little Book of HTML/CSS Frameworks”. (First Edition).
19. H. Suleman, E. A. Fox. “A Framework for Building Open Digital Libraries”. *Revista D-Lib*, Vol. 7 Nro. 12, 2001.

Recursos Educativos Abiertos en la enseñanza de los Números y los Sistemas de Numeración en la escuela primaria

Mariana Alanis, Carlos H. Salgado, Mario G. Peralta

Departamento de Informática, Facultad de Ciencias Físico-Matemáticas y Naturales

Universidad Nacional de San Luis, Ejército de los Andes 950, San Luis

mealanis@email.unsl.edu.ar, {salgado, mperalta}@unsl.edu.ar

RESUMEN

Especialmente estos últimos años, nos han demostrado que las tecnologías de la información y la comunicación pueden ser utilizadas para que el conocimiento se dé más allá de las paredes del aula, de la escuela; que es posible formar redes de conocimiento, por ejemplo, a través de las prácticas educativas abiertas (PEA) y los recursos educativos abiertos (REA). Es importante y necesario que los y las docentes, como así también los futuros y futuras docentes puedan tener una formación adecuada que les permita utilizar y potenciar estas herramientas para el beneficio de los procesos de enseñanza y aprendizaje, en particular los referidos a contenidos específicos de matemática en la escuela primaria, como es el conocimiento del número y del sistema de numeración decimal.

Para ello, en esta línea de investigación se están estudiando, analizando, adaptando y proponiendo estrategias basadas en REA y PEA para la formación de formadores en la enseñanza de Matemática en el nivel primario.

Palabras claves: Prácticas Educativas Abiertas, Recursos Educativos Abiertos, TIC, Número, Sistema de Numeración Decimal, Docentes

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-

Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-03-2020. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

La etapa de residencia de las futuras maestras y los futuros maestros de la escuela primaria se constituye en un momento primordial de iniciación a la labor docente. En esta instancia es necesario recuperar todos aquellos saberes disciplinares y didácticos de cada uno de los espacios curriculares en los que podrían desempeñarse, como lo es en particular el de Matemática.

Por otro lado, es necesario tener en cuenta que las y los estudiantes con quienes se encontrarán en sus futuras aulas, serán en su mayoría, nativos digitales, implicando esto la necesidad de incorporar las TIC a su trabajo diario. Así mismo, en el actual Diseño Curricular para el Profesorado de Educación Primaria que se dicta en el Instituto de Formación Docente Continua San Luis, aprobado por Res. N° 08 – ME – 2022, “se propone el reconocimiento de ejes centrales que actúen como orientadores de las modalidades de trabajo al interior de las carreras docentes” y uno de los ejes que se presenta es “El reconocimiento de competencias y saberes digitales relevantes para la inclusión a la cultura contemporánea y a la sociedad del futuro” [1], pretendiendo abordarlo desde las prácticas educativas

abiertas (PEA) con la creación, reconocimiento y/o modificación de recursos educativos abiertos (REA) en el área de matemática. El acceso a las TIC y la incorporación de PEA como el uso de REA abre un abanico de oportunidades para niños, niñas, jóvenes y adultos, pensando que, a medida que pasa el tiempo, los cambios tecnológicos siguen dándose de manera vertiginosa, lo que puede ser afrontado de manera más apropiada si van adquiriendo experiencia en el uso de los mismos.

Una problemática con la que nos encontramos y es la que motivó este trabajo, es que en la etapa de residencia, las y los estudiantes deben poner en práctica todo aquello que fueron aprendiendo en años anteriores, referido tanto a aspectos disciplinares como así también didácticos de la matemática en particular (también sucede con Lengua y Literatura, Ciencias Naturales y Ciencias Sociales). Sin embargo, en esta etapa demuestran poseer grandes dificultades en recordar los lineamientos didácticos estudiados previamente, recurriendo comúnmente a Internet en busca de material que, muchas veces, dista notoriamente de lo que se espera que sean sus prácticas docentes.

Si bien en diferentes contenidos matemáticos demuestran tener dificultades para seleccionar y secuenciar actividades para ser llevadas al aula, en particular el trabajo se centrará en lo referido a número y sistema de numeración decimal, en los primeros tres grados de la escuela primaria. Respecto de este contenido, las futuras y los futuros docentes deben reconocer que hay actividades que abordan el estudio de las funciones de los números, otras que permiten el estudio de la característica de posicionalidad que posee el sistema de numeración decimal, otras en cambio permiten el reconocimiento de la lectura y la escritura de los números, el orden y comparación de los mismos, entre otras. Por lo que es esperable que puedan crear y/o buscar REA respecto de esta temática en segundo año del profesorado, que puedan hacer sobre ellos un análisis adecuado, y que dichos recursos queden alojados en un repositorio interno, al que

puedan acceder cuando estén en su etapa de residencia, de ser necesario.

Se optó por los recursos educativos abiertos, dado que dentro de lo que es el movimiento educativo abierto se promueve la difusión del conocimiento y “la educación como motor del desarrollo social debería propender por incentivar la construcción y flujo universal del conocimiento, haciendo uso de múltiples canales, entre los cuales y sin duda alguna, los que se soportan en las TIC son los llamados a actuar hoy en día de manera más decidida (Ramirez Montoya & Burgos-Aguilar citado en Chiappe, 2012) [2].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Como lo plantean Ribero y Rabajoli (2016) [3] para innovar, debería tenerse en cuenta la identificación de las últimas tendencias, la implementación de buenas prácticas educativas innovadoras y las prácticas de los docentes innovadores junto con el intercambio de conocimientos y experiencias.

Esto nos lleva a poder profundizar sobre los conceptos de prácticas educativas abiertas y recursos educativos abiertos en conjunto con las y los estudiantes del profesorado de primaria, quienes ya han tenido un acercamiento en su primer año en el espacio curricular de Alfabetización Digital. Se intenta entonces, dar continuidad y ampliar esos contenidos, con vistas a ser reutilizados en su futuro.

Por lo anterior, debemos apreciar que “actualmente las exigencias competenciales de carácter genérico tienen que ver con aspectos no sólo del conocimiento de la disciplina, sino del saber hacer, por ejemplo con el uso y manejo de las TIC, de los procedimientos para trabajar en equipo y para poseer una comunicación más efectiva y del saber estar en un continuo proceso de cambio, en una sociedad globalizada, multicultural y tecnológica, lo que exige de una actitud creativa, de reflexión y de autoperfeccionamiento” [4].

También es significativo destacar la importancia que se da a la construcción del conocimiento por parte de las y los estudiantes, futuras y futuros docentes, la que en Matemática y su Didáctica I está sostenida en el constructivismo, teoría que sostiene que “el estudiante no se limita a copiar el conocimiento, sino que lo construye (constructivismo) a partir de elementos personales, experiencia e ideas previas e implícitas, para atribuir significado (eso es ahora comprender) y representarse el nuevo conocimiento con sentido adquirido (el contenido del aprendizaje)” [5], por lo que la implementación de las TIC dentro del aula, no debe responder al modelo tradicional de enseñanza.

Entonces, se entiende que es necesario generar una propuesta que permita a las y los estudiantes fortalecer su formación disciplinar en matemática, en particular los referidos a número y sistema de numeración decimal en la escuela primaria, asociados a los aspectos didácticos de su enseñanza y atravesados por la utilización de las TIC, por medio de los recursos educativos abiertos.

3. RESULTADOS OBTENIDOS Y OBJETIVOS

Las y los futuros docentes atraviesan distintas etapas a lo largo de su carrera. Es deseable, entonces, promover instancias de acercamiento y familiarización, que pueden estar promovidas por estrategias innovadoras que generen y aceleren esa apropiación. [3]. En Matemática y su Didáctica I, se trabaja con las TIC tanto en instancias de exámenes como en la práctica cotidiana, fomentando su uso desde una postura crítica y de manera cooperativa y haciendo hincapié en la revisión tanto desde lo conceptual como desde lo visual. Así, en dicho espacio curricular, se están utilizando de manera continua las TIC, con el apoyo del espacio de aulas virtuales. Algunas de las herramientas más usadas son Cmap Tools y Mindomo para la confección de mapas mentales. Canva, Genially, Power Point, para

presentaciones para exposiciones de diferentes contenidos. El uso de códigos QR para ingresar a juegos en Educaplay o Liveworksheets. Pizarra Jamboard para compartir procedimientos de resoluciones de diferentes trabajos prácticos. Visualización y análisis de videos, dirigidos especialmente a la gestión de la clase.

El uso de estas herramientas se ha centrado principalmente en la construcción del conocimiento, el trabajo colaborativo y la evaluación formativa. En la construcción del conocimiento porque los productos finales solicitados no se basan en preguntas fácticas, en el trabajo colaborativo porque “cada individuo aprende más de lo que aprendería por sí solo” [6] y en la evaluación formativa porque permite a las y los estudiantes “regular sus propios procesos de aprendizaje, sus debilidades y fortalezas” (Anijovich) [7]

Sin embargo, aún es objetivo de la cátedra, incorporarlos desde la noción de recursos educativos abiertos, con todo lo que este concepto conlleva respecto de las licencias de propiedad intelectual, que en varias ocasiones, no son tenidas en cuenta por las y los estudiantes (y docentes) en la presentación de sus trabajos. Generar este tipo de recursos, puede permitirles retomarlos en su etapa de residencia, modificarlos o no, según el grado donde deban realizar sus prácticas y según los contenidos que les sean asignados desarrollar. Se espera también, que la confección de recursos educativos abiertos, en sus diferentes formatos (imágenes, textos, videos, etc.) les permitan iniciarse en el manejo del Diseño Curricular Jurisdiccional de Educación Primaria, el cual se publicó en el año 2019 y en segundo año es el primer acercamiento que tienen con dicho material.

4. FORMACIÓN DE RECURSOS HUMANOS

Bajo esta línea de investigación, se está desarrollando el Trabajo Integrador Final correspondiente a la Maestría en Enseñanza en Escenarios Digitales, de la Asociación de Universidades Sur Andina- Universidad

Nacional de San Luis. El mismo forma parte del comienzo del desarrollo del proyecto para dar finalización a la carrera antes mencionada, y está pensado y orientado para llevarse a cabo con estudiantes del segundo año del Profesorado de Educación Primaria del Instituto de Formación Docente de San Luis, en el espacio curricular Matemática y su Didáctica I.

Dentro del proyecto de investigación se están llevando a cabo trabajos de grado para optar al grado de ingenieros en informática, licenciados en computación, profesores en informática. A través de desarrollos basados en REA y PEA han realizado trabajos integradores los técnicos universitarios web. Entre ellos, asistentes nutricionales que se utilizan en centros de salud de los barrios periféricos como ayuda a diversos profesionales como nutricionistas, médicos entre otros.

5. REFERENCIAS

- [1] Ministerio de Educación de la Provincia de San Luis. *Resolución N° 8 - ME - 2022*
- [2] Chiappe, A. (2012). *Prácticas educativas abiertas como factor de innovación educativa*. Boletín Redipe, 818, 6-12.
- [3] García J., Báez Sus M. (comp.) (2016) *Educación y Tecnologías en perspectiva*. Flacso Uruguay.
- [4] Tejada Fernandez J., (2007). *La didáctica en un entorno virtual interuniversitario: experimentación de ECTS apoyados en TIC*. Pixel – Bit Revista de Medios y Educación, Universidad Autónoma de Barcelona – España
- [5] Zapata Ross, Miguel (s.f) *Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos*. Dpto de Computación. Universidad de Alcalá. España.
- [6] Ministerio de Educación. (2019) *Trabajo colaborativo y desarrollo profesional docente en la escuela*. Gobierno de Chile.
- [7] Instituto Nacional de Formación Docente. (sf) *Rebeca Anijovich. Evaluación* [video]

MINISATÉLITE CANSAT PARA LA ENSEÑANZA DE LAS CIENCIAS EXPERIMENTALES EN DIFERENTES CONTEXTOS EDUCATIVOS

Fabián Berini, Agustín Fernandez, Ricardo Medel,
Mauro Pessino, Marcos Requena, Santiago Suárez
Universidad Siglo 21 - Córdoba, Argentina
ricardo.medel@ues21.edu.ar

RESUMEN

Los CANSAT son modelos de satélites no orbitales de bajo costo cuyo origen y principal aplicación es la enseñanza de disciplinas científicas (física, matemática, tecnología espacial, computación). Originado en 1999, este dispositivo didáctico permite simular un satélite artificial del tamaño de una lata de gaseosa representando cada una de las partes de los satélites con componentes fáciles de adquirir o construir. El CANSAT es lanzado hasta cierta altura con un cohete amateur, globo aerostático o dron”; posteriormente se deja caer con velocidad ralentizada por un paracaídas. Durante su regreso a tierra transmite la información captada por sus sensores a una *Estación Terrena*, que es, usualmente, una computadora portátil dotada de antena y software para recepción y procesamiento de los datos. Actualmente, dada la complejidad de su construcción e implementación en contextos educativos, su uso se restringe a instituciones técnicas de nivel superior o clubes de ciencia extracurriculares. Nuestra línea de investigación propone expandir su uso a contextos educativos no técnicos, desarrollando un modelo CANSAT de construcción sencilla y bajo costo, acompañado de un currículum flexible, que permita a cada docente elegir la secuencia didáctica más adecuada a los conocimientos de sus estudiantes y las competencias a desarrollar en el curso.

Palabras clave: *CanSat, Educación, Tecnología Espacial, STEAM.*

CONTEXTO

Desde el año 2018, el proyecto *Minisatélite para la enseñanza de ciencias experimentales* es un proyecto de investigación Tipo I de la Universidad Siglo 21. Que el proyecto sea de Tipo I significa que ha sido evaluado y aprobado,

y es financiado por la Secretaría de Investigación y Transferencia Científica de la universidad.

1. INTRODUCCIÓN

Soussan (2003) afirma que se debe asegurar la formación de los futuros ciudadanos y ciudadanas de modo que sepan desenvolverse en un mundo impregnado por los avances científicos y tecnológicos; esto es, que la educación debe promover la adquisición de una “cultura científica”. Este enfoque ha generado la introducción de diversas temáticas científico-tecnológicas en todos los niveles del sistema educativo. Entre tales temáticas se encuentran la programación, la robótica y, más generalmente, las ciencias de la computación, haciendo foco en el “pensamiento en acción”, que permite a los estudiantes elaborar modelos, definir reglas, resolver problemas, inventar, imaginar, crear, juzgar, evaluar y seleccionar las soluciones más adecuadas a cada problema. Frente a posiciones enciclopedistas de la educación, se han aplicado exitosamente la experimentación y manipulación en el aprendizaje y enseñanza de la ciencia y la tecnología.

En particular, en las últimas décadas han comenzado a incluirse asignaturas curriculares o actividades extracurriculares centradas en tecnologías espaciales, con temáticas que incluyen cohetes amateur, análisis de imágenes satelitales y diseño de satélites artificiales (Mermoz, 2003; Barbosa Loureda & Sobral de Araújo, 2008). Estas actividades contribuyen a la apropiación social del conocimiento científico (Massarini, 2011), en particular en un país como Argentina, considerado “satelital” debido a su extensión y particularidades geográficas, y a que tiene capacidades satelitales avanzadas en lo humano y en lo tecnológico.

La cohetaría es generalmente practicada como experiencia extracurricular y aplica conceptos de disciplinas tales como Matemática,

Física, Química y Electrónica (González, 2007), lo cual es de especial interés en los niveles secundario y universitario con orientación técnica (Descalzo & Gómez, 2009).

