

ACTAS DE LAS

# JORNADAS DE LA FACULTAD DE INFORMÁTICA

ACADÉMICA  
INVESTIGACIÓN  
EXTENSIÓN  
POSGRADO



FACULTAD DE INFORMÁTICA  
UNIVERSIDAD NACIONAL DEL COMAHUE

NOVIEMBRE 2023  
NEUQUÉN - ARGENTINA

Jornadas de la Facultad de Informática  
JFAI 2023  
6 al 10 de noviembre de 2023  
Neuquén, Patagonia, Argentina

Organiza:

Facultad de Informática  
Universidad Nacional del Comahue



**Universidad Nacional del Comahue**

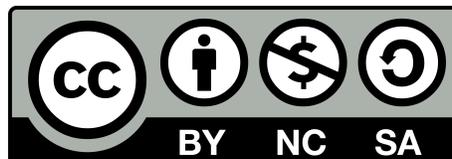
JFAI 2023 : Libro de Actas de las Jornadas de la Facultad de Informática ;  
Compilación de Gabriela N. Aranda ; Jorge P. Rodriguez. - 1a ed. - Neuquén:  
EDUCO - Universidad Nacional del Comahue. Editorial Universitaria del  
Comahue, 2024.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN: 978-987-604-653-4

1. Tecnología Educativa. 2. Actas de Congresos. I. Aranda, Gabriela N. , comp.  
II. Rodriguez, Jorge P. , comp.  
CDD 378.007





# **AUTORIDADES DE LA UNIVERSIDAD NACIONAL DEL COMAHUE**

## **Rectorado**

### **Rectora**

Dra. María Beatriz GENTILE

### **Vice-Rector**

Téc. Fernando Paul OSOVNIKAR

### **Secretario de Extensión**

Sr. Damián Andrés CANCELO

### **Director de la Editorial**

Lic. Enzo Dante CANALE

## **Facultad de Informática**

### **Decano**

Lic. Guillermo Alejandro GROSSO

### **Vice-Decana**

Lic. Silvia Nayum AMARO



## **Coordinadores JFAI 2023**

Dra. Gabriela N. Aranda  
Prof. Jorge Pablo Rodriguez

## **Comité Organizador JFAI 2023**

Prof. Susana Parra  
Lic. Silvia Amaro  
Lic. Ingrid Godoy  
Lic. Valeria Zoratto  
Tec. Romina Tealdi  
Lic. Micaela Malaspina  
Lic. Valeria Rotter  
AC. Luis Coralle  
Lic. Alejandro Mora  
APU. Federico Nahuel Mamani



# Índice

## *Proyectos de investigación*

### **Cómputo paralelo y distribuido: métricas de rendimiento, aplicaciones de big data e inteligencia artificial**

*Javier Balladini, Claudio Zanellato, Marina Morán, Rodrigo Cañibano, Agustín Chiarotto, Mariano Conchillo, Héctor Mella, Manuel Latorre, Cristina Orlandi*

2

### **04/F018: Calidad de información en comunidades virtuales**

*Gabriela N. Aranda, Valeria Zoratto, Nadina Martínez Carod, Andrés Flores, Natalia Baeza, Sandra Lucero, Lucas Cavaliere*

6

### **04/F019: Modelado de Variedad en Sistemas Big Data**

*Agustina Buccella, Alejandra Cechich*

10

### **Tecnologías Semánticas para el desarrollo de Agentes Inteligentes**

*Santiago Villarroel, Federico Solorza, Tomás Quiñonez, Matías Oyarzun, Guillermo Guerrero, Giuliano Marinelli, Christian Gimenez, Germán Braun, Sandra Roger, Laura Cecchi*

15

### **Computación Aplicada a las Ciencias y al Medio**

*Carina Fracchia, Pablo Kogan, Claudia Allan, Susana Parra, Hugo Contrera*

19

### **04/F022: Modelos formales, agentes inteligentes y aplicaciones para la enseñanza de las Ciencias de la Computación**

*Gerardo Parra, Jorge Rodríguez, Federico Amigone, Juan de la Fuente, Daniel Dolz, Guillermo Grosso, Daiana Bonet Peinado, Sandra Boari, Marcos Cortez, Ingrid Godoy*

23

### **Avances del Proyecto “Tecnologías de Datos Espaciales, Visualización y Realidad Virtual”**

*Luis Reynoso, Silvia Amaro, Lidia Lopez, Viviana Sanchez, María José Rotter, Sergio Cotal*

27

### **Robots de servicio con navegación autónoma en espacios públicos**

*Rafael Ignacio Zurita, Claudio Vaucheret, Miriam Lechner, Alejandro Mora, Luis Coralle, Ana Candelaria Alvarez*

34

## *Proyectos de extensión*

### **Programación Lógica en la Escuela Primaria**

*Laura A. Cecchi, Jorge P. Rodríguez, Valentina Villarroel, Ian Acosta, Lucas Margni, Emma Dolz, Guillermo Guerrero, Santiago Villarroel, Gabriela Aranda, Daniel Dolz, Alejandro Fierro, María Eda Cornejo, María Paz Sarrasqueta, Laura Parra, Alejandro Mora Vazquez, Valeria Roter, Micaela Malaspina, Federico Mamani, Romina Tealdi, Luis Coralle*

40

### **RA y RV aplicadas al resguardo del patrimonio cultural**

*Carina Fracchia, Amadeo Laurín*

44

## **FaIComm: Charlando Sobre Informática**

*Ian Acosta, Angel Avellaneda, Leo Bruno, Rodrigo Cañibano, Lara Galaz, Christian Gimenez, Pablo Kogan, Rafaela Mazalu, María Elvira Monserrat Vidal, Marina Morán, Jorge Rodríguez, Valeria Rotter, Valentina Villarroel*

47

## **Conectados con la Escuela**

*Ingrid Godoy, Susana Parra, Alejandro Antezana De La Rivera, Alejandro Jerson Mora Vazquez, Daniela Zacharias, Federico Nahuel Mamani, Ian Franco Matias Acosta, Ignacio Ciruzzi, Jonathan Maximiliano Cabrera, Jorge Pablo Rodríguez, Leo Bruno, Luis German Coralle, María Elvira Monserrat Vidal, Pablo Kogan, Rafael Ignacio Zurita, Rodrigo Ezequiel Piersigilli, Ruben Ramirez, and Valeria Zoratto*

51

## **Análisis forense de pericias médicos-legales asistidas por sistemas inteligentes: Preparación de Datos**

*Sandra Emilce Roger, German Alejandro Braun, Carlos Campos Fuentes, Manuel Latorre, Santiago Villarroel, Gabriel Omar Jerez, Cristian Martin Vincenzini, Fabio Marin Apablaza, Rafaela Jesus Maricela Mazalu, Romina Tealdi, Alejandro Jerson Mora Vazquez, Federico Solorza, Gabriela Noemi Aranda, Javier Ignacio Rey, Alejandro Daniel Cozzarin, Matias Ezequiel Oyarzun, and Luis Eduardo Garcia Carrillo*

55

## **ALERTAR: un sistema informático inteligente y resiliente para la vigilancia y detección temprana de severidad de pacientes en unidades de cuidados no intensivos**

*Javier Balladini, Claudio Zanellato, Marina Morán, Rodrigo Cañibano, Cristina Orlandi, Darío Semenzato, Mariano Conchillo, Héctor Mella, Manuel Latorre, Adriano Lusso, Dana Contreras, Aldana Vega, Lara Acuña, Santino Castagno, Cristian Ilabaca*

58

## **Reduciendo Brechas**

*Leo Bruno, Ian Acosta, Valentina Villarroel, Jonathan Maximiliano Cabrera, Valeria Rotter y Jorge Rodríguez*

62

## **Observatorio Electoral UNCo: Participación ciudadana en defensa de la democracia**

*Pablo Kogan, Daniela Zacharias, Christian Gimenez y Jorge Rodríguez*

66

## *Trabajos de estudiantes de grado y posgrado*

## **Observaciones al rol de la infraestructura tecnológica escolar en la educación en Ciencias de la Computación**

*Jonathan Maximiliano Cabrera*

71

## **Herramienta de diseño de aplicaciones Web para la consulta de programas Prolog**

*Santiago A. Villarroel, Christian Gimenez, Jorge P. Rodríguez, Laura A. Cecchi*

76

## **Clasificación de usuarios que preguntan en CQA**

*Adriana Mabel Cárdena, Gabriela N. Aranda, Valeria Zoratto*

80

## **Desarrollo de un entorno basado en programación en bloques para robots educativos**

*Juan Martín de la Fuente, Jorge P. Rodríguez, Laura A. Cecchi*

84

## **Modelado Bottom-up de Variedad de Dominio en Sistemas Big Data**

*Líam Osycka*

88

<b>Clasificación de posts en foros de discusión</b> <i>Jaime Daniel Saldivia Urra, Gabriela N. Aranda, Valeria Zoratto</i>	92
<b>Un enfoque para la unificación de UML 2.4.1, ORM 2 y EER</b> <i>Federico Solorza, Giuliano Marinelli, Laura A. Cecchi</i>	96
<b>Big Data en el Dominio Hidrológico: Caracterización del Contexto de Predicción de Turbidez</b> <i>Gastón Vidart</i>	100
<b>Modelos de generación de comentarios de código basados en transformers</b> <i>Cristian Martin Vincenzini, Sandra Roger</i>	104
<b>Especificación de Procesos para Preparación de Datos en Sistemas Big Data</b> <i>Sebastián Yaupe</i>	108
<b>Modelando Ontologías con Patrones en un Ambiente Gráfico Web de Ingeniería Ontológica</b> <i>Tomás Quiñonez, Christian Gimenez, Laura Cecchi</i>	112
<b>Ingeniería de reglas de asociaciones en Grafos de Conocimiento: Identificación de mejoras para lenguajes visuales</b> <i>Christian Gimenez, Germán Braun, Laura Cecchi</i>	115



ACTAS DE LAS

JORNADAS DE LA FACULTAD DE INFORMÁTICA



NOV 2023  
NEUQUÉN

FACULTAD DE INFORMÁTICA  
UNIVERSIDAD NACIONAL DEL COMAHUE

*Proyectos de investigación*

## Cómputo paralelo y distribuido: métricas de rendimiento, aplicaciones de big data e inteligencia artificial

Javier Balladini<sup>1</sup>, Claudio Zanellato<sup>1</sup>, Marina Morán<sup>1</sup>, Rodrigo Cañibano<sup>1</sup>, Agustín Chiarotto<sup>1</sup>, Mariano Conchillo<sup>1</sup>, Héctor Mella<sup>1</sup>, Manuel Latorre<sup>1</sup>, Cristina Orlandi<sup>2</sup>.

<sup>1</sup> Facultad de Informática, Universidad Nacional del Comahue  
javier.balladini@fi.uncoma.edu.ar

<sup>2</sup> Hospital Área Programa General Roca

**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 métricas de rendimiento. En particular, nos abocamos a métricas energéticas que cobraron enorme relevancia debido al gran número de unidades de procesamiento que componen los sistemas de cómputo. 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, basadas en inteligencia artificial. Una aplicación está destinada a unidades de cuidados intensivos, que debe tratar con datos masivos, y otra aplicación se orienta a unidades de cuidados no intensivos y presenta una arquitectura distribuida, tolerante a fallos y de simple instalación y mantenimiento. 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 grado, maestría y doctoral.

**Palabras claves:** computación de altas prestaciones, eficiencia energética, big data, sistema distribuido, salud, inteligencia artificial.

### 1 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, que involucra temáticas de inteligencia artificial y Big Data, 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.

## **2 Introducción**

Numerosas aplicaciones informáticas que requieren el procesamiento de grandes volúmenes de datos están surgiendo gracias al uso de tecnologías de cómputo paralelo y distribuido, que hasta hace unos años eran excepcionales. La masificación tecnológica y consecuente reducción de costos hizo que se pueda pensar en el desarrollo de nuevas aplicaciones que tengan un fuerte impacto social. En particular, nos interesamos principalmente en el sector de salud.

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 la eficiencia computacional. Actualmente, las métricas de rendimiento energéticas en muchos casos superan en relevancia a la velocidad de cómputo. En particular, nos hemos orientado a la gestión del consumo energético de estos sistemas, para reducir su impacto económico, medioambiental y social.

A continuación se introduce la motivación y problemática de ambas líneas de investigación.

### **2.1 Aplicaciones para la salud**

En una unidad de cuidados no intensivos, la metodología tradicional de atención consiste en la realización de un control de enfermería cada un determinado intervalo de tiempo fijo, idéntico para todos los pacientes. A su vez, los enfermeros no aplican un método estándar para determinar el nivel de gravedad de un paciente, que puede derivar en muertes inesperadas. Así, nos planteamos la siguiente pregunta: ¿Podemos organizar y optimizar los recursos, seleccionando y clasificando a los pacientes en diferentes niveles de riesgos para una correcta asignación de la atención?