Por otra parte, el procesamiento de imágenes satelitales comenzó recientemente a incorporarse en las actividades educativas curriculares. Un ejemplo local es el proyecto 2MP de la CONAE, que provee formación a los docentes para que apliquen esta tecnología en sus clases de diversas disciplinas asociadas, tales como Geografía o Historia (Natucci & Rodríguez, 2013; Fornerón, 2013; Marín, 2018).

El desarrollo de la microelectrónica permitió la reducción de tamaños y costos que facilita el acceso y construcción de dispositivos avanzados por parte de docentes y estudiantes. En particular, facilitó la incorporación del diseño y construcción de modelos simples y de bajo costo de satélites artificiales como parte del proceso de enseñanza, especialmente en instituciones universitarias de países centrales (Sorensen, 2017; Kawashima, 2013).

Pequeños satélites o modelos no orbitales de los mismos, ya sean “nanosatélites”, “picosatélites” o “femto-satélites” (ver Tabla 1), se utilizan en instituciones educativas para enseñar, con un enfoque de “pensamiento en acción”, tecnología espacial y sus aplicaciones, así como los conocimientos no técnicos requeridos para llevar adelante una misión satelital: planificación y gestión de proyectos, trabajo en grupo, comunicación, etcétera. (Tsuruda et al., 2009).

Tabla 1: clasificación de pequeños satélites.

Categoría	Masa
Nanosatélite	Entre 1 y 10 kg.
Picosatélite	Entre 100 g. y 1 kg.
Femtosatélite	Menos de 100 gramos

Para facilitar el diseño y construcción de estos pequeños satélites, se han desarrollado algunos estándares que permiten compartir o adquirir componentes. En particular los “CubeSat”, los “PocketQube” y los “CanSat” han tenido éxito en la educación y la industria.

El estándar “CubeSat” fue definido por Puig-Suari y Twiggs en 1999 como un modelo común de un satélite cúbico con 10 cm de arista

y un peso no superior a 1,3 kg. Este modelo es un estándar para satélites universitarios y para toda empresa que desea comenzar a desarrollar su propia línea de satélites (Rêgo et al., 2017).

Basándose en el éxito del modelo de CubeSat, en 2009 la universidad Morehead State University y la empresa Kentucky Space desarrollaron un modelo aún más simple: el “PocketQube”, que define un picosatélite cúbico de 5 cm de arista y no más de 250 gramos de masa (Twiggs, 2009).

El modelo “CanSat” fue definido por el profesor Bob Twiggs en 1998 para ser utilizado en competencias de lanzamiento por medio de cohetes amateurs. Este modelo de satélite no orbital tiene el volumen y la forma de una lata de gaseosa (por ello su nombre, “CanSat”, en inglés: lata-satélite) y no pesa más de 360 gramos. A diferencia de los “CubeSat”, e incluso de los “PocketQube”, los “CanSat” no están diseñados para ser puestos en órbita ni soportar el ambiente espacial, pues realizan vuelos suborbitales transportados por cohetes, globos o drones (Artusa et al., 2013).

A fin de ser útiles como herramienta didáctica en la enseñanza del diseño y construcción de un satélite, los CanSat tienen una estructura que simula la de un satélite: un controlador de a bordo conectado a sensores que captan datos del estado del satélite y del ambiente, una memoria y un transmisor de radio que, respectivamente, almacenan y transmiten durante el vuelo los datos recolectados a un centro de control terreno (Colín, 2016).

Como se indicó previamente, la introducción de tecnologías espaciales en los diversos niveles educativos permite la enseñanza de la ciencia y la tecnología (en algunos ámbitos, referidas por el acrónimo inglés STEM, por ciencia, tecnología, ingeniería y matemática) a través de la manipulación, la experimentación y la curiosidad.

Sin embargo, aún con modelos sencillos de satélites o simuladores, para la mayoría de las instituciones educativas y docentes sin formación específica en estas tecnologías espaciales, resulta complejo implementar actividades de diseño y desarrollo de modelos de satélites (Arruabarrena et al., 2019c).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La finalidad del proyecto “Minisatélite en la enseñanza de ciencias experimentales” es el desarrollo de un modelo de CanSat que pueda ser construido por estudiantes de nivel secundario no técnico y de nivel universitario de carreras no afines a la actividad espacial, acompañado de un currículo flexible que permita a cada docente seleccionar la secuencia didáctica más adecuada para el desarrollo de sus clases, aprovechando, a bajo costo (tanto en sentido monetario como de esfuerzo de aprendizaje y construcción) los beneficios de incluir el uso de tecnologías espaciales en el ambiente educativo.

En este proyecto de investigación nos centramos en la utilización del modelo bajo estándar CanSat debido a su accesibilidad, tanto en costos como en simplicidad de construcción.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo principal original de este proyecto es evaluar la factibilidad de utilizar un modelo CanSat de simulador satelital para la enseñanza de ciencia y tecnología en distintos ámbitos educativos de Argentina. Para lograrlo, definimos los siguientes objetivos específicos.

- ⑩ **Diseñar un modelo de CanSat** de bajo costo y fácil construcción que pueda ser fabricado en Argentina.
- ⑩ **Desarrollar y validar un plan de clases** que, simulando una misión satelital, permita enseñar los temas de ciencia y tecnología de interés para el currículum de diversas instituciones educativas argentinas.
- ⑩ Aproximarse a una **estrategia didáctica que promueva una cultura científica** a través de una nueva enseñanza de las ciencias experimentales y aplicadas superando el tradicional enciclopedismo.
- ⑩ Sentar las bases para el desarrollo futuro de **un laboratorio de estudio de tecnología espacial** en la Universidad Siglo 21.

Durante la ejecución del proyecto, que encontró dificultades importantes debido a la pandemia de COVID-19, se han obtenido los siguientes logros:

- ⑩ Se diseñó y construyó un *demostrador tecnológico* (Figura 1), que permitió realizar pruebas de funcionamiento del sistema electrónico en el laboratorio y comenzar el desarrollo del software de control requerido para el futuro CanSat.

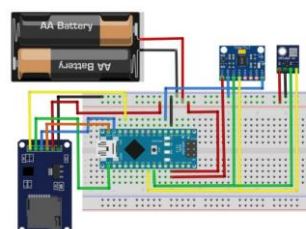


Fig. 1: Esquema de conexión y resultado de la construcción del demostrador tecnológico CanSat.

- ⑩ Se construyó un prototipo de CanSat (Figura 2) a fin de realizar pruebas más complejas en el laboratorio, sin llegar a pruebas de campo.



Fig. 2: Prototipo de CanSat.

- ⑩ Se desarrolló el software de Estación Terrena que permite la obtención y proceso de los datos adquiridos por el CanSat. Por su diseño, el software puede

ser utilizado en un curso tal como está (mostrando gráficamente los datos obtenidos de temperatura, altitud, presión, aceleración y giro en ejes de coordenadas cartesianas) o modificándose su interfaz usando un sistema de “*drag & drop*”, o bien se pueden agregar procesamiento de datos y nuevos gráficos completando un esquema de clase Java (Fernández et al., 2022).

- ⑩ Dada la falta de evidencia científica sobre las competencias que permiten desarrollar los proyectos satelitales en el aula (Arruabarrena et al., 2019c), decidimos analizar cuáles son las competencias que es posible desarrollar en este contexto. Se obtuvo una tabla de competencias, dimensiones, criterios y evidencias que será dada a validación por expertos.
- ⑩ Se comenzó el desarrollo de un conjunto de documentos (Figura 3) que implementa el currículum flexible, estableciendo diversas secuencias didácticas que explican los conceptos y proponen actividades con diferentes niveles de complejidad, desde la utilización de la Estación Terrena a la programación de clases Java para el desarrollo de gráficos de la interfaz.

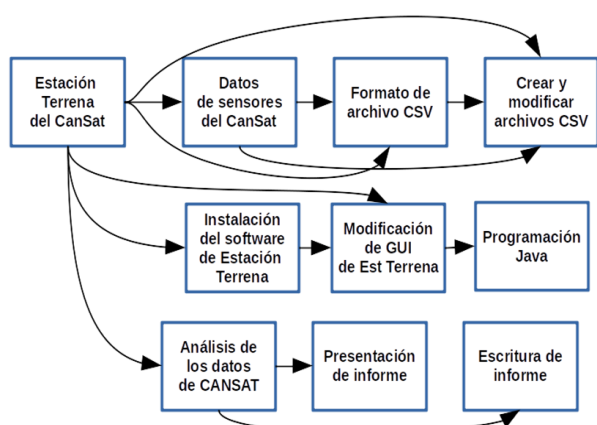


Fig. 3: Segmento del conjunto de documentos implementando el currículum flexible para CanSat.

- ⑩ Se desarrolló el curso de actualización profesional “Satélites educativos CANSAT para la experimentación

espacial”, organizado por la Universidad Siglo 21, el 23 y 30 de junio de 2021.

- ⑩ Se publicaron artículos científicos sobre los resultados obtenidos durante la ejecución del proyecto (Arruabarrena et al., 2019a,b,c; 2020), (Báez et al., 2021), (Fernández et al., 2022).

4. FORMACIÓN DE RECURSOS HUMANOS
Los equipos que ejecutan proyectos Tipo I de la Universidad Siglo 21 deben estar constituidos por un/a director/a, 2 docentes, 2 egresados adscriptos y 2 estudiantes de grado o posgrado. En los cinco años de ejecución del proyecto, el equipo ha ido cambiando de tamaño y estructura, manteniendo siempre el director, el codirector y uno de los docentes. El resto de los y las integrantes ha ido variando durante el tiempo. Por sus filas han pasado otras trece personas, entre docentes, adscriptos/as y estudiantes, provenientes de diferentes disciplinas (informática, sistemas, electrónica, educación, psicología, etc.) y cumpliendo diferentes roles. En la mayoría de los casos han realizado su primera experiencia en investigación científica, lo cual ha aportado a su formación académica.

Además de la formación en RRHH aportada por la participación de docentes, egresados/as y estudiantes en el proyecto, también se desarrolló formación a personas externas al proyecto y la institución a través del curso mencionado en la sección anterior y dos conferencias dictadas por nuestro director:

- ⑩ “Misiones Satelitales para la enseñanza de las Ciencias y la Tecnología” (8 de mayo de 2020), en el marco de un ciclo de charlas organizado por el Centro de estudiantes de la Facultad de Ingeniería y Ciencias Hídricas de la Universidad Nacional del Litoral.
- ⑩ “Tecnología Satelital en la Educación” (15 de septiembre de 2020), realizada en el marco de Edutech-Semana TIC¹.

5. BIBLIOGRAFÍA

Arruabarrena, M., Báez, L., Fernández, A., Lammers, M., Medel, R., Mori, L., Requena, M.

¹Disponible en <https://youtu.be/4Gf5MEOGc-8>

- (2019a). Propuesta de currículum flexible para enseñanza de las ciencias con CanSat. *1º Jornada de Divulgación en Ciencia del Espacio y Tecnología Satelital*. UTN Facultad Regional Córdoba.
- Arruabarrena, M., Báez, L., Fernández, A., Medel, R., Mori, L. (2019b). A Flexible Curriculum for Science and Engineering Courses using CanSat. *2nd IAA Latin American Symposium on Small Satellites*.
- Arruabarrena, M., Fernández, A., Medel, R. & Mori, L. (2019c). Estudio bibliográfico del estado del arte del desarrollo y aplicaciones educativas de CanSats. *X Congreso Argentino de Tecnología Espacial*.
- Arruabarrena, M., Báez, L., Fernández, A., Lammers, M., Marengo, E., Medel, R., Mori, L., Requena, M., Vicente, A. (2020). Minisatélite CANSAT aplicado al Aprendizaje Basado en Problemas en diversos contextos educativos. *Jornadas de Ciencia y Tecnología 2020 50 aniversario de la UTN Fac. Reg. San Francisco*.
- Artusa, J.I., Campiti, N., Competiello, M., Mori Rodríguez, J.M., Rojas, S., Rubio, C., Solares, F., González, P.M. (2013). CANSAT 2012: Metodología Espacial para la Enseñanza de las Ciencias. *VII Congreso Argentino de Tecnología Espacial, CATE*.
- Báez, L., Del Carril, R., Fernández, A., Medel, R., Requena, M. & Romero, R. (2021). Desarrollo de una Estación Terrena educativa para CanSats. *9º Congreso Nacional de Ingeniería Informática y Sistemas de Información - CONAIISI 2021*.
- Barbosa Loureda, O. & Sobral de Araújo, J.B. (2008). Educação Através de Elementos Aeroespaciais. *Revista Latino-Americana de Educação em Astronomia - RELEA*, N° 6. 67-73.
- Colín, Á. (2016). Ciencia y Sociedad: Picosatélites cansat: una herramienta para la educación en ciencias del espacio. *Ciencia UANL*, Septiembre-Octubre 2016.
- Descalzo, G. & Gómez, E.M. (2009). Evolución de Cohetes Experimentales Reutilizables en Actividades Estudiantiles. *5º Congreso Argentino de Tecnología Espacial – CATE*.
- Fernández, A., Medel, R. & Requena, M. (2022). Estación terrena adaptable y su secuencia didáctica flexible para un CANSAT educativo. *Jornadas de Ciencia y Tecnología 2022 - UTN Fac. Reg. San Francisco*.
- Fornerón, C.F. (2013). La tecnología satelital como recurso innovador para la enseñanza: Experiencia pedagógica en el nivel secundario. *I Jornadas Norpatagónicas de Experiencias Educativas en Ciencias Sociales para la Escuela Secundaria*. Río Negro, Argentina.
- González, P.M. (2007). Proyecto CANSAT: Cargas Útiles Estudiantiles en Cohetería Modelo y Experimental. *4º Congreso Argentino de Tecnología Espacial – CATE*.
- Kawashima, R. (2013). UNISEC Challenge: How Can UNISEC Contribute to Capacity Building in Space Science and Technology. *RAST 2013*.
- Marín, H.G. (2018). Parque Nacional Laguna Blanca: una investigación situada para construir en el aula. *Revista Facultad de Ciencias Exactas, Físicas y Naturales*. Vol. 5, N° 1.
- Massarini, A. (2011). El enfoque CTS para la enseñanza de las ciencias: una clave para la democratización del conocimiento científico y tecnológico. *Voces en el Fénix*. N° 8. Sept. 2011.
- Mermoz, J.F. (2003). Ciencias del Espacio en la Educación: Hacia un Nuevo Enfoque de las Ciencias. *2º Congreso Argentino de Tecnología Espacial - CATE*.
- Natucci, S.M. & Rodríguez, M.G. (2013). El programa 2MP en la enseñanza: un ejemplo de aplicación. *I Jornadas Norpatagónicas de Experiencias Educativas en Ciencias Sociales para la Escuela Secundaria*. Río Negro, Argentina.
- Rêgo, A.O., Greco, M., Pereira, M.C. (2017). Design of a 1U Cubesat Platform for Educational Purposes. *1st IAA Latin American Symposium on Small Satellites*.
- Sorensen, J. (2017) Universities in Space: Seriously Higher Education. *Spaceflight.com*
- Soussan, G. (2003). Enseñar las Ciencias Experimentales: Didáctica y Formación. *Oficina Regional de Educación para América Latina y el Caribe - UNESCO*. Chile.
- Tiggs, B. (2009). Making it Small. 2009 Cal Poly Developers' Workshop. California State Polytechnic University.
- Tsuruda, Y., Hanada, T., Van Der Ha, J.C. (2009). System Design and Project Management for University Satellites. *27th International Symposium on Space Technology and Science*.

CALIDAD EDUCATIVA: INSTRUMENTOS PARA ALCANZAR LA MEJORA CONTINUA DE LA EDUCACIÓN SUPERIOR

Juan Marcelo Reus, Carlos Humberto Salgado, Mario Gabriel Peralta, Lorena Soledad Baigorria
Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Universidad
Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina

e-mail: marceloreus@gmail.com, csalgado@unsl.edu.ar, mperalta@unsl.edu.ar, flbaigor@unsl.edu.ar

RESUMEN

En el mundo crece constantemente desde hace años la demanda de profesionales de las tecnologías de información (IT, por sus siglas en inglés) y en tiempos de pandemia se vio incrementada debido a la transformación de las condiciones de trabajo y la necesidad de digitalización y deslocalización de los servicios de las mismas empresas. La gran demanda laboral de profesionales IT, se encuentra en constante expansión y muy lejos de ser satisfecha, o por lo menos así lo demuestra la abundante información que continuamente circula en la actualidad, donde con frecuencia se nos informa que año a año hay un gran número de puestos de trabajo vinculados al sector IT que quedaron sin cubrir. Nos preguntamos entonces: *¿Por qué, las carreras relacionadas con la IT no desbordan de estudiantes? ¿Por qué es tan bajo el número de estudiantes que egresan?*

Conocer el efecto que nuestra docencia tiene en el aprendizaje de los estudiantes, hacerlo visible constituye un marco de referencia innegable, los estudios y las investigaciones pueden provenir de ciudades y culturas diferentes, pero nos interesa todo lo que sucede en la educación como fenómeno global y fundamentalmente en nuestro ambiente de trabajo más próximo, comenzando por el aula (presencial o virtual), nuestra institución, nuestra comunidad, etc.

Así, se está trabajando en la definición de un modelo de calidad que nos permita pensar y repensar los componentes esenciales del proceso enseñanza/aprendizaje. Se están

definiendo métricas, indicadores para poder aplicar el mejor método o forma de cálculo que permita tener un conocimiento de la realidad educativa que se tiene en el contexto de estudio. Para, de esta forma, poder realizar las tomas de decisiones que sean necesarias para cumplir el modelo de calidad propuesto.

Palabras Clave: Modelos de Calidad. TIC. REA. PEA. Objetos de Aprendizaje. Métricas. Indicadores.

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-03-2020. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

Cuando hablamos de calidad educativa, a menudo la atención se dirige al rendimiento de los estudiantes, lo que a nuestro criterio resulta por lo menos incompleto, dado que la calidad en este sentido debiera ser el resultado de evaluar todos los componentes involucrados, incluido el proceso mismo. Planteamos entonces la necesidad de un concepto de calidad holístico, es decir, una guía para medir

y mejorar todos los componentes, incluyendo la perspectiva y la demanda de todos los interesados. El modelo EDPC (Estudiante-Docente-Proceso-Contenido) propuesto en esta línea de investigación, se centra en los componentes básicos e indispensables y se ocupa de conocer y satisfacer a quienes participan del proceso enseñanza-aprendizaje, pero también a quienes esperan a nuestros profesionales egresados para contratarlos, esto resulta esencial para cubrir la brecha entre la oferta y la demanda laboral existente. En el mundo, diversos estudios proponen retroalimentación personalizada de datos de estudiantes y detectan factores que afectan de manera positiva y negativa.

El término “**calidad educativa**”, aparece rápidamente asociado a un sin número de investigaciones que describen innumerables factores que lo afectan de manera positiva y negativa. Así lo demuestra uno de los estudios más recientes e importantes, al menos en cuanto al volumen de datos, una investigación basada en 800 meta-análisis que han supuesto un total de 50.000 estudios y una muestra de 80.000 estudiantes (Hattie, J., 2015).

Conocer el efecto que nuestra docencia tiene en el aprendizaje de los y las estudiantes (conocer, escuchar, atender y satisfacer a quienes participan en un proceso de enseñanza-aprendizaje) resulta esencial para quien está interesado en mejorar algo. En este sentido, la técnica: “Analítica de aprendizaje” (Pardo, 2014), propone una retroalimentación personalizada de datos de los y las estudiantes. Así, el modelo EDPC se enfrenta a ese desafío, considerando a los actores esenciales en un primer nivel o versión simplificada, y a todos los que puedan sumarse en una versión más amplia.

En cuanto a los contenidos, también es posible establecer estándares o criterios de calidad, no solo desde sus aspectos técnicos, sino también a los pedagógicos. En este sentido, de las nuevas tecnologías surgen ideas y propuestas muy originales sustentadas en las TIC. Un ejemplo de ello, son los Objetos de Aprendizaje, También resulta desafiante conocer la influencia del uso de TIC, REA y PEA en los aspectos mencionados.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El principal compromiso de esta línea de investigación, es proveer una guía que permita construir un modelo de calidad, analizando y sistematizando los factores considerados de interés para una determinada comunidad, buscando la evidencia que determine debilidades, fortalezas y oportunidades para mejorar las prácticas educativas y mejorar la calidad de nuestra educación.

Como partes de una sociedad que demanda soluciones a nuevos problemas constantemente, el modelo se basa en tres áreas de la ciencia: la Ingeniería de Software, los estándares internacionales para la educación de calidad y los aportes más recientes de la neuroeducación.

La **NeuroEducación** es la disciplina que estudia el funcionamiento del cerebro durante el proceso de **enseñanza-aprendizaje**: analiza el desarrollo del cerebro humano y su reacción a los estímulos, que posteriormente se transforman en conocimientos (Cristina Saez, 2014).

El Modelo EDPC propuesto se enfoca en cuatro componentes esenciales e imprescindibles en el proceso de Enseñanza-Aprendizaje, donde **E** representa al ESTUDIANTE colocado en primer orden por ser considerado el principal componente y la razón de la existencia del proceso enseñanza-aprendizaje, **D** representa al DOCENTE, en segundo orden y es quien debe mantenerse sensible a todo lo que sucede en el escenario que lo vincula fuertemente con el estudiante y los contenidos. La **P** representa al PROCESO que involucra a todos los componentes (interacciones, datos, informes y resultados) que pretende abarcar el modelo en su mínima expresión, por último, la **C**, los CONTENIDOS, creados o propuestos y seleccionados cuidadosamente por los docentes, que finalmente serán accedidos por los estudiantes.

Para la definición del modelo no se consideró una única definición de “**calidad educativa**” resultante de una determinada corriente de pensamiento o filosofía, sino que se considera

que el concepto debe construirse en cada escenario de trabajo, teniendo en cuenta los actores que intervienen, los que serán consultados, para hallar en un primer momento las dimensiones “*aun cuando la calidad haya sido definida y/o medida en base a un número de indicadores*”, no será más que un acercamiento, una aproximación en el camino que luego se tome para intentar alcanzarla. Es por esto que, para construir los indicadores, se deben consultar las partes interesadas y tener en cuenta la apreciación que tienen sobre los componentes del modelo.

Para la aplicación del modelo, se han definido 3 etapas:

Etapla N°1: Recolección de datos: Se plantean algunas preguntas generales que se volverán más específicas de acuerdo al instrumento elegido para esta tarea (entrevista, encuesta, etc.), partiendo de la pregunta principal *¿Quién es un buen estudiante?* De la misma manera se procede con los componentes restantes. Para hallar la respuesta, existe gran cantidad de bibliografía que puede ser consultada. Pero, estamos interesados principalmente en conocer las respuestas, en principio de los mismos estudiantes, una mirada de pares, también la respuesta que tienen, sus docentes, sus familias, las organizaciones que luego los contratan, etc.

Existe en la actualidad algunos estándares que pueden ser consultados, actualizados y/o adaptados según requiera el perfil de la carrera que se trate, por ejemplo: Los estándares (ISTE 2016), creados por la Sociedad Internacional para la Tecnología en Educación, proponen una serie de competencias que estudiantes, docentes y líderes educativos deben desarrollar para el aprovechamiento de la tecnología en su desarrollo.

También podemos plantear haciendo un juego de palabras aquí, *¿Quién es un buen docente?* Es muy probable que la aplicación de este modelo se encuentre en manos de profesionales docentes, entonces se plantea aquí un doble desafío, hallar las dimensiones y construir los indicadores que requiere el modelo y una gran capacidad de autocrítica.

Edwards Deming, en (Deming 1994) se pregunta ¿Cómo podemos mejorar la educación?, y afirma que mejorar la educación, y su dirección, requiere de la aplicación de los mismos principios que deben utilizarse para mejorar cualquier proceso, también nos apunta que el primer paso de cualquier organización es trazar un diagrama de flujo que indique cómo depende cada componente de los demás. En cuanto a los contenidos, también es posible establecer estándares o criterios de calidad, no solo limitados a sus aspectos técnicos sino también a los pedagógicos. En este sentido, las nuevas tecnologías proponen y surgen ideas muy originales sustentadas en las TIC. Un ejemplo de esto son los OA (Objetos de Aprendizaje), para los que también existen en la actualidad estándares de calidad que definen requisitos mínimos técnicos-funcionales, como son: *Interoperabilidad, accesibilidad y durabilidad*.

(Beatriz Garza Gonzalez, 2009) propone un “Modelo didáctico para el diseño de objetos de aprendizaje (OA)” como una guía para la producción de OA, que además sirve de apoyo a la planeación didáctica del currículum en la modalidad no virtual. En este sentido, el modelo EDPC puede proponer o sugerir diferentes métodos para la construcción de OAs, pero también está abierto a la posibilidad de que quienes tienen la responsabilidad de ponerlo en práctica desarrollen sus propios métodos de construcción.

Etapla N°2: Hallar dimensiones: Se busca obtener, de las fuentes de datos, una tipificación que nos permita organizar la información en dimensiones, por ejemplo: Reconocimientos personales, habilidades generales y específicas, aptitudes laborales, flexibilidad ante el cambio, conocimientos, etc. En este punto, en todas las preguntas formuladas, subyace la necesidad de encontrar la percepción de calidad que tienen los actores acerca de los componentes.

Etapla N°3: Construir los indicadores para la medición y valoración de cada uno de los componentes: El modelo EDPC es flexible y tiene como premisa fundamental mejorar la calidad de todos los componentes que

intervienen en una versión reducida, como así también en una versión más amplia incluyendo todas las interacciones, comunicaciones, plataformas, recursos tecnológicos, etc. que sean parte del ambiente/escenario donde se aplica. Una vez completado el proceso que implica: Recolección de Datos, Hallar Dimensiones y Construir Indicadores para cada componente, resta asignar los pesos a cada indicador de acuerdo con los criterios de quienes se encuentren a cargo del modelo, posibilitando entonces hallar un valor de CALIDAD TOTAL para el escenario de interés.

3. RESULTADOS OBTENIDOS/ESPERADOS

Actualmente, el proyecto se encuentra en plena ejecución. La etapa 1, recolección de datos, es la que permanentemente está en ejecución y retroalimentándonos. La instanciación del modelo nos permite conocer una tendencia de las opiniones, además de incrementar la base de datos con datos e información que continuamos recibiendo. Se ha logrado identificar términos o características deseables o esperables en el proceso de enseñanza-aprendizaje, que permitieron la definición de métricas e indicadores que son los instrumentos para corregir, mejorar o cambiar las tareas y procesos que son el motor de la educación, enfocados en los distintos actores que en ella intervienen. Las bondades que presentan las tecnologías de la información, y el nuevo escenario con la pandemia y postpandemia han potenciado **las prácticas educativas abiertas** como también **el uso de los recursos educativos abiertos**. A tal punto que son el soporte, para parte de esta investigación en lo que respecta al alcance que se logra y las posibilidades que se abren.

En este trabajo se ha podido recolectar datos e información variada desde la perspectiva de los distintos actores. Entre los términos que se hacen común en las percepciones tanto de estudiantes como docentes surgen: paciencia, dedicación, puntualidad, constancia, responsabilidad, actitud, educado, ordenado, colaborador, atento, curioso, predispuesto, etc. La obtención de estas características permite tener una categorización de las opiniones y

posterior tabulación para su análisis y estudio. Estas características y atributos se obtuvieron desde la opinión de diversos actores como estudiantes, docentes, sociedad, expertos consultados, entre otros. En base a ello se han definido métricas e indicadores que posibilitaron realizar la medición de cuánto se estaba cumpliendo el modelo de calidad propuesto. La posibilidad de generar informes al utilizar el modelo, sirve como documentación que permita seguir la historia de la evolución del proceso educativo, a través de la comparación de la situación actual con la situación futura, aportando a la mejora continua del proceso de enseñanza aprendizaje.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se trabaja en lo referente a distintos modelos y métodos de evaluación de calidad. Se está colaborando con las instituciones educativas involucradas de la provincia de San Juan en lo que respecta a retroalimentación de las tareas áulicas y de interacción con los docentes y personal administrativo en lo que respecta a la instanciación del modelo de calidad. Este análisis y estudio permite hacer ajustes tanto en el modelo de calidad como en las tareas y procesos utilizados para el estudio.

Se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) del Lic. Juan Marcelo Reus, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Computación y trabajos finales de carrera de la Ingeniería en Informática e Ingeniería en Computación de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

Bibliografía

W. Edwards Deming (1994) LA NUEVA ECONOMÍA. Para la industria, el gobierno y la educación.
Hattie, J., & Zierer, K. (2020). A GUIDE TO

VISIBLE LEARNING. VISIBLE
LEARNING IN THEORY AND PRACTICE.
In R. FreeBook (Ed.).
Pardo, A. (2014). Learning Analytics.
Beatriz Garza Gonzalez (2009) Modelo
didáctico para el diseño de objetos de
aprendizaje.

Webgrafía

Cristina Saez octubre 6, 2014. QUO.
NEUROEDUCACIÓN, O CÓMO EDUCAR
CON CEREBRO
<https://cristinasaez.wordpress.com/2014/10/06/neuroeducacion-o-como-educar-con-cerebro/>

EVALUACIÓN DEL DESEMPEÑO DOCENTE EN EL NIVEL SUPERIOR

Fernando Sánchez Arroyo (1), Carlos H. Salgado (2), Mario G. Peralta (2), Alberto A. Sánchez (2)

(1) Departamento Académico de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de La Rioja, Av. Luis María de la Fuente s/n, Ciudad Universitaria de la Ciencia y de la Técnica, La Rioja, 5300.
fsanchez@unlar.edu.ar

(2) Departamento de Informática, Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis, Ejército de los Andes 950, San Luis
{salgado, mperalta, alfanego}@unsl.edu.ar

RESUMEN

En el contexto de un mundo globalizado y competitivo como nunca antes, no son pocas las organizaciones y empresas que buscan día a día permanecer en el mercado. Uno de los factores principales para lograrlo es la calidad. Si bien la calidad ya era tenida en cuenta desde épocas antiguas, según Piattini [1] es a partir de la Revolución industrial que se han logrado grandes avances en el tema, a través de diversos estudios e investigaciones que derivan en los sistemas y estándares de calidad de hoy en día.

En las organizaciones, los Sistemas de Gestión de Calidad son los encargados de establecer una relación entre distintos elementos de la organización con el objetivo de producir calidad en lo que hacen. Un ejemplo que podemos citar es la familia de normas ISO 900X, conformada por la ISO 9000: sistemas de gestión de calidad, principios y vocabulario [2]; ISO 9001: requisitos de los sistemas de gestión de calidad [3]; ISO 9004: recomendaciones para las mejoras [4].