Un problema similar ocurre en las unidades de cuidados intensivos, en donde la metodología tradicional es proclive a errores humanos, a detectar tarde el deterioro de los pacientes, y a la pérdida de datos entre registros de enfermería. En este caso, la pregunta que nos hacemos es: ¿Podemos crear un sistema de soporte a la decisión clínica basado en el procesamiento en tiempo real de señales fisiológicas y otros datos clínicos?

## **2.2 Métricas de rendimiento: eficiencia energética en HPC**

El incremento del número de unidades de procesamiento de los sistemas aumenta la probabilidad de fallos y el consumo energético. En este contexto, nos planteamos la siguiente pregunta de investigación: ¿Podemos predecir el consumo energético y ayudar a la toma de decisiones o sintonización del sistema para gestionar la energía ante la ocurrencia de fallos?

## **3 Líneas de investigación y resultados**

### **3.1 Aplicaciones para la salud**

Una de las aplicaciones que proponemos se denomina ALERTAR, que es una herramienta para la vigilancia y alerta temprana de gravedad en unidades de cuidados no intensivos. 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 instalar y mantener (no requiere el uso de servidores estándar en los hospitales). Mejoramos el diseño para superar problemas de seguridad, y para permitir la extensión de su funcionalidad, en especial para el soporte de múltiples enfermedades. Se realizaron pruebas de concepto del nuevo diseño. El objetivo en curso consiste en completar el nuevo diseño e implementarlo. El artículo [1] cubre parcialmente los avances.

En relación a las Unidades de Cuidados Intensivos, proponemos un sistema que 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. Hemos desarrollado un prototipo y estamos optimizando el sistema para aumentar su rendimiento y reducir los requerimientos de hardware. En el último año se terminó un software para análisis eficiente de la curva de electrocardiograma.

Además, se colaboró en el desarrollo de un modelo basado en agentes de la progresión de la enfermedad renal crónica, publicado en [2].

### 3.2 Métricas de rendimiento: eficiencia energética en HPC

Actualmente, nuestro principal objetivo es la gestión energética en mecanismos de tolerancia a fallos basados en checkpoints. Propusimos un modelo para estimar el consumo energético de operaciones de checkpoint y restart, y un método para su construcción. Ideamos 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 un modelo energético y un simulador que permite evaluar las estrategias, contemplando actualmente operaciones no bloqueantes y esperas en cascada causadas por dependencias entre procesos. El próximo objetivo es desarrollar un planificador para un sistema GNU/Linux que implemente el modelo. Los resultados se encuentran publicados en [3], y en el artículo [4] por aparecer.

## 4 Formación de recursos humanos

El equipo de trabajo tiene un Doctor y un Magíster, una estudiante de Doctorado en etapa de escritura de tesis, y un estudiante de Maestría. Además, hay 4 estudiantes de grado realizando su tesis de licenciatura. En 2023 se finalizó una tesis de licenciatura sobre el procesamiento eficiente de señales de electrocardiograma.

## Referencias

1. Rodrigo Cañibano, Santino Castagno, Mariano Conchillo, Guillermo Chiarotto, Claudia Rozas, Claudio Zanellato, Cristina Orlandi, Javier Balladini (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). <http://sedici.unlp.edu.ar/handle/10915/140661>
2. Candelaria Alvarez, Jose Ibeas, Javier Balladini, and Remo Suppi. Modelling Chronic Kidney Disease progression using ABM: a work in progress. In XI Conference on Cloud Computing, Big Data & Emerging Topics (La Plata, 2023). <http://sedici.unlp.edu.ar/handle/10915/155419>
3. 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 Conference on Cloud Computing, Big Data & Emerging Topics (La Plata, 2022). <http://sedici.unlp.edu.ar/handle/10915/140642>
4. Marina Morán, Javier Balladini, Enzo Rucci, Dolores Rexachs. Exploring Energy Saving Opportunities in Fault Tolerant HPC Systems. Journal of Parallel and Distributed Computing. Artículo aceptado, por aparecer.

## 04/F018: Calidad de información en comunidades virtuales

Gabriela N. Aranda, Valeria Zoratto, Nadina Martínez Carod  
Andrés Flores, Natalia Baeza, Sandra Lucero, Lucas Cavaliere

GIISCO Research Group  
Facultad de Informática - Universidad Nacional del Comahue  
Neuquen, Argentina  
{gabriela.aranda, valeria.zoratto}@fi.uncoma.edu.ar

**Abstract.** La Web 2.0 se destacó por su transformación de un mero repositorio de información en una plataforma de colaboración, dando lugar al desarrollo de servicios web, herramientas y aplicaciones que permitieron a los usuarios la creación y el intercambio de contenido en la web. Entre estos contenidos se incluyen blogs, wikis, redes sociales, así como plataformas para compartir videos y música en línea, prescindiendo de conocimientos técnicos especializados. Simultáneamente, surgieron diversas plataformas de redes sociales con objetivos diversos, abarcando desde la colaboración en entornos remotos hasta el intercambio de conocimiento técnico, académico y científico. Estas comunidades virtuales, cimentadas a través de la participación en línea, constituyeron un fenómeno de proporciones inéditas en la historia de la comunicación digital.

La Web 3.0 representa la subsiguiente fase en la evolución de la web, centrada en la creación de una web semántica. Este enfoque implica la organización y presentación de información de manera más inteligible y útil para los usuarios, priorizando la accesibilidad independientemente del dispositivo empleado, y mejorando la interacción del usuario con el fin de ofrecer resultados precisos y relevantes.

Como parte de esta transición hacia la Web 3.0, nuestro proyecto se orienta hacia la provisión de modelos de alta calidad para sistemas de software que faciliten la recuperación, análisis, clasificación y reutilización de información procedente de comunidades virtuales en la Web.

**Keywords:** Recuperación de información · Calidad de datos · Comunidades virtuales · Participación ciudadana

### 1 Sobre este proyecto

El proyecto de investigación “Reuso de información en comunidades virtuales”, de la Universidad Nacional del Comahue, tiene período de vigencia 2022-2025. Sus líneas de investigación extienden y avanzan temas desarrollados por su equipo de investigadores principales en proyectos pertenecientes al Programa “Desarrollo de Software basado en Reuso”, Parte I y II, desarrollados en la FAI entre 2013 y 2021.

En este 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 a la toma de decisiones basada en opinión pública.

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), conformado por docentes y estudiantes de la Facultad de Informática de la Universidad Nacional del Comahue, con la asesoría y colaboración de una investigadora 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 distintas 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.

## 2 Motivación

En 1993, Howard Rheingold introdujo el concepto de "comunidades virtuales" para describir colectivos sociales emergentes que se formaban a través de la interacción en línea en espacios públicos de Internet. Desde entonces, diversos entornos colaborativos han surgido, entre los cuales destacan las plataformas de redes sociales como Facebook, Twitter e Instagram, junto con las Comunidades de Preguntas y Respuestas (CQA, por sus siglas en inglés). Las CQA, similares a los foros de discusión, posibilitan a los usuarios buscar y compartir conocimientos a través de preguntas y respuestas, y han experimentado un crecimiento global, siendo ampliamente utilizadas por millones de usuarios a diario para abordar cuestiones complejas. Sin embargo, el aumento en la cantidad de información en estos sitios ha generado la necesidad de analizar y reutilizar esta información [4]. Por ende, la Integración y Reutilización de Información (IRI) desempeña un papel fundamental en la adquisición, representación, consolidación, validación y aplicación de datos para mejorar la toma de decisiones en diversos dominios de aplicación [2].

En este contexto se han realizado varias investigaciones con el fin de analizar, evaluar y extraer información de las CQA. Algunos estudios se han centrado en la evaluación de los hilos de discusión para definir criterios de calidad de las respuestas [6][1], mientras que otros han investigado las técnicas de respuesta a preguntas en las CQA, categorizándolas en sistemas basados en votación y sistemas basados en modelos [8][5]. En resumen, tanto el análisis del contenido textual de los mensajes como la investigación de las redes sociales subyacentes están intrínsecamente relacionados con este ámbito de estudio.

Además, el análisis de redes sociales en el contexto de las CQA implica la combinación de enfoques sociológicos y matemáticos. La aplicación de conceptos y operaciones de teoría de grafos permite explorar las estructuras sociales de estas redes, probar teoremas relacionados con los grafos subyacentes y someter declaraciones a pruebas [9], como por ejemplo en redes sociales de plataformas como Twitter. Sin embargo, debido a la diversidad de intereses y características de cada red, es esencial investigar el comportamiento de las personas en otras plataformas colaborativas, incluyendo la identificación de usuarios que generan información irrelevante, a los líderes de opinión, y también a los "boundary spanners" que facilitan la comunicación entre distintos grupos [7]. Esta exploración proporciona una comprensión más completa de las dinámicas sociales y de conocimiento en las CQA, y cómo estas pueden ser aprovechadas para mejorar la calidad de las respuestas y la experiencia de los usuarios.

Dada la abundante información generada por las CQA en la web, también se presenta un terreno propicio para el estudio y desarrollo de técnicas de reutilización e integración de datos [2]. Nuestro proyecto se enfoca entonces en definir métodos para la captura de información, análisis de contenido y detección y clasificación de perfiles de usuarios. Para su evaluación, se utilizan conjuntos de datos de comunidades virtuales existentes, como StackExchange (que alberga más de 173 foros de discusión sobre diversos temas y recibe mensualmente más de 100 millones de usuarios), cuyos datos están disponibles libremente para investigación. Además, el conocimiento adquirido puede ser aplicado en otras comunidades virtuales más específicas, como las creadas a partir de plataformas para la participación ciudadana, que han surgido recientemente en respuesta a los desafíos urbanos y la toma de decisiones. La participación ciudadana, a través de herramientas colaborativas como sitios web y aplicaciones móviles, se ha convertido en un medio fundamental para conocer las necesidades, demandas y perspectivas de los ciudadanos [3]. Por lo tanto, es relevante evaluar los productos existentes y adaptarlos para su aplicación en el contexto de comunidades de vecinos en la ciudad de Neuquén. Además, se pretende aprovechar las técnicas y herramientas desarrolladas en otras áreas de investigación para apoyar la toma de decisiones basada en la opinión de los ciudadanos.

### 3 Resultados obtenidos y esperados

Desde 2013, nuestro equipo de investigación ha estado trabajando en temas relacionados con las comunidades virtuales de programadores, proponiendo un modelo de calidad para foros de discusión técnicos y determinando criterios para evaluar la calidad de la información contenida en los hilos de discusión. También se ha avanzado en el procesamiento de texto de hilos de discusión, desarrollando una herramienta para recuperar información y analizarla según distintas medidas de calidad y hemos avanzado en la clasificación de roles de usuarios y propuesto un modelo de clasificación de perfiles de usuarios que preguntan.

Durante la ejecución del proyecto se han realizado las siguientes publicaciones:

- A study on influential features for predicting best answers in community question-answering forums. V. Zoratto, D. Godoy, G. Aranda, *Journal Information*. MDPI, Basel, Switzerland. *Information*, 14(9):496, ISSN 2078-2489. <https://doi.org/10.3390/info14090496>, Sep 2023.
- Calidad de información en comunidades virtuales. WICC 2023, Abr 2023, Universidad del Noroeste de la Provincia de Buenos Aires, Junín
- Reuso de información en comunidades virtuales. WICC 2022, Abr 2022, Universidad Champagnat, Mendoza. pp. 291-295

También se han desarrollado las siguientes tesis de licenciatura y se encuentra próxima a defender una tesis de doctorado.

- Clasificación de posts en foros de discusión. Tesista: Jaime Saldivia. Directoras: Gabriela Aranda y Valeria Zoratto. Junio 2023
- Clasificación de usuarios que preguntan en CQA. Tesista: Adriana Cárdena. Directoras: Gabriela Aranda y Valeria Zoratto. Junio 2023

## Referencias

1. Amancio, L., Dorneles, C.F., Dalip, D.H.: Recency and quality-based ranking question in cqas: A stack overflow case study. *Inf. Process. Manage.* **58**(4) (jul 2021). <https://doi.org/10.1016/j.ipm.2021.102552>, <https://doi.org/10.1016/j.ipm.2021.102552>
2. Day, M.Y., Ong, C.S., Hsu, W.L.: An analysis of research on information reuse and integration (2003-2008). *International Transactions on Systems Science and Applications* **6**(2), 146–157 (2010)
3. García Castro, D., Gutierrez, V.D.E., Kazak, J., Szezwanski, S., Kaczmarek, I., Wang, T.: 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)
4. Gottipati, S., Lo, D., Jiang, J.: Finding relevant answers in software forums. In: 2011 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011). pp. 323–332 (2011). <https://doi.org/10.1109/ASE.2011.6100069>
5. Khot, T., Sabharwal, A., Clark, P.: Answering complex questions using open information extraction. In: Barzilay, R., Kan, M.Y. (eds.) *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. pp. 311–316. Association for Computational Linguistics (2017). <https://doi.org/10.18653/v1/P17-2049>
6. Le, L.T., Shah, C., Choi, E.: Evaluating the Quality of Educational Answers in Community Question-Answering. In: *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*. pp. 129–138. JCDL '16, Association for Computing Machinery, New York, NY, USA (2016)
7. Matous, P., Wang, P.: External exposure, boundary-spanning, and opinion leadership in remote communities: A network experiment. *Social Networks* **56**, 10–22 (2019). <https://doi.org/https://doi.org/10.1016/j.socnet.2018.08.002>, <https://www.sciencedirect.com/science/article/pii/S0378873318301084>
8. Neshati, M.: On early detection of high voted qa on stack overflow. *Information Processing Management* **53**(4), 780–798 (2017)
9. Sanz-Menéndez, L.: Análisis de redes sociales: O cómo representar las estructuras sociales subyacentes. *Apuntes de Ciencia y Tecnología* **7**, 21–29 (06 2003)

# 04/F019: Modelado de Variedad en Sistemas Big Data

Agustina Buccella<sup>1</sup> and Alejandra Cechich<sup>1</sup>

GIISCO Research Group  
Departamento de Ingeniería de Sistemas - Facultad de Informática  
Universidad Nacional del Comahue  
Neuquen, Argentina  
`agustina.buccella,alejandra.cechich}@fi.uncoma.edu.ar`

**Abstract.** Un cambio importante con respecto a depósitos de datos tradicionales, es que en los Sistemas Big Data (SBDs) los datos pueden provenir de diferentes fuentes (como sensores, redes sociales, entorno etc.), y con diferentes formatos involucrando incluso datos desde totalmente estructurados a sin ninguna estructura definida. 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*.

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?*, nuestro proyecto propone modelar variedad a modo de líneas de productos. A diferencia de otras propuestas, la nuestra toma como partida una estructura de actividades asociadas al desarrollo de SBDs, instanciada en artefactos software producidos durante esas actividades e incorpora el modelado de variedades de manera similar a líneas de productos.

**Keywords:** Reusabilidad de Software · Variedad en Sistemas Big Data · Taxonomías

## 1 Introducción

Nuestra propuesta para modelar variedad en sistemas big data (SBDs) se ve influenciada por distintos tipos de diversidad: de *fuentes*, detectando distintas estructuras en los datos, técnicas de adquisición, etc.; de *contenido*, enfocando en las variables que son relevantes a un determinado problema de dominio; y de *procesamiento*, detectando variaciones en técnicas de análisis. Sin embargo, la variedad de los datos ha sido también considerada desde el punto de vista de incorporar semántica al proceso de modelado, incluso relacionando propiedades de calidad como interoperabilidad, seguridad, reusabilidad, etc. Entonces, en trabajos previos, hemos redefinido y extendido la clasificación de variedad para incluir la *variedad de contexto*, definiéndola como aquella que se refiere a *variaciones de dominio que pueden restringir o afectar los resultados del análisis*. Los diferentes tipos de contexto dependen de cada dominio; por ejemplo, para sistemas

hidrológicos, los cuerpos de agua pueden clasificarse en ríos, lagos, mares, etc., que pueden verse influenciados por la climatología, geología, u otros factores de entorno. Por lo tanto, estructurar conocimiento de dominio es una tarea esencial para detectar la variedad contextual.

## 2 Nuestro proceso para identificar variedad al modelar SBDs

Primeramente, hemos definido una aproximación de los elementos que componen nuestra arquitectura de referencia para SBDs basada en reuso. En ella, los aspectos de negocios (dominio), aplicación (software y análisis) y tecnológicos se abordan en niveles separados; siendo transversales aspectos como el uso/reuso de estándares, taxonomías y conocimiento.

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. Así, además de incluir a los participantes del desarrollo del SBD, involucra los requerimientos del proyecto y del dominio, restricciones, modelos y casos de uso. Es importante resaltar que estos activos deben generarse a partir de *taxonomías de dominio y estándares* y 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).

El proceso de identificación de variedad en SBDs puede iniciarse desde la evidencia que presenten los datos en datasets (enfoque bottom-up); desde el análisis conceptual del dominio (enfoque top-down); o con una mezcla de ambos casos. En diversas publicaciones esos enfoques se discuten con cada uno de sus pasos, incluyendo la caracterización de dominio que deriva en la construcción de artefactos mediante los cuales se resguardan variables y relaciones relevantes para describir contextos similares. Esos artefactos son gestionados por una herramienta de soporte (CoVaMaT), que permite almacenarlos con una estructura que facilita su recuperación y reuso.

La Figura 1, muestra los dos enfoques paso a paso:

- Enfoque top-down (T-VIP): en este enfoque, a la izquierda de la figura, (1) dado un problema de dominio, el usuario experto elabora una o más hipótesis que deberán ser probadas mediante análisis de datos (¿los datos lo corroboran?); (2) luego, en base a esas hipótesis, analistas de datos proceden a realizar las tareas del flujo de procesamiento (3); finalmente, los resultados permiten verificar las hipótesis (4) posiblemente visualizando los datos de diferentes maneras, incluso permitiendo reformulaciones (5).
- Enfoque bottom-up (B-VIP): en este enfoque, a la derecha de la figura, (1) dado un problema de dominio, el usuario experto decide iniciar un estudio exploratorio para averiguar lo que los datos pueden revelar sobre ese problema (¿qué dicen los datos?); (2) luego, el estudio es llevado a cabo por analistas

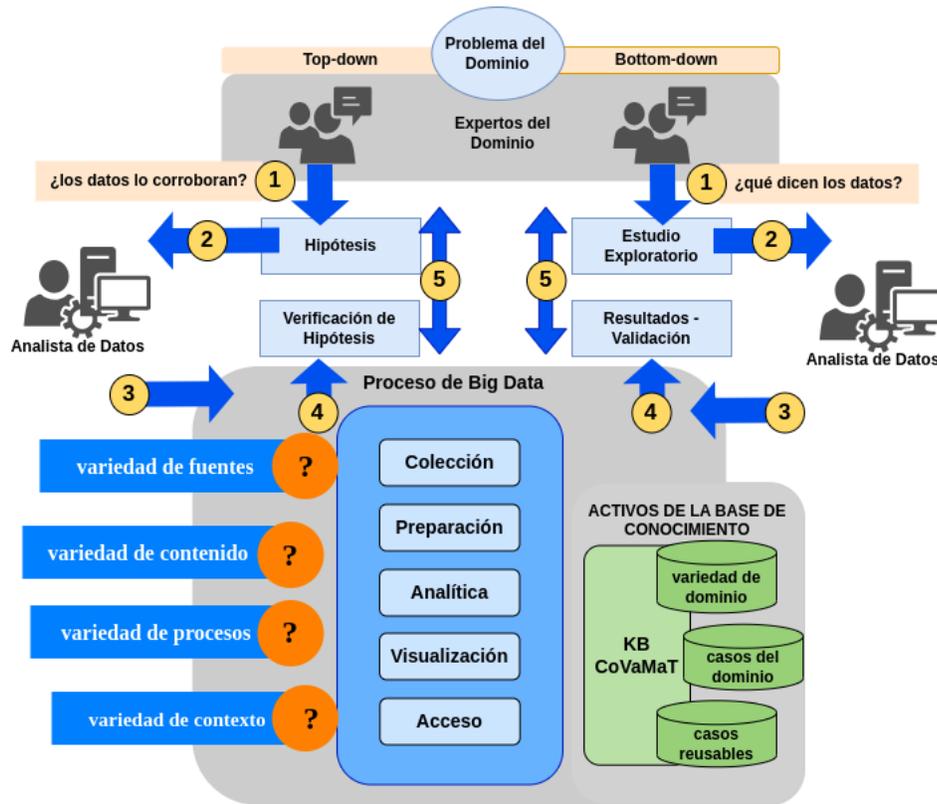


Fig. 1. Enfoques T-VIP y B-VIP para identificación de variedad

de datos aplicando las actividades del flujo de procesamiento (3); finalmente, los datos retornan (4) para ser validados por expertos, finalizando el proceso o reiniciando la búsqueda (5).

La validación del proceso, así como de los modelos y el reuso, se realiza en colaboración con INTA Alto Valle<sup>1</sup>, en el dominio de agricultura inteligente. Así, por ejemplo, se ha aplicado T-VIP para el análisis de variaciones de la napa freática en función de variables meteorológicas; o en el análisis de factores que afectan el índice NDVI, que describe la salud de la vegetación al medir la diferencia entre la luz visible (lo que la vegetación refleja) y la luz infrarroja, entre otros.

### 3 Trabajo Futuro

Como trabajo futuro se planea continuar trabajando en el desarrollo de casos de estudio en el dominio de la agricultura bajo alguno de los enfoques propuestos,

<sup>1</sup> <https://inta.gob.ar/altovalle/sobre-812000>

T-VIP y/o B-VIP, en donde se presente variedad de contexto; nutriendo así los repositorios de activos del dominio reusables.

#### 4 Integrantes, Colaboraciones y Publicaciones

El proyecto esta integrado por 11 docentes-investigadores, 6 alumnos de la Facultad, y 2 colaboradoras expertas del dominio de aplicación, específicamente pertenecientes al INTA. A su vez, el proyecto cuenta actualmente con tres doctores y un magister.

Dentro de los 2 años de ejecución del proyecto de investigación hemos realizado las siguientes publicaciones en las siguientes conferencias/congresos y revistas:

- **(2023) CoVaMaT: Functionality for Variety Reuse Through a Supporting Tool.** *Osycka, L., Cechich, A., Buccella, A.* In: Cloud Computing, Big Data & Emerging Topics - 11th Conference, JCC-BD&ET. Communications in Computer and Information Science, volume 1828, pag. 57–74, Springer, [https://doi.org/10.1007/978-3-031-40942-4\\_5](https://doi.org/10.1007/978-3-031-40942-4_5), ISBN= 978-3-031-40941-7.
- **(2023) A Model of Reusable Assets in AIE Software Systems.** *Cechich, A., Buccella, A., Villegas, C., Montenegro, A., Munoz, A., Rodriguez, A.* Journal of Computer Science & Technology 2023, 23(2). <https://doi.org/10.24215/16666038.23.e13>
- **(2023) Identificación Top-Down de Variedad de Contexto: Un Caso de Estudio en Fluctuaciones de la Napa Freática** *M. Campetella, A. Cechich, A. Buccella, A. Montenegro, A. Munoz, A. Rodriguez* . CACIC 2023: XXIX Congreso Argentino de Ciencias de la Computación, Lujan, Argentina. 09-12 de Octubre.
- **(2022) Data Variety Modeling: A Case of Contextual Diversity Identification from a Bottom-up Perspective.** *Osycka, L., Cechich, A., Buccella, A.* In: Pesado, P., Gil, G. (eds) Computer Science - CACIC 2021. CACIC 2021. Communications in Computer and Information Science, vol 1584. Communications in Computer and Information Science, vol 1584. Springer, Cham. [https://doi.org/10.1007/978-3-031-05903-2\\_9](https://doi.org/10.1007/978-3-031-05903-2_9) (124-138). ISBN: 978-3-031-05902-5
- **(2022) Un Proceso de Big Data aplicado a datos de siembra, cosecha, producción y rendimiento del girasol** *M. Conchillo, A. Buccella.* CONAI-ISI 2022: 10mo Congreso Nacional de Ingeniería Informática/Sistemas de Información. Trabajos Estudiantiles. Facultad Regional Concepción del Uruguay de la UTN. 03-04 de Noviembre.
- **(2022) Un Proceso de Big Data aplicado al transporte con tarjetas SUBE** *F. Saurin, J. Polo Sosa, A. Buccella.* CONAIISI 2022: 10mo Congreso Nacional de Ingeniería Informática/Sistemas de Información. Trabajos Estudiantiles. Facultad Regional Concepción del Uruguay de la UTN. 03-04 de Noviembre.