Parte fundamental de estos Sistemas de Gestión de Calidad son los Modelos de Calidad, los cuales presentan un guía a seguir o un marco para desarrollar actividades que aseguren cumplir con la calidad de un producto o un servicio.

A pesar de las bondades de los sistemas de gestión de calidad y de haber muchas organizaciones que utilizan estos métodos para gestionar calidad y mejorar sus productos o servicios, no han sido muy utilizados en el

ámbito de la educación, y menos en el caso que nos compete que es de “evaluación docente”. Al menos no se conoce un modelo estandarizado para tal fin.

Presentamos en este trabajo una experiencia de desarrollo de un modelo de calidad para el caso particular del “sistema de evaluación docente” de la Universidad Nacional de la Rioja (UNLaR), Departamento de Ciencias Exactas Físicas y Naturales (DACEFyN). Derivamos este modelo a partir de un conjunto de características y subcaracterísticas relevantes contenidas en la ordenanza 171/19 del Consejo Superior de la Universidad, el cual sienta las bases y condiciones para el sistema de evaluación docente de la Universidad.

Asociado a este modelo de calidad se deriva un conjunto de métricas e indicadores el cual permite la instanciación del modelo de calidad para la evaluación de desempeño y estado de cada docente respecto de las características del modelo de calidad propuesto.

En este caso la pretensión es que el modelo de calidad permita a un grupo de evaluadores emitir su opinión sobre el desempeño llevado a cabo por un docente en un período de tiempo determinado, de forma de ratificar lo realizado por el docente y de tomar las acciones correctivas que hicieran falta.

Palabras Clave: Calidad. Modelo. Desempeño docente. Evaluación.

CONTEXTO

El presente trabajo se enmarca en el Proyecto

de Investigación: Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-03-2020. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

El cuerpo docente de una institución universitaria es, junto a los estudiantes, no-docentes y graduados, uno de los pilares fundamentales de ésta. Los docentes son los que realizan las actividades de enseñanza e investigación en la universidad, por lo que está demás decir que los mismos necesitan de un desempeño adecuado y de un nivel de calidad mínimo para aspirar a la excelencia académica.

Además, las instituciones universitarias deben cumplir con estrictos procedimientos de control y de evaluación, que en el caso de nuestro país están a cargo de CONEAU (Comisión Nacional de Evaluación y Acreditación Universitaria) para que las carreras que en ellas se dictan tengan el reconocimiento oficial de éste, el máximo organismo educacional a nivel nacional. Cumpliendo con las normativas establecidas por CONEAU, una carrera universitaria puede asegurar que los títulos que otorgue a sus graduados cumplen con ciertos estándares básicos que certifican que dichos graduados poseen las competencias necesarias para llevar a cabo una determinada profesión. De alguna manera, esto garantiza la calidad de los títulos, o, dicho de otra forma, se pretende garantizar la calidad de los conocimientos de quien obtenga el título.

Es aquí donde se puede apreciar que se necesitan procesos de calidad para obtener productos o servicios de calidad. Uno de esos procesos es la enseñanza, la cual involucra al docente. Por lo tanto, cuanto mejor sea el desempeño docente más probabilidades hay de

obtener un mejor producto, o de obtener un título de calidad. Es necesario, entonces, contar con algún mecanismo que permita que la universidad pueda controlar y verificar ese desempeño.

Acorde a ello, se está desarrollando un modelo de calidad que permita medir el desempeño docente. Este modelo ayudará a mejorar la calidad de los procesos educativos, y brindará un importante aporte en cuanto al aseguramiento de las competencias adecuadas para la enseñanza, la investigación y la extensión, en otras palabras, ofrecerá una pauta para evaluar la calidad del desempeño docente.

En función de ello, y como se mencionó previamente, consideramos que es de gran utilidad proveer instrumentos que incluyan la evaluación del desempeño docente.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Para contextualizar el trabajo presentado, consiste en un modelo de calidad propuesto que pretende ser de uso continuo en la institución, debido a que la evaluación no debe ser entendida como algo momentáneo sino como algo permanente. Resaltamos esto porque la evaluación no debe ser realizada solo en procesos de acreditación de carreras, tal como sucede con algunas carreras de la UNLaR en este año, sino en todo momento si es que se quiere lograr la calidad deseada.

Analizando la ordenanza rectoral 171/19 que regula la actividad docente de la UNLaR, incluidos los procesos de evaluación docente, se determinan las características y subcaracterísticas más relevantes del modelo en base a ésta. Es por ello que, se propone un modelo de calidad cuyo objetivo es poder realizar una toma de decisión con respecto al estado de situación de cada uno de los docentes de la institución. El modelo propuesto se basa en las características establecidas en la mencionada ordenanza, la reglamentación del Consejo Nacional de Evaluación y Acreditación Universitaria (CONEAU) y el Convenio Colectivo de Trabajo para Docentes Universitarios. Además del modelo de calidad, se definió un conjunto de métricas e indicadores.

El modelo propuesto se divide en 6 características generales en función de las actividades que se considera que todo docente debe realizar. Dichas características son: (1) Gestión de la actividad académica; (2) Actualización y consistencia de los conocimientos; (3) Actividades desarrolladas en investigación; (4) Actividades desarrolladas en extensión, servicios o transferencias; (5) Actividad de gestión administrativa, gobierno universitario o gremial; (6) Formación de recursos humanos. El modelo se fue calibrando de acuerdo a las distintas instanciaciones que se llevaron a cabo en el Departamento Académico de Ciencias Exactas, Físicas y Naturales (DACEFyN). Esto permitió ajustar y redefinir métricas, así como los indicadores para un mejor entendimiento de la realidad que se está trabajando. Es decir, el modelo está en producción en este momento. Se espera que se pueda instanciar el modelo en los distintos departamentos de la UNLaR y poder realizar los ajustes y tomas de decisiones que sean necesarias. La figura 1 muestra un fragmento del modelo que refiere a la característica 3. Actividades desarrolladas en investigación.

De la desagregación de las características principales, las hojas del árbol mostrado en la Figura 1 constituyen los atributos de calidad del modelo que deberán medirse a la hora de llevar a cabo la evaluación de los docentes respecto de las actividades de investigación que ellos realicen. Para ello, se propone un conjunto de métricas que permitirán realizar la medición de cada uno de dichos atributos y, a partir de dichas métricas se definen los indicadores de calidad para ellos. Una métrica es una definición operativa que describe un atributo del producto o del proyecto. Una medida es un valor real [5].

-
- 3. Actividades desarrolladas en investigación
 - 3.1 Docente investigador categorizado
 - 3.1.1 Categoría del investigador
 - 3.2 Dirección o participación en programas y/o proyectos con evaluación externa o interna
 - 3.2.1 Direcciones en proyectos y/o programas con evaluación interna o externa
 - 3.2.2 Participaciones en proyectos y/o programas con evaluación interna o externa
 - 3.3 Formación de RRHH en investigación
 - 3.4 Publicación, comunicación de resultados de investigación en congresos, eventos científicos
-

Figura 1: Fracción del Modelo de Calidad - Característica 3 y subcaracterísticas

Así, en la tabla 1 solo se detallan las métricas e indicadores para los atributos de calidad de la característica: 3: Actividades desarrolladas en investigación.

En los indicadores para los casos de PCRCE (Publicación o comunicación de resultados en congresos o eventos) se asigna una escala entre dos variables “x máximo” y “x mínimo” (Xmax y Xmin) los cuales toman valores de acuerdo al entendimiento de un experto (docente e investigador con al menos diez años de trayectoria y con antecedentes en comités o comisiones de evaluación). En estos casos la evaluación será “satisfactoria” si se alcanzan o superan los valores de “x máximo” y será “no satisfactoria” cuando los valores sean iguales o inferiores a “x mínimo”. Cualquier valor entre “x máximo” y “x mínimo” es calculado como el cociente entre la variable PCRCE y el “x máximo”, obteniendo de esta manera un valor intermedio que indica la satisfacción parcial del requerimiento.

Tabla 1: Métricas e indicadores para la dimensión 3 Actividades desarrolladas en investigación

ID	Nombre	Descripción	Función de medición	Indicador
3.1	Docente investigador categorizado	¿Es docente investigador categorizado? ¿Cuál es su categoría?	CI=Consultar en el legajo del docente. Si es categorizado debe haber una Resolución al respecto y en donde consta su categoría.	1: Si CI = 1 0.7: Si 1 < CI ≤ 3 0.3: Si 3 < CI ≤ 5 0: Si CI < 5 o CI = VACÍO

3.2	Dirección o participación en programas y/o proyectos con evaluación interna o externa	¿Ha dirigido o ha participado en proyectos de investigación con evaluación interna o externa?		
3.2.1	Dirección en programas y/o proyectos con evaluación interna o externa	¿Ha dirigido proyectos de investigación con evaluación interna o externa?	$DPPEIE = \#DPPEI + \#DPPEE$	Se define: $X_{max}=4$, $X_{min}=1$ (de acuerdo a juicio experto) 1: si $DPPPEEI \geq X_{max}$ $DPPPEEI/X_{max}$: si $X_{min} < DPPPEEI < X_{max}$ 0: si $DPPPEEI \leq X_{min}$
3.2.2	Participación en programas y/o proyectos con evaluación interna o externa	¿Ha participado en proyectos de investigación con evaluación interna o externa?	$PPPEIE = \#PPPEI + \#PPPEE$	Se define: $X_{max}=10$, $X_{min}=2$ (de acuerdo a juicio experto) 1: si $DPPPEEI \geq X_{max}$ $DPPPEEI/X_{max}$: si $X_{min} < DPPPEEI < X_{max}$ 0: si $DPPPEEI \leq X_{min}$
3.3	Formación de RRHH en investigación	¿Ha formado RRHH mediante proyectos de investigación?	$FRHI = \text{Contar los recursos humanos a cargo en los distintos proyectos}$	Se define: $X_{max}=5$, $X_{min}=1$ (de acuerdo a juicio experto) 1: si $FRHI \geq X_{max}$ $FRHI/X_{max}$: si $X_{min} < FRHI < X_{max}$ 0: si $FRHI \leq X_{min}$
3.4	Publicación, comunicación de resultados de investigación en congresos, eventos científicos	¿Ha publicado o comunicado los resultados de alguna investigación?	$PCRCE = \text{Contar la cantidad de publicaciones en los congresos o eventos}$	Se define: $X_{max}=2$, $X_{min}=1$ (de acuerdo a juicio experto) 1: si $PCRCE \geq X_{max}$ $PCRCE/X_{max}$ si $X_{min} \leq PCRCE < X_{max}$ 0: si $PCRCE < X_{min}$

3. RESULTADOS OBTENIDOS/ESPERADOS

Para la instanciación del modelo se recurrió a la solicitud de los datos correspondientes, cuyas fuentes constan de legajos docentes, solicitados a las Direcciones de carrera y Subsecretaría de investigación y extensión, así como de una encuesta personal al docente. Para ello se ha tomado como muestra a los docentes de la carrera de Ingeniería en Sistemas de Información del DACEFyN de la

UNLaR.

Una vez instanciados los datos se pudieron obtener conclusiones acerca de las características analizadas para un conjunto de docentes. Esto permite tener un escenario individual por cada docente y a la vez global, en este caso, del departamento (DACEFyN). En cuanto a lo individual y siguiendo la rigurosidad del régimen de carrera docente se puede establecer si el desempeño de un docente en particular ha sido satisfactorio en el

período o no. De igual manera, por extensión, se puede saber acerca de un grupo de docentes, por ejemplo, de un departamento o una carrera en particular, donde se pueden encontrar las falencias y por consiguiente qué medida correctiva o que recomendación se puede dar al respecto. Por ejemplo, si en una determinada carrera se detecta que los docentes tienen una evaluación no satisfactoria en el punto 3.4 antes mencionado, es un indicador de que los docentes no hacen publicaciones, entonces cabe preguntarse si es porque no tienen trabajos de investigaciones o si al tenerlos no pueden publicarlos. Luego habrá que tomar una acción correctiva: incentivos de investigación, difusión de congresos y eventos por ejemplo.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se trabaja en lo referente a distintos modelos y métodos de evaluación de calidad. Se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis, del Ing. Fernando Sánchez Arroyo, quien se desempeña como docente de la Universidad Nacional de La Rioja.

Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Computación y trabajos finales de carrera de la Ingeniería en Informática e Ingeniería en Computación de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis y del Departamento Académico de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de La Rioja.

5. REFERENCIAS

- [1]. M. Piattini, F. Ó. Garcia Rubio, and I. Caballero, Calidad de Sistemas Informáticos: Alfaomega-RA-MA, (2007).
- [2]. Norma ISO 9000:2015
<https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:es>
- [3]. Norma ISO 9001:2015
<https://www.iso.org/obp/ui/#iso:std:iso:9001:ed-5:v1:es>
- [4]. Norma ISO 9004:2015
<https://www.iso.org/obp/ui/#iso:std:iso:9004:ed-4:v1:es>
- [5]. DeMarco, T. Controlling Software Projects: Management, Measurement and Estimation. USA, (1986).

SOFTWARE PARA GAMIFICAR TERAPIAS DE REHABILITACIÓN AD-HOC

Ferrarini Oliver, Cintia¹, González de Doña, Mónica Gilda², Ormeño G. Emilio⁴, Renzo Sessa⁵ y Aballay Alicia³

^{1,2y3}.Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.

⁴ Instituto de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, UNSJ.

⁵ Centro Educativo Terapéutico Casita Amarilla (CET)

{¹ferrarinicintia, ²gonzalez.monicaagilda, ³prof.alicia}@gmail.com; ⁴eormeno@iinfo.unsj.edu.ar,

⁵sessarenzo14@gmail.com

RESUMEN

En este trabajo se presenta los resultados obtenidos del Proyecto de Desarrollo Tecnológico y Social PDTs UNSJ. Resolución N°589/R, denominado “Software para Gamificar Terapias de Rehabilitación Ad-Hoc”, cuyo objetivo fue desarrollar un software basado en componentes para dispositivos móviles que facilite el diseño, aplicación y evaluación de experiencias gamificadas de rehabilitación ad-hoc, destinado a los profesionales del Centro de Rehabilitación, Diagnóstico y Asesoramiento (CERyDAR). Se realizó una investigación aplicada, en la que se priorizó una metodología cuanti-cualitativa y se optó por un diseño descriptivo y experimental. Se desarrolló un prototipo, las validaciones de usabilidad se realizaron con terapeutas que realizan intervenciones terapéuticas en el Centro Educativo Terapéutico Casita Amarilla (CET).

Palabras claves: Gamificación, intervenciones terapéuticas ad-hoc, software, videojuegos, dispositivos lúdicos.

CONTEXTO

Las actividades de investigación y desarrollo se llevaron a cabo en el Gabinete de Gamificación del Instituto de Informática de la Universidad Nacional de San Juan, con domicilio en calle Meglioli Sur 1050, edificio CUIIM, Rivadavia, San Juan, Argentina.