- **(2022) Caracterización de Variables para el Análisis del Índice de Vegetación** *C. Villegas, A. Buccella, A. Cechich*. CACIC 2022: XXVIII Congreso Argentino en Ciencias de la Computación, La Rioja, Argentina. 03-06 de Octubre.
- (2022) Análisis de Turbidez basado en Caracterización de Contextos *G. Vidart, A. Cechich, A. Buccella, A. Montenegro*. Congreso Argentino de AgroInformática (CAI) de las 51<sup>o</sup> Jornadas Argentinas de Informática, Buenos Aires, Argentina. Octubre 2022.
- **(2022) Modelado de Variedad de Activos de Dominio en Sistemas Big Data** *A. Buccella, J. Luzuriaga, A. Cechich, L. Osycka, C. Villegas, M. Cruz, F. Corgatelli, R. Martínez, R. Mazalu, M. Moyano*. XXIV Workshop de Investigadores en Ciencias de la Computación, Mendoza, Argentina. Abril 2022.
- **(2021) Identificación de variedad contextual en modelado de sistemas big data** *L. Osycka, A. Buccella and A. Cechich*. XXVII Congreso Argentino en Ciencias de la Computación, Modalidad Virtual. Octubre 2021.(367-376). ISBN 978-987-633-574-4
- **(2021) Reusabilidad en el contexto de Desarrollo de Sistemas para Big Data**. *A. Buccella, J. Luzuriaga, A. Cechich, L. Osycka, F. Paterno, M. Pol'la, M. Cruz, R. Martínez, R. Mazalu, M. Moyano*. XXIII Workshop de Investigadores en Ciencias de la Computación, La Rioja, WICC, Argentina. 15 y 16 de Abril. ISBN 978-987-24611-4-0
- **(2021) Decisiones de Diseño en un Proceso de Big Data para la Predicción de Temperatura del Agua** *L. Osycka, G. Siculo, A. Buccella, A. Cechich*. CONAISI 2021: 9no Congreso Nacional de Ingeniería Informática/Sistemas de Información. Trabajos Estudiantiles. Facultad Regional Mendoza de la UTN. Noviembre.
- **(2020) Un Proceso de Big Data aplicado a datos del COVID-19** *G. Canevello, N. Saddi, G. Vidart, C. Villegas, S. Yaupe, A. Buccella*. CONAISI 2020: 8vo Congreso Nacional de Ingeniería Informática/Sistemas de Información. Trabajos Estudiantiles. Universidad Tecnológica Nacional. Facultad Regional San Francisco, Argentina. 5-6 de Noviembre. ISBN 978-950-42-0202-8

Al mismo tiempo se han realizado las siguientes Tesis de la Carrera de Licenciatura en Sistemas de Información, en el marco del proyecto:

- Especificación de Procesos para Preparación de Datos en Sistemas Big Data. Alumno: Sebastian Yaupe. Mayo 2023. Directora: Alejandra Cechich
- Big Data en el Dominio Hidrológico: Caracterización del Contexto de Predicción de Turbidez. Alumno: Gastón Vidart. Noviembre de 2022. Directora: Alejandra Cechich
- Modelado Bottom-up de Variedad de Dominio en Sistemas Big Data. Alumno: Lám Osycka. Noviembre 2022. Directora: Alejandra Cechich

# Tecnologías Semánticas para el desarrollo de Agentes Inteligentes

Santiago A. Villarroel<sup>1</sup>, Federico Solorza<sup>1</sup>, Tomás L. Quiñonez<sup>1</sup>, Matías Oyarzun<sup>1</sup>, Guillermo Guerrero<sup>1</sup>, Giuliano Marinelli<sup>1,2</sup>, Christian N. Gimenez<sup>1</sup>, Germán A. Braun<sup>1</sup>, Sandra E. Roger<sup>1</sup>, and Laura A. Cecchi<sup>1</sup>

<sup>1</sup> Grupo de Investigación en Lenguaje e Inteligencia Artificial  
Facultad de Informática  
Universidad Nacional del Comahue

<sup>2</sup> CONICET

{santiago.villarroel,federico.solorza,tomas.quinonez,matias.oyarzun,  
guillermo.guerrero}@est.fi.uncoma.edu.ar {giuliano.marinelli,  
christian.gimenez,german.braun,roger,lcecchi}@fi.uncoma.edu.ar

**Palabras Clave:** Agentes Inteligentes - Inteligencia Artificial - Programación Lógica - Web Semántica - Ontologías - Grafos de Conocimiento - Procesamiento de Lenguaje Natural - Aplicaciones para la Enseñanza de Ciencias de la Computación

## 1. Objetivos

Este proyecto de investigación persigue un doble objetivo general. En primer lugar, generar conocimiento especializado en el área de agentes inteligentes que accedan, procesen y recuperen información mediados por tecnologías semánticas. En este sentido, se propone estudiar técnicas de representación de conocimiento y razonamiento, tecnologías del lenguaje natural, metodologías para la interoperabilidad e integración de datos, y generar nuevos principios, metodologías formales y herramientas basadas en la gestión semántica de los datos.

El segundo objetivo, complementario al anterior, es la formación y consolidación de recursos humanos. Esto será llevado a cabo a través de becas de investigación y perfeccionamiento, como así también de la realización de post-gradados en Ciencias de la Computación.

## 2. Resultados Alcanzados

En el transcurso de los años 2022 y 2023 se ha desarrollado diversos trabajos de investigación:

- Modelos de generación de comentarios de código basados en transformers: Este trabajo aborda el desafío de generar automáticamente comentarios claros y relevantes para describir fragmentos de código, utilizando un modelo de aprendizaje automático basado en la arquitectura transformer. Los resultados revelan que el modelo especializado en el lenguaje GO superó al modelo

de referencia, y aunque el rendimiento del modelo PROLOG fue inferior, una métrica adicional de similitud semántica demostró que las descripciones generadas mantienen una alta semejanza con las referencias humanas[14,15,16].

- Un lenguaje visual unificador de EER, UML y ORM 2: Este trabajo desarrolla un lenguaje visual para el manejo del metamodelo KF, considerando aspectos fenomenológicos, semióticos, ontológicos y cognitivos para el diseño de sus símbolos. El nuevo lenguaje visual busca facilitar la comunicación eficiente y convencional en el diseño de ontologías. Se presenta una representación formal parcial a través de la lógica descriptiva[13].
- Un marco para la interoperabilidad entre modelos con herramientas híbridas: El desarrollo y mantenimiento de sistemas complejos enfrenta el desafío de lidiar con diferentes tipos de modelos debido a las posibilidades del lenguaje, preferencias, tamaños, etc., que implican la interacción entre usuarios con diferentes niveles de competencia. Proponemos FaCIL, un marco para herramientas de modelado híbrido que apoya la capacidad de representar modelos en diferentes formatos mientras se preserva su semántica. FaCIL estructura y facilita la interacción entre modelos conceptuales visuales y textuales, sus especificaciones formales y abstracciones, así como el seguimiento y propagación de actualizaciones en todas las representaciones[1].
- Modelando Ontologías con Patrones en un Ambiente Gráfico Web de Ingeniería Ontológica: Este trabajo define y diseña una metodología para el modelado de ontologías basado en patrones, que sugiere Patrones de Diseño Ontológico de Contenido (CODP) a partir del modelo en desarrollo. Además, se desarrolló un prototipo de herramienta, denominado ODPR-API, que implementa dicha metodología, recomendando automáticamente CODP que pueden ser aplicados en una ontología, con el objetivo de acelerar el proceso de desarrollo de ontologías[12,11,10].
- Ingeniería de reglas de asociaciones en Grafos de Conocimiento. Identificación de mejoras para lenguajes visuales: Este trabajo propone una metodología para detectar deficiencias en un lenguaje visual para reglas de mapeo y sugerir modificaciones para una mayor efectividad cognitiva. La metodología se aplica a un lenguaje visual específico para mostrar un ejemplo concreto de aplicación, revelando varias formas de modificar el lenguaje para mejorar su eficiencia cognitiva[5].
- Datos observacionales/hidrográficos del Océano Atlántico Sur publicados como LOD: Este trabajo describe la publicación de datos de observación y mediciones hidrográficas de Elefantes Marinos del Sur como LOD, proporcionando un recurso accesible para la comunidad internacional[17].
- Question Answering aplicado a la web semántica: predicción de la respuesta esperada. Este plan se centra en la investigación y desarrollo de soluciones para sistemas de Búsqueda de Respuestas (Question Answering - QA), con el objetivo de dar respuestas correctas a preguntas formuladas en lenguaje natural. Se busca realizar una clasificación correcta del tipo de pregunta para optimizar el proceso de búsqueda, y luego localizar, extraer y presentar la información deseada al usuario[8,9].

- Enseñanza de Prolog: Estos trabajos argumentan que la Programación Lógica es una opción viable y deseable para niños de escuela primaria para introducir la programación y desarrollar habilidades de Pensamiento Lógico y Computacional[2,3].
- Plataforma para la Enseñanza de la Programación basada en Eye Tracking: Este trabajo presenta la definición y diseño de una plataforma para la enseñanza de programación que utiliza eye tracking como principal mecanismo de interacción, destinada a estudiantes con dificultades motrices. El objetivo es reducir las brechas en el aprendizaje de programación, permitiendo a estos estudiantes desarrollar, modificar y ejecutar programas usando la vista como principal medio de interacción[7,6].
- Desarrollo de un entorno basado en programación en bloques para robots educativos: en este trabajo se definió un lenguaje formal de programación por bloques y diseñó una herramienta web compatible con la plataforma de robótica educativa FrankLab. Como resultado, se ha desarrollado un entorno de programación por bloques que permite a los usuarios manipular los robots, observar su visión, guardar y compartir programas, y obtener una traducción a un lenguaje textual compatible con Frankestito[4].

### 3. Líneas Futuras de Investigación

Entre las líneas de investigación futuras se encuentran:

- Aplicar la metodología propuesta sobre los lenguajes visuales para reglas de mapeo, y evaluar el impacto de las modificaciones sugeridas en la eficiencia cognitiva de los usuarios.
- Desarrollar recursos educativos y estrategias pedagógicas adicionales para facilitar enseñanza de la programación y particularmente Prolog en todos los niveles educativos. Evaluar la eficacia de las plataformas educativas desarrolladas
- Investigar y desarrollar estrategias de clasificación, predicción y generación de contenido textual a partir de imágenes forenses. Estas estrategias tienen como finalidad proporcionar respaldo y facilitar la elaboración de informes periciales para el cuerpo médico forense.
- Investigar sobre metalenguajes para el modelado conceptual con el objetivo de interactuar de manera inteligente con lenguajes lógicos y ontologías. Diseñar e implementar herramientas para lograr un modelado conceptual inteligente basado en los metalenguajes definidos.

### Referencias

1. Braun, G., Fillotrani, P.R., Keet, C.M.: A framework for interoperability between models with hybrid tools. *Journal of Intelligent Information Systems* **60**(2), 437–462 (2023)

2. Cecchi, L.A., Rodríguez, J.P., Dahl, V.: Logic programming at elementary school: Why, what and how should we teach logic programming to children? In: Prolog: The Next 50 Years, pp. 131–143. Springer (2023)
3. Dahl, V., Cecchi, L.A.: Introducing prolog in language-informed ways. In: Prolog: The Next 50 Years, pp. 154–167. Springer (2023)
4. Fuente, J.M.D.L., Rodríguez, J.P., Cecchi, L.A.: Tesis de Licenciatura en Ciencias de la Computación: Desarrollo de un entorno basado en programación en bloques para robots educativos. Tech. rep., Facultad de Informática - Universidad Nacional del Comahue (2022)
5. Gimenez, C., Braun, G., Cecchi, L.A.: Tesis de Magister en Ciencias de la Computación(en desarrollo): Ingeniería de reglas de asociaciones en Grafos de Conocimiento: Identificación de mejoras para lenguajes visuales. Tech. rep., Facultad de Informática - Universidad Nacional del Comahue (2023)
6. Guerrero, G., Rodríguez, J., Cecchi, L.: EYE TRACKING PARA LA ENSEÑANZA DE LA PROGRAMACIÓN. In: XXV Workshop de Investigadores en Ciencias de la Computación (WICC 2023, UNNOBA, Junín) (2023)
7. Guerrero, G.A., Rodríguez, J.P., Cecchi, L.A.: Plataforma para la Enseñanza de la Programación basada en Eye Tracking. In: Actas de las IX Jornadas Iberoamericanas de Interacción Humano-Computadora (2023)
8. Oyarzun, M., Roger, S.: Question aswering aplicado a la web semántica: predicción de la respuesta esperada. In: XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza) (2022)
9. Oyarzun, M., Roger, S.: Predicción de la respuesta en un sistema de búsqueda de respuesta semántico. In: XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)(La Rioja, 3 al 6 de octubre de 2022) (2023)
10. Quiñonez, T., Giménez, C., Cecchi, L., Fillottrani, P.R.: Hacia la recomendación automática de patrones de diseño ontológico. In: XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)(La Rioja, 3 al 6 de octubre de 2022) (2023)
11. Quiñonez, T.L., Giménez, C.N., Cecchi, L., Fillottrani, P.R.: Modelando ontologías con patrones en un ambiente gráfico web de ingeniería ontológica. In: XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza) (2022)
12. Quiñonez, T.L., Gimenez, C.N., Cecchi, L.A.: Tesis de Licenciatura en Ciencias de la Computación: Modelando Ontologías con Patrones en un Ambiente Gráfico Web de Ingeniería Ontológica. Tech. rep., Facultad de Informática - Universidad Nacional del Comahue (2023)
13. Solorza, F., Marinelli, G., Cecchi, L.A.: Tesis de Licenciatura en Ciencias de la Computación: Un lenguaje visual unificador de EER, UML y ORM 2. Tech. rep., Facultad de Informática - Universidad Nacional del Comahue (2023)
14. Vincenzini, C., Roger, S.: Generación de comentarios a partir de código fuente utilizando transformers. In: XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)(La Rioja, 3 al 6 de octubre de 2022) (2022)
15. Vincenzini, C., Roger, S.: Tecnologías Semánticas para el desarrollo de Agentes Inteligentes: Generación de comentarios a partir de código fuente. In: XXV Workshop de Investigadores en Ciencias de la Computación (WICC 2023, UNNOBA, Junín) (2023)
16. Vincenzini, C.M., Roger, S.E.: Tesis de Licenciatura en Ciencias de la Computación: Modelos de generación de comentarios de código basados en transformers. Tech. rep., Facultad de Informática - Universidad Nacional del Comahue (2023)
17. Zárate, M., Braun, G., Lewis, M., Fillottrani, P.: Observational/hydrographic data of the south atlantic ocean published as lod. *Semantic Web* **13**(2), 133–145 (2022)

## COMPUTACIÓN APLICADA A LAS CIENCIAS Y AL MEDIO

Carina Fracchia<sup>1</sup>, Pablo Kogan, Claudia Allan, Susana Parra,  
Hugo Contrera  
Facultad de Informática  
Universidad Nacional del Comahue  
{carina.fracchia, pablo.kogan, claudia.allan, susana.parra}@fi.uncoma.edu.ar,  
hac.nqn@gmail.com

**Resumen.** Los avances de las Tecnologías de la Información y la Comunicación (TIC) han contribuido al desarrollo de recursos más visuales e interactivos, cuya incorporación en el ámbito educativo puede favorecer un incremento de la motivación en los estudiantes, mejorar la comprensión y de esta manera propiciar aprendizajes más significativos. En relación a esto, hace más de dos décadas que se está investigando, desarrollando e implementado diversas herramientas basadas en software libre y objetos de aprendizajes, tanto para el apoyo a la enseñanza y aprendizaje, como para propiciar dar respuestas a problemáticas de nuestro medio. Otra de las líneas en las que se ha avanzado es la gestión de los procesos de sistemas de votación y de publicación de resultados. A continuación se presentan algunos de los avances alcanzados hasta el momento en las principales líneas de investigación.

**Palabras clave:** Educación, Realidad Aumentada, Realidad Virtual, Voto Electrónico, Objetos de Aprendizaje, Métodos Computacionales.

### 1 Introducción

El presente trabajo está enmarcado en el proyecto de investigación “04/F021: Computación Aplicada a las Ciencias y al Medio”. Participan docentes y estudiantes avanzados de la carrera de Licenciatura en Ciencias de la Computación (FAIF) e investigadores de la Universidad Antonio Nariño de la ciudad de Bogotá, Colombia. Se está trabajando en las siguientes líneas de investigación: 1) Desarrollo y Uso de Recursos TIC, 2) Métodos Computacionales y 3) Realidad Aumentada y Realidad Virtual. Estas líneas están interrelacionadas, persiguiendo como propósito general acercar la Universidad al medio.

En la actualidad existe un gran abanico de recursos tecnológicos que pueden favorecer el desarrollo de escenarios digitales más motivadores, visuales e interactivos, con la ventaja de poder ser aplicados a una gran variedad de contextos, no sólo el educativo. En este sentido los desarrollos teóricos y tecnológicos realizados en proyectos anteriores han mostrado los resultados positivos obtenidos con el trabajo de gestores online de torneos de programación, evidenciando además problemáticas a seguir investigando como lo es el agregado en estas herramientas de opciones que faciliten el

trabajo colaborativo y la comunicación dentro de los equipos que compiten. Uno de los desarrollos propuestos persigue la generación de un gestor de torneo online, que mediante el uso de lenguajes de programación basado en bloques y mecanismos diversos de comunicación brinde soluciones a estas problemáticas [1,2].

Dentro de las tecnologías que han emergido este último tiempo se encuentran la Realidad Aumentada (RA) y la Realidad Virtual (RV), que además pensando en nuestro contexto pueden ser de gran ayuda en las áreas de salud, turismo y difusión del patrimonio cultural. La RA permite agregar información sintética a un contexto real complementándolo, para esto sólo se necesita de un dispositivo provisto de una cámara. Esta información que se agrega puede ser a través de texto, audio, imágenes, videos u objetos 3D, estos últimos son los que más atraen la atención por sus características visuales y por sus posibilidades de manipulación e interacción [3]. Pensando en la educación no formal, que se da en las instituciones culturales que tienen a su cargo el acopio y la preservación del patrimonio, como es el caso de los museos de nuestra región, la RA facilitará la generación de modelos 3D de piezas allí resguardadas, dando al público en general la posibilidad de verlas con más detalle y manipularlas fuera de las vitrinas donde están exhibidas normalmente. Se prevé el desarrollo de modelos 3D y software RA que permitan, desde un punto de vista didáctico, aprovechar las posibilidades de interacción, interpretación e intermediación que ofrecen este tipo de recursos para atraer el interés del público hacia el patrimonio, introduciendo mejores formas de comunicarse con él durante la visita. Respecto a la Realidad Virtual, esta tecnología se diferencia de la anterior en que permite a los usuarios sumergirse totalmente en un ambiente sintético simulado por un dispositivo, requiriendo para ello el empleo de estímulos visuales y auditivos entre otros [4].

El uso de tecnología ofrece ventajas en una gran variedad de contextos, respecto al sistema electoral en los últimos años se han propuesto cambios tanto a nivel universitario, municipal, provincial y nacional. Las mejoras obtenidas en la velocidad del proceso de escrutinio sobresalen como principal argumento de su aplicación. El sistema Gukena ha sido desarrollado en el año 2015 por integrantes de este proyecto y empleado en seis elecciones de la Universidad Nacional del Comahue, mostrando que es posible en el sistema tradicional de boleta de papel, mejorar considerablemente la velocidad con que se obtienen los resultados de la elección agilizando la comunicación de los datos en el sistema de escrutinio. Los desarrollos realizados para el centro de cómputos de la UNCo sirvieron de base para analizar los sistemas electorales del país, en particular nos interesa seguir profundizando la investigación enfocada en nuestra región. El objetivo es el análisis técnico y pragmático de un sistema de votación para que pueda ser utilizado en entornos académicos, institucionales y corporativos [5, 6].

Dentro de la gran variedad de recursos TIC existentes se encuentran los Objetos de Aprendizaje (OA). Teniendo en cuenta la dificultad que representa para los docentes del área de matemáticas la enseñanza de algunos conceptos, tales como los algoritmos numéricos iterativos y el cálculo para funciones de dos variables, se continuará trabajando en el desarrollo de OA como unidades pedagógicas virtuales, específicas y de granularidad baja, para desarrollar material educativo dinámico e interactivo en la enseñanza de métodos numéricos para funciones de una y dos variables. El estudio del

espacio tridimensional requiere que el estudiante desarrolle habilidades de visualización, que tiene que ver con los procesos y capacidades de los individuos para realizar tareas que requieren ver o imaginar mentalmente objetos geométricos espaciales, así como relacionar y realizar operaciones o transformaciones geométricas con los mismos. Específicamente se busca desarrollar material que mejore el proceso de enseñanza aprendizaje de importantes conceptos de cálculo de varias variables, mediante la visualización 3D de las superficies cuádricas y sus operaciones [7].

## **2 Algunos avances y resultados esperados**

Se prevé continuar con la transferencia de resultados por medio de la presentación de artículos en congresos, jornadas, entre otros y su publicación en revistas nacionales e internacionales. Además está prevista la transferencia a través del dictado de diferentes cursos de posgrado, en la Maestría en Ciencias de la Computación de la Facultad de Informática y en la Maestría en Enseñanza en Escenarios Digitales. La organización de dos ediciones del Workshop sobre Enseñanzas en Escenarios Digitales ha contribuido a la difusión de buenas prácticas educativas desarrolladas en los distintos niveles educativos, además de la compilación en un libro de las que han resultado más creativas e innovadoras [8].

Particularmente, el período de pandemia 2020-2021 ha evidenciado la importancia de tener docentes formados en el uso de TIC, además de contar con infraestructura tecnológica adecuada para la formación en estos nuevos escenarios digitales. En este sentido se ha trabajado en el desarrollo y uso de diferentes herramientas digitales empleando OA provistos de tecnología móvil, posibilitando así el acceso y generación de contenidos educativos reutilizables, independientes de la plataforma de uso [9,10]. De la misma manera en que se ha trabajado en proyectos anteriores, se prevé continuar con la extensión al medio. Mediante la creación de un observatorio electoral de la Universidad Nacional del Comahue se prevé contribuir a la transparencia y seguimiento de los procesos electorales de las Provincias de Río Negro y Neuquén. En este sentido se han producido algunos avances [11].

En relación al trabajo con las tecnologías RA y RV se han realizado algunos desarrollos [12,13], y se continúa trabajando buscando posibles usos en sistemas de asistencia a la conducción (Advanced Driver Assistance Systems - ADAS), además de favorecer el desarrollo de recursos que potencien la difusión y ampliación de información sobre sitios turísticos, circuitos productivos de la región, y resguardo del patrimonio cultural.

## **3 Formación de Recursos Humanos**

El grupo de investigación viene trabajando y formándose en proyectos anteriores a través de la realización de cursos de posgrado, extensión, entre otros. El proyecto comenzó en el año 2022, hasta el momento ha finalizado el trabajo de un becario CIN,

y se defendieron 7 tesis, 4 tesis de LCC y 3 tesis de maestría. Siguen en desarrollo dos tesis de grado de la carrera de Licenciatura en Ciencias de la Computación, una de maestría y tres de doctorado. Un docente investigador se encuentra en etapa de tesis en la Maestría en Ciencias de la Computación (UNCo), y dos docentes investigadoras están finalizando el cursado de la Maestría en Enseñanza en Escenarios Digitales (UNCo, AUSA).

## Referencias

1. Fracchia, C. C., & Bramardi, S. (2020). Torneos de programación: combinando los aprendizajes competitivo y cooperativo. *Revista TEYET*, (27), e6-e6.
2. Stacco, J. M., Kogan, P., Godoy, I., Rodríguez, J., & Fracchia, C. (2020). Zorzal: un ambiente colaborativo basado en bloques para competencias de programación. In *II Simposio Argentino de Educación en Informática (SAEI 2020)-JAIIO 49*
3. Azuma, R. T. (2019). The road to ubiquitous consumer augmented reality systems. *Human Behavior and Emerging Technologies*, 1(1), 26-32.
4. Sherman WR, Craig AB (2003). *Understanding Virtual Reality: Interface, Application, and Design*. Amsterdam: Morgan Kaufmann Publishers.
5. Soto, S., Ramos, C., Kogan, P., Vaucheret, C. A., & Rodríguez, J. (2018). Gukena: escrutinio descentralizado para voto ponderado. In *XII Simposio de Informática en el Estado (SIE 2018)-JAIIO 47*
6. Kogan, P., Soto, S., & Vaucheret, C. A. (2019). Aportes al proceso de fiscalización electrónica: experiencia sobre uso de la boleta única electrónica en elecciones 2019, provincia de Neuquén. In *XIII Simposio de Informática en el Estado (SIE 2019)-JAIIO 48 (Salta)*.
7. Allan, C., Parra, S., & Martins, A. (2017). Objetos de aprendizaje para la interpretación geométrica de métodos numéricos: uso de GeoGebra. *TE & ET*.
8. Fracchia, C., Kogan, P., & Aranda, G. (2023). WEED: I workshop enseñanza en escenarios digitales. <http://rdi.uncoma.edu.ar/handle/uncomaid/17413>
9. Rodríguez, J., Kogan, P., Guerrero, G., Pereyra, G., & Torrico, F. (2022). Diseño participativo de secuencias didácticas basadas en el desarrollo de aplicaciones móviles en la escuela. In *XXVIII CACIC*.
10. Rodríguez, J., Parra, G., Gili, G., Parra, S., Dolz, D., & Roumec, H. (2023). Una Máquina de Turing en la Escuela. In *XXVIII CACIC*.
11. Kogan, P., Soto, S., Vaucheret, C., & Zacharias, D. (2022). ¿Cómo aumentar la velocidad de publicación de resultados de una elección?. *Memorias de las JAIIO*, 8(12), 69-79.
12. Castro M.E., Fracchia, C.C., Willging, P. (2022). Aplicación de Realidad Aumentada en la enseñanza de la asignatura Anatomofisiología. 9° Seminario Internacional de Educación a Distancia “ESCENARIOS INÉDITOS EN LA EDUCACIÓN SUPERIOR. PERSPECTIVAS, HUELLAS Y EMERGENTES”, RUEDA.
13. Rodríguez, C. A., Valderrama, S., Vargas, D., Eliseo, M. A., Fracchia, C. C., & Roa, K. (2023). Quizzes via Augmented Reality on Learning Management System: A Case Study of Moodle. *Journal of Educators Online*, 20(1), 137-144. [https://doi.org/10.9743/jeo.2023.20.1.5\\*](https://doi.org/10.9743/jeo.2023.20.1.5*)

## 04/F022: Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación

Gerardo Parra, Jorge Rodríguez, Federico Amigone, Juan de la Fuente, Daniel Dolz, Guillermo Grosso, Daiana Bonet Peinado, Sandra Boari, Marcos Cortez, and Ingrid Godoy

Grupo de Investigación en Lenguajes e Inteligencia Artificial  
Facultad de Informática  
Universidad Nacional del Comahue

**Abstract.** Actualmente las Ciencias de la Computación en los niveles obligatorios del sistema educativo están adquiriendo un papel significativamente importante. Por lo tanto, es prioritario trabajar en la producción de recursos educativos en el área.