INTRODUCCIÓN

Actualmente en las intervenciones terapéuticas los especialistas (kinesiólogos, psicopedagogos, terapeutas ocupacionales, entre otros) que asisten a personas con algún tipo de discapacidad, con frecuencia desarrollan actividades y/o construyen elementos físicos específicos para contribuir a su rehabilitación o mejora de alguna

condición. La utilización o adaptación de juegos, juguetes, o elementos de uso diario, en el proceso de rehabilitación (físico como: ángulo, flexibilidad, duración, fuerza, entre otros; o psicológico como atención, memoria, concentración). Aun cuando los terapeutas están familiarizados con estos recursos, a través de entrevistas semiestructuradas con profesionales de la institución adoptante (Centro de Rehabilitación, Diagnóstico y Asesoramiento (CERyDAR)), expresaron la importancia de contar con un mecanismo integrador que les permitiera: a) evaluar el progreso de un paciente, b) obtener una retroalimentación de advertencias o incentivos y c) "mejorar la experiencia" de rehabilitación a través de la incorporación de estímulos sonoros, luminosos, vibraciones, entre otros. Las actividades de esta institución se focalizan de modo que abarquen la adquisición de la noción de sí, el autovalimiento, que según Rodríguez (2019) es la destreza de valerse por sí mismo, que facilitan a la persona realizar actividades diarias como asearse, cambiarse, alimentarse, comunicarse, relaciones con los otros y con objetos de la realidad en sus niveles más prácticos o conceptuales, mediante el recurso de vinculación personalizada, el juego y el trabajo con pares y adultos, incluyendo tanto los vínculos primarios, familiares, como secundarios y sociales. Para "mejorar la experiencia", existen dispositivos electrónicos abiertos con tecnología inalámbrica bluetooth de baja energía (Bluetooth Low Energy), cuyo propósito es la difusión de la programación y de la electrónica con la filosofía DIY (Do It Yourself); que podrían usarse para mejorar la intervención terapéutica. Sin embargo, para lograr los requerimientos de "evaluación" y "retroalimentación", se requiere el desarrollo de una plataforma de software para vincular, ensamblar y sincronizar estos dispositivos

entre sí, para que el terapeuta diseñe y construya una terapia específica y adecuada a las características de su destinatario.

El objetivo del proyecto, fue desarrollar un software basado en componentes para dispositivos móviles que faciliten el diseño, aplicación y evaluación de experiencias gamificadas de rehabilitación ad-hoc, destinado a los profesionales de CERyDAR. Se realizó una investigación aplicada, en la que se priorizó una metodología cuantitativa y se optó por un diseño descriptivo y experimental. Se desarrolló un prototipo del software, las validaciones de usabilidad se realizaron con terapeutas que realizan intervenciones terapéuticas en el Centro Educativo Terapéutico Casita Amarilla (CET).

El software generado significó un apoyo a los terapeutas en el proceso de diseño, desarrollo y evaluación de experiencias de rehabilitación ad-hoc. Para alcanzar este objetivo se realizó en un primer momento entrevistas focalizadas. Luego se realizaron instancias experimentales y por último se desarrolló un prototipo de videojuego como se presenta en el apartado de resultados.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La investigación está orientada al área de Videojuegos para la salud y la educación. El uso de dispositivos lúdicos pueden ensamblarse en forma física y lógica, para mejorar la experiencia en las intervenciones terapéuticas de rehabilitación ad-hoc, a la vez brindan información valiosa sobre la evolución del destinatario. Constituye no sólo un enfoque original en el ámbito local, sino que a nivel internacional son escasas las investigaciones que contemplan el desarrollo de las funciones cognitivas y motoras desde ese enfoque.

"Gamificación" refiere al uso de elementos de videojuegos en sistemas ajenos al juego para mejorar la experiencia y la participación del usuario (Deterding, Sicart, Nacke, O'Hara y Dixon, 2011). En los sistemas de gamificación, los elementos de juego más utilizados son puntos, insignias y tablas de

clasificación, que también se conocen como "La Tríada PBL". El término "dispositivos lúdicos" refiere a tecnologías que, como requisito fundamental, deben proveer comunicación por Bluetooth o radiofrecuencia: Entre otras, Tecnologías de hardware y software que promueven el aprendizaje de la programación y la electrónica como BBC Micro:bit, Little bit o Makey Makey; Dispositivos para monitoreo de la salud (relojes inteligentes, contador de pasos, y pulseras de respuesta galvánica de la piel); Cámaras de profundidad como kinect; Sensores con comunicación inalámbrica.

3. RESULTADOS OBTENIDOS / ESPERADOS

En cuanto a las entrevistas focalizadas, cuyo objetivo fue conocer qué significado tenía para los terapeutas consultados; los "juegos" o "juguetes" como recursos para intervenciones terapéuticas ad-hoc y en qué medida estarían dispuestos a usar videojuegos en las mismas. Se establecieron las siguientes categorías de análisis:

-comprensión sobre qué significa juegos o juguetes,

-tipos de juegos y/o juguetes que emplean

-dificultades detectadas en el uso de juegos y/o juguetes,

-uso de pantallas /videojuegos con fines terapéuticos

-conocimientos sobre softwares gamificados

Se reconoce en la información relevada, la necesidad de generar innovación en la forma de abordaje de las intervenciones terapéuticas.

Diseño y desarrollo de instancias experimentales

Se diseñaron instancias experimentales utilizando el dispositivo de realidad virtual Oculus Quest 2 con la participación de niños y terapeutas del Centro Educativo Terapéutico Casita Amarilla (CET). Se seleccionó para las instancias experimentales esta institución porque garantizaba mayor campo de prueba por la cantidad de profesionales y de niños que la que cuenta, que se podía hacer uso del casco y por qué contaba con un equipo

profesional interdisciplinario, cuyos aportes enriquecieron el desarrollo del software propuesto. Esta institución surge con una nueva oportunidad de aplicación de experiencias debido a la imposibilidad de llevar a cabo instancias experimentales con los integrantes de la institución destinataria por cuestiones organizativas de la misma. Las instancias experimentales se desarrollaron en las siguientes cuatro fases:

Primera fase: Exploración Inicial

Se realizó una prueba piloto de uso del dispositivo de realidad virtual Oculus Quest 2 con algunos integrantes del equipo de investigación y participantes voluntarios. Para identificar posibilidades de aplicación, barreras, entre otros aspectos. Figura N°1



Figura N°1

Luego, se realizó un encuentro en el que se presentó el dispositivo de realidad virtual Oculus Quest 2 a los directivos y terapeutas del CET, con fin de conocer sus inquietudes y necesidades en relación al uso de esta tecnología para ser aplicada a las intervenciones terapéuticas ad-hoc. Como resultado de esta primera fase se definió un procedimiento denominado “Secuencia de Reconocimiento y Aprendizaje del Dispositivo” (SERAD). El mismo facilitó, según señalaron los integrantes de la institución CET, el aprendizaje en la forma de uso correcta de los visores y el reconocimiento de elementos importantes para la transición del plano físico al virtual y la configuración de la conexión entre el Oculus Quest 2 con otros dispositivos. El tiempo de aprendizaje para el manejo del sistema, arrojó un valor aproximado de 30 minutos en promedio.

Segunda Fase: Exploración del dispositivo por profesionales

En esta fase se seleccionó en forma aleatoria 10 profesionales del CET y se les solicitó que

aplicaran el procedimiento denominado SERAD obtenido de la fase anterior. Los profesionales involucrados pertenecían a diferentes disciplinas: psicología, psicopedagogía, terapia ocupacional, kinesiología, fonoaudiología y docentes de educación especial dedicados a la estimulación y rehabilitación.

Como resultado de esta segunda fase fue posible conocer la percepción de los profesionales con respecto a factibilidad en el uso de esta tecnología como medio para viabilizar las intervenciones terapéuticas ad-hoc. Quienes afirmaron que podrían utilizar el recurso en sus espacios de intervención. En promedio el tiempo de reconocimiento y utilización fue de 35 minutos.

Tercera Fase: Exploración del dispositivo por los estudiantes

A partir de la exploración a cargo de los especialistas, el dispositivo/recurso es utilizado en cada espacio. A continuación, se plantean objetivos de cada profesional y resultados obtenidos.

Área psicopedagogía

Objetivos: trabajo de atención y grafismo en letras. Cantidad de estudiantes: dos

Experiencia: La experiencia se dividió en dos instancias, una por cada estudiante. En el primer caso, el estudiante presentaba gran dificultad de conectar con el entorno y reiterados automatismos, con el objetivo de mejorar su capacidad de fijar o focalizar la atención, se utilizó la aplicación “First steps” (Primeros Pasos). Como resultado: el alumno, al momento de colocarle el visor, logró regular su cuerpo, movilizar su cabeza hacia el personaje y fijar la atención ante los sonidos.

En el segundo caso, el estudiante presentaba dificultades motrices lo que impedía el trabajo en el cuaderno. El objetivo fue diseñar letras. Para ello se trabajó con la aplicación “gravity sketch”. Como resultado, el estudiante comenzó a garabatear, se presentan algunas imágenes. Figura N°2. Desde la perspectiva del profesional esto significó un gran logro a seguir trabajando.

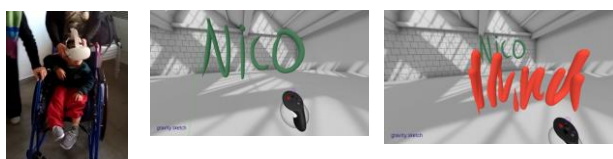


Figura N°2

Área: Kinesiología- Ed. Física

Objetivos: control postural, movilización de miembros superiores. **Cantidad de estudiantes:** dos con dependencia de silla de ruedas. **Experiencia:** Se realizó una secuencia breve de reconocimiento del dispositivo con cada alumno. El terapeuta utilizó la aplicación “Vr fit” y “Dance Center”, durante un periodo de una hora.

Resultados: El terapeuta señaló la importancia de este proceso ya que los estudiantes lograron tener mayor atención al momento de realizar actividades básicas, como llevar los miembros superiores a un determinado punto. Un aspecto importante a destacar es que fue posible coordinar al paciente en sus movimientos. Tal como se muestra en la Figura N°3



Figura N°3

Área fonoaudiología

Objetivo: identificación de secuencias, socialización comunicativa y memoria. **Cantidad de estudiantes:** tres. **Experiencia:** Se realizó una secuencia breve de reconocimiento del dispositivo con cada alumno. Se presentó, acorde a las preferencias de cada alumno, un video de YouTube VR en el que se realizaron, vuelos, cortos animados y experiencias de paisajes. Al quitarse el dispositivo, la terapeuta mostró un conjunto de imágenes a fin de reconocer qué elementos estaban y cuáles no, dentro de la experiencia, con algunos de los concurrentes, realizó la reproducción de la secuencia. **Resultados:** El profesional señaló que es un gran recurso ya que asegura eliminar otros estímulos lo que permite hacer foco en la historia. A partir de

esta técnica se retomó en la siguiente sesión de trabajo con elementos representativos de lo visto en la realidad virtual. Puede apreciarse en la Figura N°4.



Figura N°4

Últimas experiencias realizadas

Un terapeuta de niños con discapacidad describe cómo utilizar la tecnología de realidad aumentada del Oculus Quest 2 en sus prácticas terapéuticas. Para la Experiencia distribuyó a los chicos en tres salas de acuerdo a sus necesidades y utilizó una aplicación de juego, que consistía en usar unos sables luminosos para golpear cubos que se iban acercando al usuario al ritmo de la música, se trabajó en las lateralidades. También empleó la aplicación para trabajar en el reconocimiento de colores. Con algunos pacientes que necesitaban asistencia, utilizó streaming en un celular o computadora para ayudarles a comprender y reforzar lo que debían hacer. Con uno de ellos, en particular, que presentaba un problema conductual grave, utilizó una aplicación de realidad aumentada denominada “Arlupa” para jugar a un juego similar al propuesto por el software denominado “Pokémon Go”, lo que le ayudó a aumentar su atención de menos de cinco minutos a 20 minutos. Finalmente, el terapeuta comenzó a trabajar con Oculus Quest 2 y utilizó videos de diferentes animales, entre otros, dinosaurios en la selva, a fin de continuar trabajando en su atención y curiosidad.

Cuarta Fase: difusión entre directivos

Se generó una instancia de concientización-capacitación destinada a los diferentes equipos directivos. Participaron 20 directivos de las principales instituciones educativas de gestión privada, para personas con discapacidad de la provincia de San Juan. Los participantes exploraron la secuencia de

aprendizaje para la comprensión de la utilización de los visores. Acorde a sus intereses experimentaron con las aplicaciones mencionadas. De la experiencia se rescató el interés por implementar este tipo de tecnologías en sus instituciones. Se muestran algunas imágenes en Figura N° 5 y N°6



Figura N°5



Figura N°6

5. Desarrollo del Prototipo de Software

Teniendo en cuenta los requerimientos y necesidades detectadas en las instancias experimentales, se desarrolló una aplicación para Android que permitiría a los terapeutas crear juegos ad-hoc vinculando diversos BBC-Microbits (BBC micro:bit. (2021)). La aplicación utiliza el sensor de Bluetooth para detectar los BBC-Microbits cercanos. Luego, lee los servicios habilitados de cada uno y permite ensamblarlos para generar una experiencia de juego personalizada para los terapeutas. El objetivo es crear juegos que promuevan la interacción social y el trabajo en equipo.

A continuación se muestran en la Figura N°6

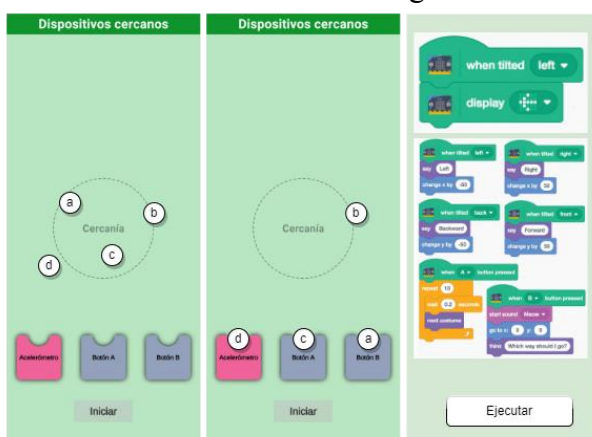


Figura N°6

La aplicación presenta varios beneficios para los terapeutas y sus pacientes. En primer lugar, permite crear juegos ad-hoc que se ajustan a las necesidades específicas de cada grupo de pacientes. En segundo lugar, promueve la interacción social y el trabajo en

equipo, lo que puede tener beneficios positivos para la salud mental de los pacientes. En tercer lugar, utiliza tecnología de vanguardia, como Bluetooth y microcontroladores programables, lo que puede aumentar el interés de los pacientes en la terapia. La aplicación presenta una forma innovadora de utilizar tecnología para mejorar las intervenciones terapéuticas. Al combinar tecnología de vanguardia con la experiencia clínica de los terapeutas, se pueden crear juegos personalizados que promuevan la interacción social y el trabajo en equipo. Sin embargo, es importante tener en cuenta que la aplicación puede presentar ciertos desafíos técnicos para los terapeutas que no estén familiarizados con la tecnología de los microcontroladores programables. Por otra parte, se necesitan más investigaciones para evaluar la efectividad de la aplicación en las intervenciones terapéuticas ad-hoc. La innovación se comprende como creación (invención): El foco de atención está en el uso de recursos (gente, tiempo y dinero) para inventar o para desarrollar un producto o servicio nuevos, o bien, una nueva forma de hacer las cosas, una nueva forma de pensar acerca de ellas (Ahmed, Shepherd, & Ramos Garza, 2012). Para lograr mejoras en la calidad de vida de los pacientes, lo que exige actuar con Responsabilidad Social, y si fuera necesario, a través de la cooperación entre diferentes organizaciones. (Norma internacional ISO 26000, 2010, p.13). La creatividad y la innovación de los profesionales, técnicos y personal de apoyo cumplen un rol importante en la generación de instrumentos y dispositivos, en este caso, juegos y juguetes, que se adapten a las necesidades de los pacientes. La iniciativa, el compromiso, la perseverancia y la empatía son valores esenciales en todo proceso de rehabilitación, desde su diseño, pasando por la prueba hasta el desarrollo y luego, en la aplicación y evaluación, en función de la solución del problema para el cual fueron pensados y desarrollados.