Este proyecto de investigación persigue el objetivo de desarrollar modelos y construcciones teóricas que fortalezcan y redimensionen los procesos de enseñanza de las Ciencias de la Computación.

**Keywords:** Educación en Ciencias de la Computación · Escuela Secundaria · Recursos Educativos Desenchufados.

### 1 Contexto

Este proyecto de investigación, iniciado en Enero de 2022 y con finalización prevista para Diciembre de 2025, se fundamenta en resultados anteriores de los siguientes proyectos de investigación:

- Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación (Cod: 04/F015).
- Agentes Inteligentes en Entornos Dinámicos (Cod: 04/F006).
- Informática Aplicada al Medio Académico y Científico (Cod: 04/F004).

Además, el proyecto busca construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el ámbito de la Extensión Universitaria en relación a la aplicación de la tecnología de Agentes y Modelos Formales al contexto educativo. Las iniciativas de referencia se expresan en los siguientes proyectos de Extensión:

- 2023 - Conectados con la Escuela.
- 2023 - FaIComm: Charlando sobre Informática.
- 2022 - Proyecto de Extensión Institucional — Reduciendo Brechas.
- 2021 - Conectados con la Escuela: Desafíos para Aprender Computación.

- 2020 - Conectados con la Escuela: Desafíos virtuales para aprender computación.
- 2019 - Vamos a la Escuela: Acercando las Ciencias de la Computación a la Escuela Media.
- 2018 - Vamos a la Escuela: Acercando las Ciencias de la Computación a la Escuela Media.
- 2017 - Vamos a la escuela: Acercando las Ciencias de la Computación a la escuela Media.
- 2016 - Agentes Robots: Construcción de Prototipos en la Escuela Media.
- 2016 - Agentes Robots: Vamos a la Escuela.
- 2016 - Agentes Robots: Divulgando Computación en la Escuela Media.
- 2016 - La Competencia como Eje Motivador para Acercarnos a la Programación.
- 2015 - Agentes Robots: Acercando la Escuela Media a la Universidad a través de la Computación.
- 2015 - Olimpiadas de Programación para Escuela Media.
- 2014 - Con la Escuela Media Programamos Robots y conocemos más de Computación.
- 2014 - Entrenamiento en la Programación de la Computadora a partir de una Aplicación para Competencias.
- 2013 - Divulgando Computación con Robots en la Escuela Media.

## 2 Objetivos

El objetivo general de este proyecto de investigación es desarrollar construcciones teóricas, modelos y agentes inteligentes que fortalezcan los procesos de enseñanza de las Ciencias de la Computación. A partir de este objetivo general, a continuación enumeramos las metas específicas que nos proponemos alcanzar:

1. Desarrollar agentes inteligentes orientados al procesamiento de información.
2. Desarrollar e implementar prototipos de agentes inteligentes que soporten los modelos conceptuales para el aprendizaje colaborativo.
3. Avanzar en la identificación y descripción de los elementos que permitan demarcar la didáctica de la informática y su relación con otros campos del conocimiento.
4. Formular modelos para la investigación y el desarrollo en Educación en Ciencias de la Computación generales y situados en campos problemáticos particulares.
5. Fortalecer la formación de recursos humanos en las áreas de conocimiento relacionadas.

## 3 Resultados Obtenidos y Esperados

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[6, 10, 11].

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 entró en ejecución en el año 2023, incorpora, entre otros, contenidos de Teoría de la Computación, Representación del Conocimiento y Redes de Computadoras, entre otros nudos disciplinares[3].

Estos contenidos usualmente son considerados abstractos y con cierta complejidad para estudiantes sin formación previa en la disciplina. Los recursos educativos desenchufados, sobre todo los desarrollados en el ámbito de la iniciativa *CS Unplugged*, están ampliamente difundidos y suelen ofrecer un primer contacto satisfactorio[1, 8, 12].

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[7, 2]. 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 nociones fundamentales de teoría de la computación. Los estudios desarrollados aportan resultados alentadores en relación a la efectividad de los recursos elaborados[9, 5, 4].

El diseño y estudio de efectividad fueron orientados por los modelos teóricos definidos en esta 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 de este proyecto 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, otros dos miembros están cursando la Maestría en Enseñanza en Escenarios Digitales, que desarrollan de manera conjunta las Universidades Na-

cionales 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 estudiantes integrantes 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.

## References

1. Bell, T., Vahrenhold, J.: Cs unplugged—how is it used, and does it work? In: *Adventures Between Lower Bounds and Higher Altitudes*, pp. 497–521. Springer (2018)
2. Colmenares E, A.M.: Investigación-acción participativa: una metodología integradora del conocimiento y la acción. *Voces y Silencios. Revista Latinoamericana de Educación* **3**(1), 102–115 (2012)
3. Consejo Provincial de Educación de la Provincia de Neuquén: Diseño Curricular Jurisdiccional de los tres primeros años de la Escuela Secundaria Neuquina. Resolución N° 1463/18 (2018)
4. 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. In: *XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, U. Champagnat. Mendoza, Argentina)*. (2022)
5. Dolz, D., Martínez, R., Parra, G., Rodríguez, J., Ginez, N.: Recursos educativos desenchufados para la enseñanza de las ciencias de la computación en la escuela secundaria. In: *XV TE&ET* (2020)
6. K-12 Computer Science Framework Steering Committee: *The K–12 Computer Science Framework*. ACM (2016)
7. Martí, J.: *La investigación-acción participativa: estructura y fases* (2017)
8. Nishida, T., Kanemune, S., Idosaka, Y., Namiki, M., Bell, T., Kuno, Y.: A cs unplugged design pattern. *ACM SIGCSE Bulletin* **41**(1), 231–235 (2009)
9. Rodríguez, J., Parra, G., Gili, G., Parra, S., Dolz, D., Roumec, H.: Una máquina de turing en la escuela. In: *XXVIII CACIC* (2022)
10. Sadosky, F.: *CC – 2016 Una propuesta para refundar la enseñanza de la computación en las escuelas Argentinas*. Fundación Sadosky, Argentina (2013)
11. Society, R.: *After the reboot: Computing education in UK schools*. Policy Report (2017)
12. Taub, R., Armoni, M., Ben-Ari, M.: Cs unplugged and middle-school students’ views, attitudes, and intentions regarding cs. *ACM Transactions on Computing Education (TOCE)* **12**(2), 1–29 (2012)

# Avances del Proyecto “Tecnologías de Datos Espaciales, Visualización y Realidad Virtual”

Luis Reynoso<sup>1</sup>, Silvia Amaro<sup>1</sup>, Lidia Lopez<sup>1</sup>, Viviana Sanchez<sup>1</sup>, María José Rotter<sup>1</sup>, Sergio Cotal<sup>1</sup>

<sup>1</sup>Facultad de Informatica - Universidad Nacional del Comahue  
Neuquen, Argentina  
{luis.reynoso, silvia.amaro}@fi.uncoma.edu.ar

**Abstract.** Muchos organismos del país de diferentes jurisdicciones (nacional, provincial y municipal) aplicando estándares internacionales del Consorcio Geoespacial Abierto (OGC) publican información geoespacial con el propósito que pueda consumirse e integrarse en cartografía digital para estudios y análisis. En los últimos años, y alrededor del mundo se han multiplicado las iniciativas con las Infraestructuras de Datos Espaciales (IDE) logrando interoperabilidad de la información geoespacial. Además, la accesibilidad de la realidad virtual está en constante crecimiento. En combinación con la realidad aumentada y la integración del análisis de big data en este entorno, se persigue añadir un valor esencial al proceso en su conjunto. El objetivo principal es emplear estas tecnologías para crear aplicaciones y paneles de control que aprovechen datos abiertos y geoespaciales, proporcionando una manera efectiva de interactuar con la información y los resultados.

**Keywords:** Información geoespacial · Infraestructura de datos espaciales · Visualización de datos · Realidad virtual · Big data

## 1 Contexto

Este estudio se lleva a cabo en el marco del Proyecto de Investigación Tecnologías de Datos Espaciales, Visualización y Realidad Virtual de la Facultad de Informática de la Universidad Nacional del Comahue. La implementación de estas tecnologías implica la concepción de estrategias, la creación de aplicaciones y tableros de control que involucran a la comunidad académica, la sociedad civil y los sectores público y privado, para la explotación de datos mediante tecnologías de infraestructura de datos espaciales, visualización, big data, realidad virtual y datos abiertos. Cabe destacar que el proyecto recibe financiamiento de la propia Universidad Nacional del Comahue.

## 2 Introducción

La visualización de datos [13], la infraestructura de datos espaciales [4], el big data [11], los datos abiertos (open data) [1], el gobierno abierto (open gov)

y la realidad aumentada son modelos tecnológicos disruptivos que, ya sea de forma independiente o en combinación, están siendo cada vez más utilizados en nuestra vida cotidiana. Estas tecnologías están despertando un creciente interés en diversas entidades gubernamentales y empresas debido a su inmenso potencial estratégico.

Una de las principales métricas para evaluar el impacto y el alcance de estas tecnologías, al igual que en el caso de los datos abiertos y los geoservicios [6], es su recopilación y reutilización. Tanto los generadores de datos abiertos como los desarrolladores y los enriquecedores que los explotan, contribuyen a la cadena de valor de los datos abiertos y los geoservicios. La combinación de diferentes conjuntos de datos permite la creación de nuevos datos que pueden dar lugar a nuevos servicios o productos.

Numerosos organismos gubernamentales ponen a disposición cada vez más datos abiertos en línea [3], principalmente para garantizar el ejercicio efectivo del derecho de acceso a la información pública, fomentar la participación ciudadana y promover la transparencia en la gestión pública, de acuerdo con la Ley 27275 de Acceso a la Información Pública. Estos conjuntos de datos abiertos (datasets) están en constante crecimiento. A pesar de que, cada vez con más frecuencia, se accede a los datos abiertos a través de los portales nacionales de datos abiertos, seguidos por aquellos que se obtienen directamente de las administraciones públicas, los conjuntos de datos abiertos pueden ser tan extensos que la visualización tabular resulta insuficiente para su comprensión. En cambio, la geolocalización de la información en mapas y la explotación de datos mediante técnicas de big data y visualización [7, 12, 14] nos ayudan a comprender mejor los datos, extraer conclusiones y resultan altamente ilustrativos cuando se representan en un mapa. De manera similar, los paneles de control constituyen una herramienta de visualización de los indicadores clave y los datos principales que facilitan y simplifican la comprensión del estado general de un tema. 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 ciencia de datos aplicadas a conjuntos diversos.
- 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.