6. FORMACIÓN DE RECURSOS HUMANOS

-Ormeño Emilio Gustavo. Director del proyecto. Actualmente doctorando del Doctorado en Ciencias de la Informática, y cuya tesis está orientada al área de Videojuegos para la salud.

-Ferrarini Oliver, Cintia. Integrante del proyecto. Actualmente realizando una Diplomatura Tecnologías de la Información y Comunicación (TIC) y Accesibles al Conocimiento (TAC): Desafíos inclusivos en la enseñanza a estudiantes con discapacidad visual, auditiva, motora, intelectual y TEA (Trastorno del espectro autista), UNSJ.

-Olguín Luis Alberto. Integrante del proyecto. Actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan. Título ``Red de Co-Préstamo en Bibliotecas”.

-Morales Rubén. Actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan, cuya tesis se denomina “Videojuegos serios para favorecer el desarrollo de competencias para la programación de software”.

5. BIBLIOGRAFÍA

-Ahmed, P. K., Shepherd, C. D., & Ramos Garza, L. (2012). *Administración de la innovación*. Pearson.

<http://190.57.147.202:90/xmlui/bitstream/handle/123456789/531/Administracion%20de%20la%20Innovacion%20K.%20Ahmed.pdf?sequence=1>

-BBC micro:bit. (2021). Retrieved September 14, 2021, from <https://microbit.org/>

-Biswas, D., & Soh, P. J. (2021). A review of micro:bit applications in education: Trends, challenges and future directions. *Education and Information Technologies*, 26(3), 2735-2752. <https://doi.org/10.1007/s10639-020-10484-5>

-Bluetooth. (2021). Retrieved September 14, 2021, from <https://www.bluetooth.com/>

-Brown, T., Ansell, J., & Preece, A. (2021). Micro:bit-based digital interventions for mental health: A systematic review. *Journal*

of Medical Internet Research, 23(1), e24303. <https://doi.org/10.2196/24303>

-Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. In *CHI'11 extended abstracts on human factors in computing systems* (pp. 2425-2428). Vancouver, BC, Canada, 2011, pp. 2425–2428 [Online]. Available: <http://doi.acm.org/10.1145/1979742.1979575>

-Johnson, C., & McDonald, C. (2019). The micro:bit as a tool for learning to code and physical computing. *TechTrends*, 63(3), 312-318. <https://doi.org/10.1007/s11528-019-00389-4>

-Norma internacional ISO 26000, 2010, p.13.

-Rodriguez, E. S. C. (2019). Habilidades básicas sobre desarrollo de autovalimiento. *Revista Pedagógica*, 1(1). IESPP Oxford, Trujillo - Perú.

<http://www.iesppoxford.edu.pe/wp-content/uploads/2019/12/RP1-A09.pdf>

Tecnologías innovadoras aplicadas en contextos educativos

Claudia Russo¹, Mónica Sarobe², Benjamin Cicerchia³, Nicolás Alonso⁴, Gustavo Gnazzo⁵,
Mariana Adó⁵, Natalia Bendatti⁵, Valeria Cassera⁶, Matías Contreras⁷, Gustavo Iglesias⁷,
Tamara Ahmad⁸

{ccrusso, msarobe, lbcicerchia, nfalonso, ggnazzo, mado, nsbendati,
vcassera}@docentes.unsada.edu.ar, {mjcontrera, gviglesias}@alumnos.unsada.edu.ar,
tamaraahmad@unnoba.edu.ar

Universidad Nacional de San Antonio de Areco, Buenos Aires, Argentina
Universidad Nacional del Noroeste de la Provincia de Buenos Aires, Instituto de
Investigación y Transferencia (ITT-CIC)

RESUMEN

En los próximos años las tecnologías innovadoras impactarán en el avance tecnológico, cobrando importancia los desarrollos que apliquen dicha tecnología, en el proyecto se trabajará principalmente con las siguientes tecnologías: Realidad Aumentada, EV3D, Realidad Virtual, Inteligencia Artificial, entre otras

Para el presente trabajo se presenta la continuidad de un proyecto de *I+D+i* (Investigación, Desarrollo e innovación) centrado no sólo en las tecnologías innovadoras y su relación con la Educación a Distancia (EaD) sino su relación con las distintas disciplinas. En el marco del presente y considerando que las tecnologías actuales plantean paradigmas que implican impacto directo en los modelos educativos, se propone investigar sobre las tecnologías informáticas aplicadas a diferentes contextos, haciendo énfasis en sus variantes y evoluciones en las diferentes disciplinas.

A su vez se busca continuar indagando sobre distintas metodologías y técnicas que permitan dotar de interoperabilidad a los sistemas de información de la Universidad Nacional de San Antonio de Areco (UNSAdeA), que tengan relación con tecnologías planteadas en el presente proyecto, realizando experiencias concretas.

PALABRAS CLAVES

Entornos Virtuales, Innovación en Educación, Recursos Digitales, Realidad aumentada y virtual, EV3D, IA.

CONTEXTO

Las líneas de investigación a describir se enmarcan en el proyecto de investigación: Educación a Distancia e innovación Tecnológica, con lugar de trabajo en la Universidad Nacional de San Antonio de Areco (UNSAdeA) presentado en la

¹ Doctora en Ciencias Informáticas, Profesor Asociado CIC, UNNOBA, UNSADA

² Maestrando de la UNLP, Lic. en Informática, Profesora, UNNOBA, UNSADA

³ Doctorando de la UNLP, Profesor, UNSADA UNNOBA

⁴ Lic. en Informática, Profesor, UNLP, UNNOBA, UNSADA

⁵ Lic. en Sistemas, Profesor, UNNOBA, UNSADA

⁶ Analista Programador Universitario, Profesora, UNSADA

⁷ Alumno avanzado de la Carrera Lic. en Informática, UNSADA

⁸ Investigador externo UNNOBA

convocatoria de Subsidios a la Investigación 2022 ante la Secretaría de Investigación de la Universidad y aprobado mediante Resolución (CS) N° 302/2022. Su objetivo es continuar con la investigación y la formación de recursos humanos sobre los aspectos tecnológicos del desarrollo de aplicaciones utilizando tecnologías innovadoras en diferentes contextos. El proyecto está integrado por un equipo interdisciplinario de investigadores, docentes y estudiantes pertenecientes a la Escuela de Desarrollo Productivo y Tecnológico, y se trabajará en conjunto con la Escuela de Desarrollo Social y Humano de la UNSAdA.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las bases del proyecto se sientan en el proyecto que lo antecede, en el cual se trabajó en el desarrollo de la diagramación y planificación para la implementación de la educación digital en la UNSAdA. En este período se continuará trabajando con los fundamentos pedagógicos y didácticos incorporando herramientas tecnológicas emergentes a la educación a distancia y a la educación presencial. De esta forma se investigará cómo la informática impacta en el desarrollo de tecnologías innovadoras de manera de analizar, definir y desarrollar herramientas y estrategias innovadoras que impacten de manera responsable en el desarrollo de la educación y de la sociedad en general.

En los próximos años las tecnologías emergentes impactarán en el avance tecnológico y cobrarán importancia los desarrollos que apliquen dicha tecnología, en el proyecto se trabajará principalmente con:

Realidad Aumentada (RA): que, por definición, es el conjunto de tecnologías que permiten la superposición, en tiempo real, de imágenes, marcadores o

información generados virtualmente, sobre imágenes del mundo real. En la realidad aumentada el mundo real y el mundo virtual se entremezclan para crear una realidad mixta en tiempo real. Es un recurso tecnológico que ofrece experiencias interactivas al usuario a partir de la combinación entre la dimensión virtual y la física, con la utilización de dispositivos digitales. La RA se aplica en contextos y campos diversos como la medicina, la publicidad, el campo militar, los dispositivos de navegación, la prospección hidrológica y geológica, agropecuario, ambiental o el mundo industrial entre otros.

Realidad virtual (RV): que es un sistema tecnológico, basado en el empleo de distintos dispositivos, cuyo fin es producir una apariencia de realidad que permita al usuario tener la sensación de estar presente en ella. Su aplicación, aunque centrada inicialmente en el terreno de los videojuegos, se ha extendido a otros muchos campos, como la medicina, educación generando aulas virtuales o simulaciones de vuelo y laboratorios entre otros. La realidad virtual puede ser de dos tipos: inmersiva y no inmersiva. Los métodos inmersivos de realidad virtual con frecuencia se ligan a un ambiente tridimensional creado por una computadora, el cual se manipula a través de cascos, guantes u otros dispositivos que capturan la posición y rotación de diferentes partes del cuerpo humano. La realidad virtual no inmersiva también utiliza la computadora y se vale de medios como el que actualmente nos ofrece Internet, en el cual podemos interactuar en tiempo real con diferentes personas en espacios y ambientes que en realidad no existen sin la necesidad de dispositivos adicionales a la computadora, ofreciendo un nuevo mundo a través de una ventana de escritorio.

Inteligencia artificial (IA): que es la rama de las ciencias de la computación dedicada al desarrollo de agentes racionales no vivos,

entendiendo como agente a cualquier cosa capaz de percibir su entorno, procesar tales percepciones y actuar en su entorno (proporcionar salidas). La inteligencia artificial convencional se basa en el análisis del comportamiento humano ante diferentes problemas. Los sistemas de IA actualmente son parte de la rutina en campos como educación, economía, medicina, ingeniería y la milicia, y se ha usado en gran variedad de aplicaciones de software, juegos de estrategia como ajedrez de computador y otros videojuegos. Otros ejemplos de sus usos se encuentran en el área de control de sistemas, planificación automática, la habilidad de responder a diagnósticos y a consultas de los consumidores, reconocimiento de escritura, reconocimiento del habla y reconocimiento de patrones.

Entornos virtuales 3D: que implica que con la tecnología tridimensional se pueden realizar diferentes herramientas que las personas u organizaciones pueden involucrar en su día a día para brindar un mejor servicio o para facilitar algunas de sus necesidades, las cuales son, pantallas 3D, juegos en 3D, impresiones 3D y entornos virtuales 3D. Los entornos virtuales 3D o espacios de simulación 3D son sistemas inmersivos, interactivos, personalizables, accesibles y programables, que permiten diseñar actividades complejas.

RESULTADOS OBTENIDOS/ESPERADOS

Como se mencionó anteriormente, las bases del proyecto se sientan en el proyecto que lo antecede, en el cual se trabajó en el desarrollo de la diagramación y planificación para la implementación de la educación digital en la UNSAdA

Durante esa instancia se realizó la presentación del Sistema Institucional de Educación a Distancia (SIED) de la Universidad ante CONEAU obteniendo la

aprobación por parte del Ministerio A su vez, se construyó, mediante herramientas de software de código abierto, un Entorno Virtual de Enseñanza y Aprendizaje (EVEA) que integra los sistemas de información para lograr una administración y gestión académica de forma centralizada que permite crear, administrar y llevar a cabo aulas virtuales, herramientas de aprendizaje, comunicación y colaboración.

Se presentó una Becas EVC – CIN Convocatoria 2019, en la que durante el desarrollo de la misma, se relevaron estudios de las diferentes propuestas para el desarrollo del EVEA se vislumbraron las ventajas de utilizar un entorno basado en un LMS de Software Libre como lo es Moodle, alojado sobre una infraestructura *cloud* (en la nube) teniendo de manifiesto el contexto de la UNSAdA y su desarrollo tecnológico actual.

Se desarrolló una tesis de maestría en tecnología aplicada a la educación (UNLP) que actualmente se encuentra en evaluación para exposición.

Se busca en este sentido continuar con la investigación y la formación de recursos humanos sobre los aspectos tecnológicos del desarrollo de aplicaciones utilizando tecnologías innovadoras en diferentes contextos; y a su vez se pretende difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes congresos, jornadas y workshops de carácter nacional e internacional relacionados con el uso de las tecnologías emergentes.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por docentes e investigadores formados y en formación pertenecientes a la Universidad Nacional de San Antonio de Areco (UNSAAdA), algunos de los cuales dirigieron becas CIN, y diversos trabajos

finales de la carrera de Analista en Informática. Uno de los investigadores de este equipo defendió sus tesis de maestría obteniendo el título de Magíster en Tecnologías aplicadas a Educación otorgado por la Universidad Nacional de La Plata (UNLP). Otros de los docentes investigadores del equipo se encuentra actualmente esperando fecha para la defensa de su tesis también en el Magíster en Tecnologías aplicadas a Educación otorgado por la Universidad Nacional de La Plata (UNLP). Se postularon durante 2022 dos presentaciones de Becas CIN de alumnos avanzados y una beca CIC.

BIBLIOGRAFÍA

Campi, W. (2018) “La educación a distancia en Argentina a través de sus normas: de la Ley 1597/1885 a la Resolución Ministerial 2641-E/2017” en Dari, N. y Bauman, P. , Marcos regulatorios y modelos pedagógicos: un camino hacia la virtualización de la educación superior en el MERCOSUR, Bernal, Universidad Nacional de Quilmes

Lezcano, H. y Suárez, M. (2021). “Educación Superior, tecnologías digitales y educación a distancia: impacto en la comunidad de estudiantes de la Universidad Nacional de San Antonio de Areco en relación a la aplicación de los dispositivos virtuales implementados en la enseñanza” Tesis de grado de la Lic. en Gestión Educativa, UNSAdA.

López, S. (2018) “Modelos pedagógicos en la educación a distancia: el caso de la Universidad Nacional de Quilmes en Argentina” en Dari, N. y Bauman, P. Marcos regulatorios y modelos pedagógicos: un camino hacia la virtualización de la educación superior en el MERCOSUR, Bernal, Universidad Nacional de Quilmes.

Pontoriero, F. A. (2021). E-learning en la educación superior argentina - Modelo de evaluación de calidad a partir del aporte de referentes clave Virtualidad, Educación y Ciencia, 22 (12), pp. 22-45.

Solla, L. M. (2019). “Integración de sistemas de información institucionales y entornos virtuales de aprendizaje: ¿necesidad operativa o un componente ausente en el SIED?” En el 8º Congreso internacional de la Red Universitaria de Educación a Distancia (RUEDA 2019).