### 3 Líneas de Investigación y Desarrollo

En este proyecto podemos identificar las siguientes líneas de investigación y desarrollo:

- Infraestructuras de Datos Espaciales (IDE) Una IDE se compone de un conjunto de políticas, estándares, organizaciones y recursos tecnológicos que facilitan la producción, obtención, uso y acceso a la información geoespacial [4]. Las IDE permiten publicar información disponible para todos los interesados lo cual permite la democratización de la información geoespacial. La aplicación de estándares de OGC posibilita que los datos puedan ser publicados en la web en conjunto con sus metadatos y catálogos [2, 4], e integrados unos con otros en el caso de utilizar el mismo Sistema de Referencia de Coordenadas (CRS, del inglés Coordinate Reference Systems). Estas infraestructuras son un importante patrimonio en la construcción de ecosistemas digitales y son un tipo de datos fundamental para la localización de Objetivos de Desarrollo Sostenibles (ODS). Sin embargo la tecnología empleada por las IDE no es un contenido que se desarrolle en las curricula de titulaciones Geoespaciales, por ello es importante el desarrollo de capacidades no solo de profesionales egresados sino también de estudiantes e incluso agentes técnicos que trabajan en distintos organismos y requieren de información geoespacial. Asimismo, el cúmulo de información geoespacial muchas veces requiere de herramientas efectivas para su visualización a partir de aplicaciones, API, visualizadores y tableros de control, lo cual permite la generación de valor agregado, posibilitando que las IDE se conformen como un poderoso instrumento en la gobernanza de datos.
- Visualización de Datos, Big Data y Realidad Virtual Big data no se trata únicamente de los datos, se trata de conocimientos y percepciones en los datos, y su interpretación. La visualización de datos se limita a una pantalla

o a un papel plano (ambos planos). En este sentido, es aquí donde entra en juego la realidad virtual. Es necesario estudiar, probar e implementar espacios o plataformas que integren la percepción humana de los patrones y las relaciones entre los datos. Es decir, mirar los datos desde adentro hacia afuera en lugar de desde afuera hacia adentro. Existen desarrollos que unen Big Data y realidad virtual, combinando análisis de datos modernos y aprendizaje automático con la visualización de datos de muy buena performance.

- **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.
- **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.

## 4 Resultados Esperados

A continuación se enumeran y describen algunos resultados alcanzados en 2022 y 2023 a partir de la investigación y aplicación de las tecnologías de interés.

- **Tablero Mapa Web de la UNCo:** En conjunto con la Secretaría de Planificación y Desarrollo Institucional (SPDI) se desarrolló en el año 2021 un tablero que visualiza 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. El tablero fue actualizado en mayo de 2023 con información proporcionada por la SPDI y fue publicado por la SPDI: <https://planeamiento.uncoma.edu.ar/?p=3721>

- Información geoespacial e Objetivos de Desarrollo Sostenibles (ODS): Nuestro proyecto participó en el segundo taller “Diálogo de Saberes en clave de ODS” organizado por la Secretaría de Planeamiento y Desarrollo Institucional. A partir de la asociación de todos los proyectos de investigación de la UNCo a Objetivos de Desarrollo Sostenibles (ODS) se construyó un tablero de datos con información sobre proyectos de investigación, ODS y Unidades Académicas. El mismo se puede consultar en: [https://opendata.fi.uncoma.edu.ar/SPDI/SPDI\\_TableroODS.html](https://opendata.fi.uncoma.edu.ar/SPDI/SPDI_TableroODS.html)  
Se desarrolló una presentación sobre como las Infraestructuras de Datos Espaciales y el Marco Integrado de Información Geoespacial permiten localizar ODS. La presentación fue subida a youtube por la SPDI: <https://www.youtube.com/watch?v=uAGBqxAbv1s>. El contenido de ambos resultados ha sido publicado en novedades de SPDI: <https://planeamiento.uncoma.edu.ar/?p=4200>
- Seminarios sobre Ciencia de Datos Geoespaciales en distintas Universidades. Se desarrollaron cursos presenciales sobre Ciencia de Datos Geoespaciales en las siguientes Universidades:
  - Seminario “Aportes de Ciencias de Datos a la investigación en Geografía”, realizado en la Universidad del Centro de la Provincia de Buenos Aires (UNCPBA), 1 al 3 de Junio de 2023, un seminario para alumnos y docentes de 60 horas; se brindó una charla abierta en UNCPBA sobre “Datos abiertos y nuevas tecnologías para la investigación en Ciencias Sociales: Nuevas formas de interacción, tratamientos de datos y visualización”;
  - Se disertó en las Jornadas presenciales de Sensibilización y Uso de Información Geoespacial, que se desarrollarán en Hall central de Rectorado de la Universidad Nacional de San Juan (UNSJ) y se brindó un taller teórico práctico “Ciencia de Datos a partir de información Geoespacial”, el 15 y 16 de Agosto de 2023;
  - Se desarrolló el seminario “Ciencia de Datos: Reportes, Mapas Web y Tableros con R Studio” de 36 horas en la Universidad de la Patagonia San Juan Bosco (UNPSJB) el 25, 26 y 27 de Octubre de 2023.
- Participación en la Infraestructura de Datos Espaciales del país: IDERA Nuestra activa participación en IDERA durante los últimos en el grupo técnico de trabajo de Academia y Ciencia ha permitido alcanzar los siguientes resultados:
  - Organización del Ciclo de Desarrollo de Capacidades. Se organizaron cuatro encuentros virtuales abiertos para fomentar el desarrollo de capacidades en información geoespacial, sobre: (1) Evaluación de Calidad de la Información Geográfica mediante ISO 19157 (2) Metodología para Evaluar la exactitud posición de límites territoriales. (3) Geoportal Federal para la Gestión Local. Una herramienta de análisis territorial para municipios. (4) Transformando Datos en Acción: SDMX. Experiencias de GIS y Estadísticas en la Implementación de Estándares Globales.
  - Organización del programa académico de las XVI y XVII Jornadas IDERA, La tarea consistió en la gestión de ponencias, short-papers y posters

recibidos en las jornadas, la asignación de trabajos a miembros del comité evaluador, notificación a los autores y, confección, diseño y comunicación del programa académico. Las XVI jornadas IDERA se desarrollaron en Córdoba (28 de Junio al 1 de Julio de 2022) y las XVII jornadas IDERA en Santa Rosa, La Pampa (del 28 al 30 de Junio de 2023).

- Coordinación del Grupo Técnico de Trabajo Academia y Ciencia de IDERA. La actividad consiste en coordinar la tarea de los subgrupos del Comité Editor, el Comité Evaluador, el Ciclo de Desarrollo de Capacidades, el Programa de Desarrollo de Capacidades y el de traducción de documentos de Vías Estratégicas del IGIF. Se realizan reuniones permanentes de los subgrupos en conjunto con miembros de 34 universidades adheridas a IDERA. También se participa en el Consejo Directivo de IDERA y en otros grupos técnicos de trabajo como Marco Institucional y Provincias.
- Publicación de Libros digitales. Se compiló y publicó el libro digital de las Jornadas Virtuales de Academia y Ciencia de IDERA, año 2020: enlace publicado en el sitio web de IDERA. Se están editando los libros de las jornadas del 2022 y 2023.

## 5 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.

## 6 Bibliografía

1. Berends, J., Carrara, W., Engbers, W., Vollers, H. (2020) Re-using Open Data. A Study of Companies transforming Open Data into Economic and 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](https://data.europa.eu/sites/default/files/re-using_open_data.pdf)
2. Guisande Gonzalez, C., Vaamonde Liste, A., Barreiro Felpeto, A.. Tratamiento de datos con R, Statistica y SPSS. Madrid: Ediciones Diaz 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:  
<https://ri.conicet.gov.ar/handle/11336/27085>

5. Ministerio de Salud de la Nación, Argentina (2021b). COVID-19. Casos registrados en la República Argentina. Datos Abiertos del Ministerio de Salud.
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 Jersey: 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. Khattak, 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.

# Robots de Servicio con Navegación Autónoma en Espacios Públicos

Rafael Ignacio Zurita<sup>1</sup>, Claudio Vaucheret<sup>1</sup>, Miriam Lechner<sup>1</sup>, Alejandro Mora<sup>1</sup>, Luis Coralle<sup>1</sup>, and Ana Candelaria Alvarez<sup>2</sup>

<sup>1,2</sup>04/F024: Identificación, Evaluación y Adaptación de Componentes de Hardware y Software para Sistemas Embebidos - Parte III

<sup>1</sup>Facultad de Informática, Universidad Nacional del Comahue.

<sup>2</sup>Escuela de Ingeniería, Universidad Autónoma de Barcelona.

{rafa, cv, mtl, alejandro.mora, luyo}@fi.uncoma.edu.ar,

anacandelaria.alvarez@uab.cat

<http://se.fi.uncoma.edu.ar/>

**Resumen** El eje del proyecto de investigación *04/F024: Identificación, Evaluación y Adaptación de Componentes de Hardware y Software para Sistemas Embebidos - Parte III* es la investigación aplicada de las áreas de conocimiento involucradas en los sistemas que conforman los robots autónomos de servicio. En particular, en la programación de sistemas embebidos e inteligentes combinados, para la construcción de robots autónomos de servicio FAI. Se espera, a largo plazo, que estos robots provean asistencia en la mejora de la calidad de vida universitaria en los espacios físicos de la Facultad de Informática.

**Keywords:** Robótica. Sistemas Embebidos. Robots de Servicio. Calidad de Vida. Robótica en fruticultura. Navegación Autónoma.

## 1. Introducción

En la actualidad los robots autónomos tienen un alto potencial para colaborar en la calidad de vida en ambientes públicos [1]. Esta posibilidad surgió a partir de una combinación de factores. Por un lado, la cuarta revolución industrial plantea el uso de robótica interconectada para trabajos con riesgos para la vida humana. Segundo, la tecnología de las computadoras embebidas ha alcanzado el rendimiento necesario para la ejecución de sistemas que requieren alto poder de cómputo, como los que conforman un robot autónomo. Finalmente, a partir de la pandemia covid-19 surgió la necesidad inmediata de utilizar robots de servicio en áreas públicas [2–4]. Por ejemplo, para la prevención de enfermedades a través de la desinfección de superficies con radiación ultravioleta (UV) sin contacto. Se ha demostrado que los dispositivos de luz ultravioleta (como PX-UV) son efectivos para reducir la contaminación en superficies de alto contacto en hospitales. En lugar de la desinfección manual, el cual aumenta el riesgo de exposición del personal de limpieza, existe un alto potencial para realizar una desinfección

## 2 Sistema Embebido para Posicionamiento de un Tractor en Fruticultura

eficaz con robots autónomos (o controlados remotamente) [5]. Las oportunidades residen en la navegación inteligente autónoma, y la detección de áreas de alto riesgo y de alto contacto, combinadas con otras medidas preventivas.

Por lo tanto, si un robot puede navegar sin intervención humana, y lograr trayectorias necesarias a sus objetivos, su colaboración en tareas peligrosas, pesadas y/o rutinarias podrían ser innumerables. Algunas de ellas son: desinfección de lugares públicos; entrega de medicamentos o transporte de sustancias químicas peligrosas (entre oficinas en laboratorios, o entre habitaciones de hospitales y hogares); transporte de cargas pesadas (monitores, conjunto de libros, material embalado, etc). También puede realizar tareas rutinarias como navegar entre las oficinas y aulas de una institución, y observar que las luces estén apagadas durante el día, y puertas y ventanas cerradas en invierno; detectar si hay presencia de gas en alguna habitación u oficina, o si la temperatura es demasiada elevada (detectando un posible foco de incendio o peligro para la vida humana); detectar si la temperatura es muy baja (y mejorar el ambiente si es necesario) previa a la actividad de personas en el lugar; detectar si hay presencia de humedad elevada y lograr evitar una inundación; etc.

Estas funcionalidades de un robot móvil autónomo se pueden categorizar. La siguiente lista enumera algunas de las categorías más importantes:

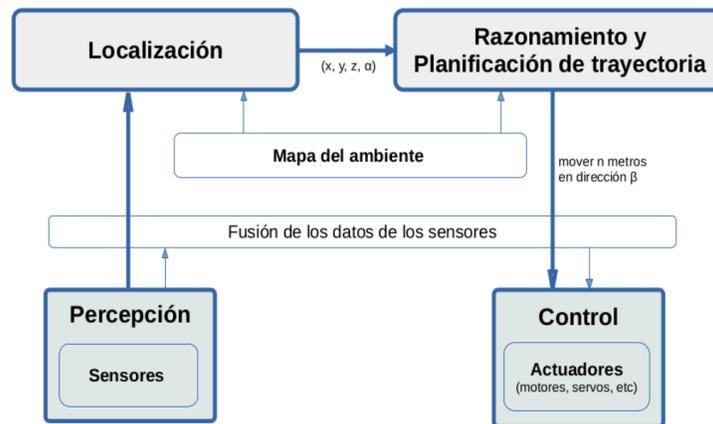
1. Evitar siniestros o peligros para la vida humana.
2. Conciencia y ahorro energético.
3. Mejorar la calidad de vida evitando enfermedades.
4. Brindar seguridad física edilicia y de las personas.
5. Contabilizar stock y evitar faltantes antes que se necesiten.
6. Realizar transporte de equipamiento pesado o sustancias peligrosas.
7. Colaborar en la búsqueda de objetos y personas.
8. Realizar asistencia a personas que lo requieran.

Las áreas recién enumeradas requieren una plataforma base de experimentación común: un robot de servicio móvil autónomo de interior. A partir del trabajo iniciado en el proyecto anterior -una plataforma móvil- se continúa en este proyecto con la construcción de los subsistemas restantes, para convertir la misma en un robot autónomo. Es posible luego agregar los mecanismos necesarios para investigar (y evaluar) algunas de las funcionalidades de servicio listadas anteriormente, y que permitan validar el trabajo de investigación aplicada base alcanzado con el robot móvil.