Belloch, C. (2012) Entornos virtuales de aprendizaje. Unidad de Tecnología Educativa (UTE). Universidad de Valencia. Disponible en https://www.uv.es/bellohc/pedagogia/EV_A3.pdf

Cabero, J., & Llorente, M. (2007). La interacción en el aprendizaje en red: uso de herramientas, elementos de análisis y posibilidades educativas. RIED. Revista Iberoamericana de Educación a Distancia, 10(2), 97-123. Recuperado de: <https://doi.org/10.5944/ried.2.10.995>

Cataldi, Z. Lage, F. J. Dominighini, C. (2013). Fundamentos para el uso de simulaciones en la enseñanza. Revista de Informática Educativa y Medios Audiovisuales. Vol. 10(17), págs.8-16, ISSN 1667-8338. Recuperado de: <http://laboratorios.fi.uba.ar/lie/Revista/Articulos/101017/A2mar2013.pdf>

Alfadatizando: innovación en el proceso de enseñanza aprendizaje en escuelas secundarias usando visualización de datos para desarrollar pensamiento computacional en materias de Humanidades y Ciencias Sociales

Alejandra B. Lliteras^{1,2}, Alejandro Artopoulos^{2,3,4}, Alejandro Fernandez^{1,2}, Jimena Huarte³

¹ UNLP, Facultad de Informática, LIFIA - ² CIPBA - ³ Escuela de Educación, Universidad de San Andrés - ⁴ Universidad de Buenos Aires

alejandra.lliteras@lifia.info.unlp.edu.ar, alepoulos@udesa.edu.ar,
alejandro.fernandez@lifia.info.unlp.edu.ar,

RESUMEN

El tránsito de la sociedad industrial a la informacional, de la mano del proceso de globalización y el desarrollo de las tecnologías de la información y la comunicación llevan a la necesidad de un cambio en la educación de los jóvenes. Adquirir habilidades de Pensamiento Computacional resulta de suma importancia para el siglo XXI y debe ser desarrollada desde la escuela. Si bien existen diversos esfuerzos por abordar esta habilidad desde la programación, resulta fundamental, tener una mirada integral de este tipo de pensamiento de manera transversal a las materias del curriculum y no como una acción aislada y relegada a la hora específica de TIC. Aún más, si bien se asocia este tipo de pensamiento a la acción de programar, es posible además desarrollarla desde la visualización de datos.

diferentes niveles educativos, abordando pensamiento computacional desde la visualización de datos y co-creando secuencias didácticas para su potencial uso en aprendizaje basado en proyectos y basado en acontecimientos, en donde, entre otras materias, se incluían algunas correspondientes a las ciencias sociales y humanidades. En abril de 2022 se defendió una tesina de grado [Bértora & Spaltro, 2022] relacionada a la visualización de datos en el ámbito educativo de secundario y en octubre de ese mismo año se presenta la primera versión de Alfadatizando [Lliteras et al., 2022] plataforma educativa que permite la gestión de actividades de visualización de datos para desarrollar pensamiento computacional en el nivel secundario.

En la actualidad se trabaja en la redefinición de la arquitectura, diseño y desarrollo de la plataforma Alfadatizando.

CONTEXTO

Finalizando el año 2020, se presenta la propuesta de doctorado llamada “Aprendo Sociales con Datos. Con el fin de construir conocimiento, en Ciencias Sociales, a partir del análisis de conjuntos de datos usando modelización, visualización y razonamiento algorítmico para escuelas secundarias de Argentina.” [Lliteras, 2020] a partir de ese momento, se dictaron cursos, a docentes de

1- INTRODUCCIÓN

El tránsito de la sociedad industrial a la informacional, de la mano del proceso de globalización y el desarrollo de las tecnologías de la información y la comunicación implican que “el mundo está cambiando alrededor de la escuela” [Halinen, 2014], lo que nos enfrenta a nuevos desafíos para un futuro sostenible. En este contexto, se requiere que los jóvenes conozcan el mundo digital [Kong & Abelson, 2022], incorporando nuevas habilidades. Esto

remarca la necesidad de renovar tanto lo que se enseña como también la manera en que se lo hace [Halinen, 2014].

En 2006, Wing [Wing, 2006] difunde el término Pensamiento Computacional e indica que este tipo de pensamiento es necesario para todos los individuos, no solo para quienes se dedican a las ciencias de la computación sino también para quienes se desempeñan en diferentes profesiones. Wing considera que este tipo de pensamiento debe ser incorporado a la enseñanza de todas aquellas disciplinas que requieren de habilidades analíticas. La autora destaca que el Pensamiento Computacional desarrolla las capacidades de abstracción y descomposición, lo que permite mejorar la resolución de problemas, el entendimiento del comportamiento humano y reformular un problema complejo en otro posible de resolver. Finalmente, este tipo de pensamiento complementa y combina pensamiento matemático e ingenieril y permite trabajar en diferentes niveles de abstracción

El Pensamiento Computacional (PeCo) es considerado actualmente una de las habilidades principales para el siglo XXI por lo tanto abordarlo desde el sistema educativo permitiría cerrar brechas sociales [Hazzan et al., 2020].

Además de verlo como una nueva habilidad que requieren los individuos hoy para su futuro desarrollo laboral, el Pensamiento Computacional es requerido para actividades cotidianas como la búsqueda de información en la web y la interpretación adecuada de los resultados obtenidos. Lo anterior nos indica que desarrollar la capacidad de búsqueda e interpretación de la información para integrarla y aplicar pensamiento crítico sobre los resultados lo que conlleva a pensar en la necesidad de un dialogo fluido entre diferentes disciplinas curricular, y aun más en la actualidad, donde plataformas que incluyen inteligencia artificial penetran una vez más los muros de la escuela, debiendo procurar las instituciones educativas, los medios para que

se conviertan en oportunidades de aprendizaje y no amenazas al sistema educativo

2- LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación presentada apunta al estudio de plataformas y tecnologías que permitan la innovación en el aula para favorecer el proceso de enseñanza aprendizaje en el marco del nivel secundario usando visualización de datos para desarrollar pensamiento computacional, considerando la importancia de los datos abiertos y la datificación [Mejias & Couldry, 2019] de la educación

3- RESULTADOS ESPERADOS/OBTENIDOS

El Pensamiento Computacional permite construir, pero además comprender aspectos complejos de las plataformas, siendo además una habilidad requerida en los jóvenes del Siglo XXI, pudiendo ser abordado en los ámbitos educativos (informales o formales para cualquiera de sus niveles) así como en espacios que permitan repensarse como profesionales.

En particular, los campos de la investigación en Humanidades Digitales (HD) requieren de nuevos skills que muchas veces no son brindados en las formaciones específicas, y que además no se condicen con la manera en la que los estudiantes secundarios aprenden las disciplinas de las humanidades en el colegio secundario.

Basados en la experiencia de una primera versión de plataforma educativa para desarrollar pensamiento computacional a partir de la visualización de datos, se trabajará en la mejora y ampliación de funcionalidades de esta. Se analizarán plataformas educativas y nuevas tecnologías que penetraron los muros

de las instituciones educativas con el fin de generar guías de uso y buenas prácticas para que puedan ser adoptadas por docentes en sus prácticas con alumnos, uno de los casos emblemáticos en la actualidad es el uso de ChatGPT¹ en diferentes materias de la curricula secundaria.

4- FORMACIÓN DE RECURSOS HUMANOS

La actualidad del tema, así como la diversidad de líneas de investigación previamente presentadas, presenta un tema de investigación concreto y relevante para la formación de recursos humanos, los cuales se espera trabajar tanto para tesinas de grado, trabajos de títulos intermedio y proyectos con alumnos.

5- BIBLIOGRAFÍA

[Bértora & Spaltro, 2022] Bértora, S., & Spaltro, F. (2022). Visualización web de datos para usuarios finales en entornos educativos Tesina de grado. UNLP, Facultad de Informática.

[Halinen, 2014] Halinen, I. (2014). General aspects of basic education curriculum reform 2016 Finland. Retrieved October, 4, 2016.

[Hazzan et al., 2020] Hazzan O., Ragonis N., Lapidot T. (2020) Computational Thinking. In: Guide to Teaching Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-030-39360-1_4

[Kong & Abelson, 2022] Kong, S. C., & Abelson, H. (Eds.). (2022). Computational Thinking Education in K-12: Artificial Intelligence Literacy and Physical Computing. MIT Press.

[Llitas et al., 2022] Llitas Alejandra, Artopoulos Alejandro, Fernandez Alejandro., & Huarte Jimena. (2022, October). AlfaDatizando: a Data Visualization Platform to work Computational Thinking in Digital Humanities. In 2022 XVII Latin American Conference on Learning Technologies (LACLO) (pp. 1-6). IEEE.

[Llitas, 2020] “Aprendo Sociales con Datos. Con el fin de construir conocimiento, en Ciencias Sociales, a partir del análisis de conjuntos de datos usando modelización, visualización y razonamiento algorítmico para escuelas secundarias de Argentina.”. Propuesta doctoral. UNLP. Facultad de Informática. Diciembre de 2020.

[Mejias & Couldry, 2019] Mejias, U. A., & Couldry, N. (2019). Datafication. Internet Policy Review, 8(4).

[Wing, 2006] Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.

¹ <https://openai.com/blog/chatgpt>

Autoras/es

Aballay, Alicia
Aballay, Laura
Abásolo, María José
Acciaio, Matías
Aciar, Silvana
Acosta, Denis
Acosta, Nelson
Acuña, Cesar
Ado, Mariana
Adorno, Sebastián
Agamennoni, Osvaldo
Agesilao, Agustín T.
Agostini, Federico
Aguado, Cintia
Aguil Mallea, Daniel
Aguilar, Norma
Aguirre, Juan Jose
Aguirre, Verónica
Ahmad, Tamara
Alanis, Mariana
Alario, Rocío
Alba, M. Daniela
Albarracín, Gonzalo
Albornoz, Enrique Marcelo
Albornoz, María Claudia
Alderete, Claudia
Alfonso, Hugo
Alfonzo, Pedro L.
Alonso, Nicolás
Alvarado, Y.
Alvarez, Diego
Alvarez, Eduardo
Alvarez, Leonel
Alvarez, Margarita
Alvez, Carlos
Amadeo, Ana Paola
Amaro, Silvia N.
Amigone, Federico
Angelucci, Tatiana
Antonelli, Leandro
Antonini, Antonela
Antonini, Antonella
Aparicio, Noelia
Aranda, Gabriela
Aranda, Marcos
Archuby, Federico
Arcidiacono, José
Argañaraz, Julián
Argento, facundo
Armana, Silvana V.
Armando, Silvana
Arroyuelo, Jorge
Artola, Eugenia C.
Artola, Verónica
Artopoulos, Alejandro
Asiain, Lucía
Asselborn, Miriam
Asteasuain, Fernando
Astudillo, Gustavo Javier
Ayala, Florencia
Azar, Paola
Baeza, Natalia
Baigorria, Lorena S.
Baiutti, Paola Beatriz

Baldassarri, Sandra
Balich, Franco
Balich, Nestor
Balladini, Javier
Ballesteros, Carlos
Balmaceda, Silvina
Banchoff Tzancoff, Claudia M.
Baquel, Ruben
Barba, Lautaro
Bareiro, Hernán
Barrera, María Alejandra
Barrionuevo, Mercedes
Basciano, Ivan
Basgall, María José
Bast, Silvia
Bateca, Carlos
Battaglia, Manuel
Battaglia, Nicolás
Bazán, Patricia
Bazterrica, Sebastián
Becerra, María del C.
Becerra, Martín Ezequiel
Bellino, Franco
Beltramini, Paola
Bencardino, Antonio
Benchoff, Delia Esther
Bendatti, Natalia
Benegas, Marcos
Benemerito, Ignacio
Benquerencia Mendez, Nicolás
Berini, Fabián
Bermúdez, Carlos
Bernal, Rubén Alfredo
Bernaldo de Quirós, Franco
Bernardis, Edgardo
Bernardis, Hernán
Bernoco, Cristian
Berón, Mario
Bertero, Regina
Bertoglio, Sebastián
Bertone, Roberto
Bertone, Rodolfo
Bessoni, Luciano
Bianchini, Germán
Bilbao, Martín
Biron, Martin Ferreyra
Bjerg, Ernesto
Blanco, Francisco
Bond, Román
Boracchia, Marcos
Borda, Emanuel
Borgna, Federico
Borgnia Giannini, Erik
Bossaro, Julio C.
Bourse, Francisco
Brancolini, Alessandro
Britos, Luis
Bruno, Laureano
Buccella, Agustina
Buckle, Carlos
Buffarini, Abril
Bustos, Diego
Busum Fradera, Matías
Cabrera, Andrea
Caceres, Katherine

Caffetti, Yanina Andrea
Calabrese, Julieta
Calderón, Sergio Leandro
Cambarieri, Mauro
Camele, Genaro
Candia, Agustín
Canessa, Carolina
Canessa, Luis
Cano, Eduardo
Cañibano, Rodrigo
Capdevila, Juan
Cappelletti, Marcelo
Cardenas, Marina E.
Cardozo, Claudia
Cardozo, Martín
Carini, Abril
Carnero, Mercedes
Carnuccio, Esteban A.
Carrizo, Claudio
Carruthers, Juan A.
Casalini, María Clara
Casanova, Belén
Casanova, Carlos
Casanova, Guillermo
Cassera, Valeria
Castelli, Alan
Castillo, David
Castillo, Julio
Castro, Alicia
Castro, Silvia
Catala, Nelson
Catania, Carlos
Cavaliere, Lucas
Caymes-Scutari, Paola
Ceballos, Darío
Cecchi, Laura
Cechich, Alejandra
Cedaro, Karina
Celador, Leonardo
Cenci, Karina M.
Chaglasian Sgang, Ayelen
Challiol, Cecilia
Chapperon, Gabriela
Charne, Javier
Chatterjee, Parag
Chávez, Edgar
Chavez, Susana
Chayle, Carolina I.
Chazarreta, Facundo
Chediak, Ernesto
Chichizola, Franco
Chuk, Oscar Daniel
Cicerchia, Lucas Benjamin
Cintora, Federico
Cipriano, Marcelo
Citate Gomez, Ignacio M.
Clavijo, Eugenio
Cocconi, Diego
Colato, German
Collado, Federico
Colombres, Mariela
Coma-Roselló, Teresa
Company, Ana María
Conchillo, Mariano
Contesti, Pablo

Conti, Laura
 Contreras, Matias
 Conturso, Gustavo
 Corbalán, Leonardo
 Corgatelli, Franco
 Correa, Martín Sebastián
 Corso, Cynthia
 Cortez, Alberto
 Costanzo, Manuel
 Cottal, Sergio
 Cristina, Federico
 Cruz, Marcos
 Cuña, Cintia
 Dal Bianco, Pedro
 Damiano, Luis Esteban
 Dangiolo, Federico
 Dapoto, Sebastián
 Dapozo, Gladys N.
 Dasso, Aristides
 De Giusti, Armando
 De Giusti, Laura
 De Vincenzi, Marcelo
 Decoud, Carla
 Del Do, Marcelo
 Del Do, Marcelo
 Del Gener, Aldana
 del Río, Nicolás
 Delfino, Hugo
 Delgado, Ciro
 Delia, Lisandro
 Delrieux, Claudio
 Deroche, A.
 Di Cicco, Carlos
 Di Felice, Martin
 Di Genaro, María E.
 Di Paola, Gonzalo
 Díaz Lapérgola, María Ayelén
 Díaz, Adrián
 Díaz, Gabriela
 Díaz, Javier
 Diedrichs Escudero, Ana Laura
 Diedrichs, Ana
 Dielschneider, M. Juliana
 Dieser, María Paula
 Dimase, Matías
 Ditz, Yanina
 Dogliotti, Mariano
 Dolz, Daniel
 Dorrego, Braian
 Dorzán, María Gisela
 Drazic, Leandro
 Duarte, David
 Dubinsky, Manuel
 Dufour, Fernando
 Durán, Elena
 Eguren, Sebastián Rodríguez
 Elgueta, Rodrigo Atilio
 Elkfury, Fernando
 Encinas, Diego
 Errecalde, Marcelo
 Escalante, Julián
 Escudero, Consuelo
 Esnaola, Leonardo
 Espíndola, María Cecilia
 Espinosa, Lucia Vilma
 Esponda, Silvia
 Estevez, Elsa

Estrebou, César
 Estrella, Micaela
 Etchart, Graciela
 Etcheverry, Pablo
 Eterovic, Jorge E.
 Fachal, Adriana
 Falappa, Marcelo A.
 Falcioni, Ludmila
 Fava, Laura
 Felipe, Cáceres
 Fernandez Bariviera, Aurelio
 Fernández Sosa, Juan
 Fernandez, Agustín
 Fernandez, Alejandro
 Fernandez, Erika
 Fernandez, J.
 Fernández, Juan Manuel
 Fernandez, Rocío
 Fernandez, Valentín
 Ferracutti, Gabriela
 Ferrara, Damian
 Ferrarini Oliver, Cintia
 Ferrarini, Cintia
 Ferraro, María de los Angeles
 Ferreira, Pablo
 Ferreyra, Gastón
 Ferreyra, Juan Pablo
 Figueroa, Karina
 Filippi, Jose Luis
 Fillottrani, Pablo Rubén
 Flores, Andrés Pablo
 Flores, Sebastián
 Flores, Sergio Rafael
 Fontana, Daniel
 Fontana, Eliana
 Forestiero, Camila
 Fraga, Hugo
 Frati, Emanuel
 Funes, Ana
 Furlani, Federico
 Futo, Gustavo
 Gaetan, Gabriela
 Gagliardi, Olinda Edilma
 Galarza, Brian
 Galdame, Elisa
 Gallardo, César
 Gallina, Sergio
 Gallo, Natalia Verónica
 Gallo, Silvana
 Ganga, Leonel
 Ganuza, María Luján
 Garay, Francisco
 García Krahm, Gustavo
 García, A.
 García, Edith
 García, Edith
 García, Sergio
 Gargano, Cecilia
 Gargiulo, Florencia
 Garí, Yisel
 Garrido, Nelson
 Gatica, Julieta
 Gaudiani, Adriana Angélica
 Gellon, Ivonne
 Gho, Edgardo
 Giardinelli, Agustín
 Gil Costa, Graciela Verónica

Gil, Esteban
 Gil, Fernando López
 Giorlando, Agustín
 Gira, Facundo Díaz
 Giro, Juan Matías
 Gnazzo, Gustavo Luis
 Godoy, Diego Alberto
 Godoy, Pablo Daniel
 Goin, Martin M. J.
 Gomez DOrazio, Lucas
 González de Doña, Mónica Gilda
 Gonzalez, Alejandro
 Gonzalez, Elian
 Gonzalez, Eugenia Alejandra
 Gonzalez, Jeremías
 Gonzalez, María Paula
 Gonzalez, Nahuel
 Gonzalez, Rodrigo
 Gramática, Martín Nicolás
 Greiner, Cristina
 Groller, Eduard
 Gronda, María Laura
 Grosso, Alejandro
 Guanco, Leonardo
 Guasch, María Mercedes
 Guerra Moronta, Jonathan
 Nicolás
 Guerra, Jorge
 Guerrero, Guillermo
 Guerrero, R.
 Guevara, Andrea
 Guiguet, Marcelo
 Gutierrez, Matías H.
 Guzmán, Jérica V.
 Harari, Ivana
 Hasperué, Waldo
 Hernandez, José Luis
 Hernandez, Nicolás
 Herrera Conegliano, Oscar A.
 Herrera De Rosa, Exequiel
 Herrera, Claudia M.
 Herrera, M.
 Herrera, Norma Edith
 Hnatiuk, Jair
 Huapaya, Constanza
 Huarte, Jimena
 Hurtado, Liliana
 Iannini, Juan
 Ibañez, Bárbara
 Ibañez, Eduardo
 Ierache, Jorge
 Iglesias, Gustavo
 Iglesias, Luciano
 Illescas, Gustavo
 Inda, Kevin Marcelo
 Irrazábal, Emanuel
 Isgró, Valentino
 Istvan, Romina Mariel
 Iuliano, Pablo J.
 Jara, Carlos
 Jara, Jimena
 Jason, Sofia
 Jaszczyszyn, Adrian
 Jofré, Mónica
 Jofré, Nicolás
 Kasián, Fernando
 Klenzi, Raúl Oscar

Kucuk, Lucas G.
 Kuna, Horacio
 La Red Martinez, David
 La Red Martinez, David L.
 Labella, Danilo
 Lacuesta, Gastón Axel
 Lafuente, Guillermo Javier
 Lafuente, Gustavo
 Lamas, C.
 Lanza-Castelli, Silvia
 Lanzarini, Laura
 Larrea, Martín
 Lasagna, Valeria
 Lattuca, Ana
 Lazarte, Ivana
 Lazurri, Guillermo
 Ledesma, Emiliano
 Ledesma, Viviana
 Lencina, Paula
 Leo, Rosana
 Lépez, Héctor
 Lescano, Germán Ezequiel
 Lewis, Mirtha
 Lezcano Airaldi, Andrea
 Libutti, Leandro
 Ligorria, Karina
 Lliteras, Alejandra B.
 Loiseau, Matías
 Lomoro, Jorge
 Loor, Fernando
 López, Carlos G.
 López, Juan J.
 López, Lidia
 Lopresti, Mariela
 Lopumo, Serenela
 Lovos, Edith
 Lucero, Sandra Viviana
 Ludueña, Veronica
 Luengo, Pablo
 Lugani, Carlos Fabian
 Luna, Adriana
 Luna, Diego
 Lund, María Inés
 Luque, Emilio
 Luque, Leandro
 Luzuriaga, Juan Manuel
 Maccallini, Lucas
 Macías, Patricio
 Maffei, Fabian
 Magris, Natalia Anahí
 Maidana, Carlos
 Maiorano, Ariel
 Maldonado, Andrea
 Maldonado, Calixto
 Maleh, F.
 Malvacio, Eduardo
 Manassero, Martín
 Mancilla, Matías
 Manganelli, Silvina
 Mansfeld, Santiago
 Manzaraz, Ana Karina
 Marecos Brizuela, Terecio Diosnel
 Marianetti, Osvaldo
 Marin Aranda, María Alejandra
 Mariño, Sonia
 Marrero, Luciano
 Martín, Adriana

Martín, Adriana C.
 Martín, Adriana E.
 Martin, Sofía
 Martinez Carod, Nadina
 Martínez, Carla
 Martínez, Carlos
 Martínez, Enrique
 Martínez, Malvina
 Martínez, Nicolas Garcia
 Martínez, Rodolfo
 Martínez, Roxana
 Martino, Lucio
 Mas, Fernando
 Masanet, María Isabel
 Masilla, Juan
 Masuet, Juan Pablo
 Matkovic, Kresimir
 Matteo, Lorena
 Mazalú, Rafaela
 Mazuran, Clara
 Mazzanti, Renato
 Medel, Ricardo
 Medina, Patricio
 Medina, Santiago
 Medina, Yanina
 Mellino, Nicolás
 Mendez Garabetti, Miguel
 Mendoza, Dora
 Menvielle, María Alejandra Paz
 Mercado, Gustavo
 Mercado, Viviana
 Merino, Mariano
 Micieli, Gustavo
 Migani, Silvina
 Minetti, Gabriela
 Mir, Sebastián
 Mirabete, M.
 Miranda, Natalia
 Molina, D.
 Molina, Maria Laura
 Molina, Maricel
 Monfroglio, Romeo Lorenzo
 Montejano Massa, José Pedro
 Montejano, Germán
 Montejano, German A.
 Montella, Yanina
 Montenegro Aguilar, G.
 Montezanti, Diego
 Monti, Marcos Adrián
 Moralejo, Raul
 Morales, Daniel Martín
 Morales, Martin
 Morales, Oswaldo W. N.
 Morán, Marina
 Moreno, Leandro
 Moreno, Marcelo
 Moyano, Ezequiel
 Moyano, Joaquín
 Moyano, José Hipólito
 Moyano, M. Emilia
 Muñoz, Rocío
 Murazzo, María
 Murazzo, María A.
 Naiouf, Marcelo
 Nasiff, Gabriel
 Navas, Guillermo Sergio
 Naveda, Claudia

Neil, Carlos
 Nicolía, Valeria
 Nobal, María Celeste
 Norscini, J.
 Novais, Paulo
 Nuñez, Gustavo Marcelo
 Ojeda, Juan
 Oliva, Facundo
 Olivares, Diego
 Olivera, Lucas
 Oloriz, Mario G.
 Olsowy, Verena
 Ontiveros, Patricia
 Orlandi, Cristina
 Ormeño, Emilio
 Orozco, Sergio
 Ortiz, Felipe
 Osés, Lucía
 Osio, Jorge R.
 Osorio, Alejandra
 Osycka, Liam
 Pacini, Elina
 Padilla, Franco
 Paez, Martin
 Pagano, Matias
 Pagnoni, Verónica Karina
 Palacio, Franco
 Palacios, Cristian Gabriel
 Palma, Marcelo
 Palmero, Pablo Rafael
 Pandolfi, Daniel
 Paniego, Juan Manuel
 Panizzi, Marisa
 Pantano, Juan Cruz
 Parada, Gonzalo
 Paredes, Rodrigo
 Parlanti, Tatiana Sofía
 Parra, Gerardo A.
 Parra, Susana
 Pasini, Ariel
 Pauletto, Ana C.
 Pavez, Eduardo
 Paz, Carla
 Pazos, Fernando
 Pedraza, Virginia
 Peliza, Carlos
 Pendent, Horacio
 Peña, Pablo
 Peralta, Mario
 Pereira, Maria Joao V.
 Perero, Lucrecia R.
 Perez Villar, Gustavo
 Perez, Gabriel
 Perez, Marisa
 Pérez, Santiago
 Pesado, Patricia
 Pessino, Mauro
 Petkoff Bankoff, Kristian
 Petris, Raquel H.
 Pezet, Braian
 Piccoli, María Fabiana
 Picotto, Ignacio
 Piergallini, Maria Rosana
 Pinto, Noelia
 Piray, Eduardo
 Pirondo, Franco
 Poblete, Claudia

Poch, Miguel
 Pollo Cattaneo, María Florencia
 Poncio, Silvia
 Ponte Ahon, Santiago
 Portugal, Juan Salvador
 Pousa, Adrián
 Printista, Marcela
 Procopio, Gastón
 Puig, Martin Pi
 Puricelli, Fernando
 Pytel, Pablo
 Quintana, Fabio
 Quintero, Franco
 Quiroga Marin, Ariel
 Quiroga, Facundo
 Quiroga, Ruben
 Ramón, H.
 Ramon, Hugo
 Ramón, Hugo
 Ramos, L.
 Rangel Henriques, Pedro
 Rasjido, José
 recabarren, Facundo
 Requena, Marcos
 Reus, Juan Marcelo
 Rexachs, Dolores
 Reyes Zambrano, Gary
 Reyes, Nora
 Reynoso, Luis
 Ricciardi, Ezequiel
 Riesco, Daniel
 Ríos, Gastón
 Ríos, Gastón I.
 Rivero, José
 Rivero, Julieta
 Robino, Luciano
 Rodríguez Medina, Carlos Gustavo
 Rodríguez Eguren, Sebastián
 Rodríguez, G.
 Rodríguez, Guillermo
 Rodríguez, Ismael
 Rodríguez, Jorge
 Rodríguez, Marina
 Rodríguez, Nelson
 Rodríguez, Rocío Andrea
 Roger, Sandra
 Roggero, Patricia
 Romagnano, María Rosalía Gema
 Romano, Gabriela
 Romera, Liliana
 Romero, Fernando
 Romero, Laura
 Romero, María Soledad
 Romero, Rubén
 Ronchetti, Franco
 Rosatto, Daniel
 Roseinstein, Javier
 Rosenstein, Javier
 Rosete-Suárez, Alejandro
 Rotella, Carina
 Rotter, M. José
 Ruano, Darío
 Rucci, Enzo
 Ruiz Díaz, Juan
 Russo, Claudia
 Sabelli, Agustín
 Saez, Silvia

Saizar, Victoria
 Salamanca, Andrés
 Salas, Milagros
 Salazar Mesía, Natalí
 Salazar, Nevelin Irene
 Saldaño, Viviana E.
 Saldarini, Javier
 Salgado, Carlos
 Salina, Mauro
 Salto, Carolina
 Salvatore, Juan E.
 Samat, Pablo
 Sanchez Arroyo, Fernando
 Sanchez, Alberto
 Sanchez, Mariano
 Sanchez, Viviana
 Santana, Sonia
 Santinelli, Paola
 Sanz, Cecilia
 Sanz, Diego Rubén
 Sanz, Victoria
 Sarmiento, Adriana
 Sarobe, Monica C.
 Sattolo, Iris
 Savoie, Ramiro
 Schab, Esteban
 Schiavoni, Alejandra
 Schiffino, Cristian
 Schpetter, Analía
 Schroeder, Romina
 Selva, Ricardo
 Selzer, Matias Nicolas
 Semino, Luciana
 Serafino, Sandra
 Serra, Ariel
 Serra, Silvio
 Serrano, Diego J.
 Sessa, Renzo
 Sifón, Ricardo
 Sigampa Paez, Elvio
 Silva Pavón, José Miguel
 Silvestri, Alejandro
 Silvestri, Valeria
 Simon, Agustin
 Sinde, Natalia
 Sinopoli, Alesio Esteban
 Soligo, Pablo
 Sosa, Hernan
 Spositto, Osvaldo Mario
 Stanchi, Oscar
 Straccia, Luciano
 Strappa, Jan
 Suarez, Omar
 Suarez, Santiago
 Suarez, Tomás
 Suppi, Remo
 Taffernaberr, Juan Carlos
 Taffernaberry, Carlos
 Tagarelli, Sandra
 Tanzola, Juan E.
 Tapella, Daniel
 Taranilla, María Teresa
 Taruschio, Marcelo
 Tejada, I. Millán
 Tesone, Fernando
 Tessore, Juan Pablo
 Thomas, Pablo

Tinetti, Fernando G.
 Tintenfich, Dylan
 Tissera, Pablo Cristian
 Tobar, Sebastián
 Tomaselli, Gabriela
 Torres, Juan Ignacio
 Torres, Luis
 Tortosa, Nicolás
 Trabes, Guillermo
 Trasmontana, Julio
 Trottini, Ana María
 tsiro, sebastian
 Urribarri, Dana K.
 Useglio, Gustavo
 Varela, Raúl
 Vargas Navarro, Rocío
 Vargas, Javier Leonardo
 Vazquez, Alejandro
 Vecchi, Norma
 Vega, Alejandro
 Vegega, Cinthia
 Venosa, Paula
 Vera, Andrea
 Vera, Cristina
 Vera, Pablo
 Verdejo, Ariel
 Verdún, Noelia
 Verino, Claudia
 Viadana, Alejandra
 Victoria, José
 Vila, S. Ledesma
 Vilaboa, Pablo
 Vilches, Diego
 Villagarcía Wanza, Horacio
 Villagra, A. M.
 Villagra, Andrea
 Villagra, Silvia
 Villagrán, Luis Daniel
 Villegas, Carolina
 Villodas, Martín
 Vincenzini, Cristian
 Violini, Lucía
 Vivas, Luis
 Vozella, Nicolás
 Welch, Daniel
 Yee, Cristina Manresa
 Zabala, Hernan
 Zacaria, Cintia
 Zaccardi, Gonzalo
 Zaldua, Analía Magdalena
 Zanellato, Claudio
 Zangara, María Alejandra
 Zanitti, María Ayelén
 Zárate Álvarez, Nicolás Ignacio
 Zarate, Marcos
 Zeballos, Martín
 Zeballos, Matías
 Zeligueta, Laura
 Zogbe, A. Sara
 Zogbe, Sara
 Zoratto, Valeria