## 2. Contexto

La línea de investigación se inserta en el contexto del Proyecto 04/F024 Identificación, Evaluación y Adaptación de Componentes de Hardware y Software para Sistemas Embebidos - Parte III, dirigido por el Mg. Claudio Vaucheret, y co-dirigido por el Lic. Rafael Zurita. Además, integrantes de este proyecto colaboran con el proyecto de investigación *Sistemas de visión, robótica y navegación autónoma de vehículos para la fruticultura digital*, dirigido por el Ing. Dario

Mendieta, y co-dirigido por el Dr. Marcelo Moreyra, del Depto. de Electrotecnia, Facultad de Ingeniería.



**Figura 1.** Subsistemas de un robot con capacidades de navegación autónoma.

### 3. Líneas de Investigación y Desarrollo

Un robot móvil autónomo de servicio está compuesto de varios subsistemas: sensado, percepción, localización, planificador de rutas, y control (Figura 1) [6]. Se detallan a continuación las líneas de investigación que deben ser trabajadas para lograr estos subsistemas.

- I+D de un sistema de percepción para la localización y detección de obstáculos, basado en sensores de bajo costo.
- I+D de un sistema de localización basado en un sistema de posicionamiento global, y uno local.
- I+D de un sistema planificador de ruta, capaz de indicar la ruta más adecuada en términos de diferentes métricas (más corta, más segura, más rápida, etc).
- I+D de un sistema de control capaz de operar los actuadores del robot (motores, servos, etc). El sistema debe contar con un control PID y ser capaz de lograr las instrucciones de operación del planificador de rutas.
- I+D Desarrollar un sistema de detección de obstáculos y personas, capaz de identificar la distancia de los objetos y su ángulo de localización.

También, para las etapas de integración del robot en ambientes públicos con personas es requerido implementar un sistema de interacción humano-computadora que permita la comunicación con el robot móvil, para poder solicitarle que haga

#### 4 Sistema Embebido para Posicionamiento de un Tractor en Fruticultura

algo o hacerle una pregunta cuya respuesta contenga información accesible al robot. Finalmente, validar las capacidades de los prototipos al navegar en un ambiente de interior dinámico, con personas. Por ejemplo, midiendo y evaluando los avances en la mejora de la calidad de vida universitaria una vez que los robots de servicio coexistan en el edificio de la Facultad de Informática con la comunidad.

### 4. Resultados Obtenidos y Esperados

El objetivo principal de este proyecto es el estudio, diseño y construcción de robots móviles autónomos de servicio, a partir de las plataformas móviles ya construidas, para asistir a la calidad de vida de las personas en ambientes interiores [7].

Hasta el presente, se han construido tres plataformas móviles electromecánicas, de dimensiones adecuadas para un futuro robot de servicio que asista a personas. Además, ya se han desarrollado, durante este proyecto, los sistemas de sensado, percepción y localización, lo que le permite a los futuros robots determinar por su cuenta su localización (ubicación y orientación). Actualmente, el robot puede ser operado remotamente, a través de su visión.

En el área de bajo nivel se alcanzó un objetivo esencial para cada sistema empujado, y que consiste en el port del sistema operativo XINU, para ser utilizado como sistema operativo de tiempo real en los microcontroladores que adquieren señales y activan actuadores [8].

Se espera continuar con los subsistemas de planificación, y con un sistema de localización secundario, que debe estar activo para cuando el sistema de localización primario se encuentre inoperable.

### 5. Formación de Recursos Humanos

Tres estudiantes de grado comenzaron con un plan de trabajo. Dos de ellos han finalizado sus estudios, realizando sus tesis en el área de investigación de este proyecto. Uno de los docentes investigadores inició el Doctorado en Ingeniería, en la FAIN, UNCOMA; bajo la dirección del Dr. Favio Masson (UNS) y el Dr. Marcelo Moreyra (UNCo), trabajando con el tema central de tractores autónomos en fruticultura.

La Dra. Luciana Benotti (FAMAF, UNC), nos asesora de manera externa en la formación de conceptos de IA, ética en IA, y lingüística computacional.

### Referencias

1. How Robots Are Being Used for COVID-19. Robots for infectious diseases web page, <https://roboticsforinfectiousdiseases.org/how-robots-are-being-used.html>, últ. vez accedido 20/10/2023.
2. Li, G., Hou, Y. & Wu, A.: Fourth Industrial Revolution: technological drivers, impacts and coping methods. *Chin. Geogr. Sci.* 27, 626–637 (2017).

3. Sifat, A.H., Bharmal, B., Zeng, H. et al.: Towards computational awareness in autonomous robots: an empirical study of computational kernels. *Complex Intell. Syst.* 9, 6269–6295 (2023).
4. Gupta, A., Singh, A., Bharadwaj, D. et al.: Humans and Robots: A Mutually Inclusive Relationship in a Contagious World. *Int. J. Autom. Comput.* 18, 185–203 (2021).
5. C. R. Kovach, Y. Taneli, T. Neiman, E. M. Dyer, A. J. A. Arzaga, S. T. Kelber: Evaluation of an ultraviolet room disinfection protocol to decrease nursing home microbial burden, infection and hospitalization rates. *BMC Infect. Dis.* 17, 186 (2017).
6. Robin R. Murphy: *Introduction to AI Robotics*. MIT Press, Cambridge, MA, USA, 2nd edition, 2019.
7. Proy. de Inv. UNCo 04/F011 Identificación, Evaluación y Adaptación de Componentes de Hardware y Software para Sistemas Embebidos – Parte II. Director: Rodolfo del Castillo.
8. Zurita, R.I., Alvarez, C., Lechner, M., Mora, A., Zárate, M.: A New Academic RTOS: Xinu Port to the AVR Architecture. In: Pesado, P. (eds) *Computer Science – CACIC 2022. CACIC 2022. Communications in Computer and Information Science*, vol 1778. Springer, Cham. [https://doi.org/10.1007/978-3-031-34147-2\\_17](https://doi.org/10.1007/978-3-031-34147-2_17)

ACTAS DE LAS

JORNADAS DE LA FACULTAD DE INFORMÁTICA



NOV 2023  
NEUQUÉN

FACULTAD DE INFORMÁTICA  
UNIVERSIDAD NACIONAL DEL COMAHUE

*Proyectos de extensión*

# Programación Lógica en la Escuela Primaria

## PROYECTO DE EXTENSIÓN

Directores: Laura A. Cecchi, Jorge P. Rodríguez

Valentina Villarroel, Ian Acosta, Lucas Margni, Emma Dolz,  
Guillermo Guerrero, Santiago Villarroel, Gabriela Aranda, Daniel Dolz,  
Alejandro Fierro, María Eda Cornejo, María Paz Sarrasqueta, Laura Parra,  
Alejandro Mora Vazquez, Valeria Roter, Micaela Malaspina,  
Federico Mamani, Romina Tealdi, and Luis Coralle

Grupo de Investigación en Lenguaje e Inteligencia Artificial  
Facultad de Informática  
Universidad Nacional del Comahue  
`{j.rodri, lcecchi}@fi.uncoma.edu.ar`

**Palabras Clave:** Pensamiento Lógico. Pensamiento Computacional. Formación en computación. Escuela Primaria. Inteligencia Artificial. Programación Lógica. Prolog. Contenidos Escolares sobre Medio Ambiente.

### 1. Objetivos

Desarrollar habilidades de Pensamiento Lógico y Computacional en los estudiantes de nivel primario, a través de la programación lógica y contribuir a la formación en este campo disciplinar de los docentes del nivel primario sin conocimiento previo sobre programación. A los efectos de promover visión integrada entre las Ciencias de la Computación y las problemáticas ambientales se buscará modelar con herramientas de la programación lógica contenidos curriculares relacionados a la ecología y sus conflictos. Las actividades se plantean de modo de favorecer la igualdad de género.

### 2. Fundamentación

Actualmente, existe un consenso internacional sobre la importancia de enseñar Ciencias de la Computación y de desarrollar habilidades de Pensamiento Computacional y Pensamiento Lógico al resolver problemas de propósito general. En la Provincia de Neuquén, en las escuelas primarias no existe aún un espacio curricular específico para la enseñanza de esta disciplina, por lo que esta población de estudiantes se encuentra en situación de vulnerabilidad. La Programación Lógica se presenta como un dispositivo pedagógico para enseñar a programar de un modo diferente al de la Programación Imperativa, promoviendo el pensamiento computacional, lógico, crítico y creativo. Prolog (derivado de “PROgramation en LOGique”) es un lenguaje de Programación Lógica que

nació a comienzos de la década de los 70, de la mano de Alain Colmerauer, Philippe Roussel y Robert Kowalski. Este lenguaje es utilizado en el desarrollo de sistemas en el campo de la Inteligencia Artificial. En el contexto planteado, la propuesta consiste en acercar la Programación Lógica y en particular el lenguaje Prolog a niños de 4to, 5to y 6to grado de Escuelas Primarias, cuyas edades rondan entre los 8 y 10 años al inicio del ciclo lectivo y entre 9 y 11 años al final. Las intervenciones didácticas propuestas en esta iniciativa se plantean en el marco del diseño participativo, donde la interacción entre la Facultad, Educación Digital de la Provincia de Neuquén y las escuelas supone la construcción colaborativa de experiencias educativas, que consideran las singularidades de cada contexto. Se trabajará especialmente sobre ejes relacionados a problemas ambientales, que sean parte de los contenidos que integren las propuestas de enseñanza de cada escuela primaria. Se abordarán estas problemáticas con la Programación Lógica como herramienta para representar soluciones posibles. El equipo de trabajo afectado al proyecto es de carácter interdisciplinar, ya que está conformado por profesionales de Ciencias de la Educación, particularmente de enseñanza primaria y educación mediada por tecnologías, de la Dirección Provincial de Educación Digital, y por profesionales de las Ciencias de la Computación de la Facultad de Informática de la UNCo. Por otra parte, el trabajo que se lleva a cabo en forma conjunta con miembros de la Dirección Provincial de Educación Digital, del Ministerio de Educación y Gobierno de la Provincia de Neuquén, nos ubica como referentes en aspectos relacionados a la enseñanza de la disciplina y nos acerca a los responsables de los desarrollos de las curriculas de la escuela primaria. Trabajar con docentes de las escuelas primarias y organismos públicos otorga mayor persistencia a la mejoras. Actualmente, hay un interés remarcado de la comunidad científica internacional en rescatar ideas de este tipo de experiencias, considerando la importancia de esta clase de pensamiento abstracto y de acercar en forma temprana este paradigma a los niños. Particularmente, la directora de este proyecto participa activamente del Comité “Prolog Education”, cuyo Chair es la Dra. Verónica Dahl asesora en esta propuesta (<https://prologyear.logicprogramming.org/Education.html>). Asimismo, parte del equipo está desarrollando líneas de investigación referidos a esta temática enmarcados en los Proyectos de Investigación “Tecnologías Semánticas para el Desarrollo de Agentes Inteligentes” (04/F020) y “Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación” (04/F022). Integrantes alumnos y docentes de este proyecto, están desarrollando su tesis en temáticas relacionadas a esta iniciativa. Se espera por una parte que el resultado de estas tesis aporten nuevas alternativas a la concreción de este tipo de experiencias. Por otro lado, el proyecto busca enriquecer la calidad de las tesis al proporcionar información para mejorar su consistencia, al mismo tiempo que fomenta la reusabilidad de los resultados obtenidos y la creación de una herramienta más efectiva y sólida. Durante 2022 y 2023 se desarrollaron experiencias educativas que promovieron el contacto temprano entre estudiantes de las escuelas primarias 125 y 180 con conceptos sobre programación lógica. Estas experiencias alentadoras se enfocaron en conceptos fundamentales del Pensa-

miento Computacional y Lógico, tales como la abstracción, la generalización, la descomposición y el razonamiento. (Video ilustrativo de una de las experiencias <https://www.instagram.com/reel/Cuz3GGKgii/?igshid=NjFiZTE0ZDQ0ZQ==>) Es interesante remarcar que estas experiencias fueron sociabilizadas en el “The Prolog Day Symposium” enmarcado en la celebración “2022: The Year of Prolog. Celebrating the 50th anniversary of Prolog”, ubicándose a nivel internacional como un punto de referencia. Asimismo, esta propuesta apuesta a contribuir a resolver la necesidad de despertar vocaciones en las disciplinas STEM y en Computación, motivando la incorporación de saberes en forma amena. Finalmente, la experiencia se plantea con el objetivo de favorecer la igualdad de género, a fin de empoderar a las niñas a elegir esta clase de profesiones.

### 3. Destinatarios

Docentes, estudiantes y comunidad educativa de las escuelas primarias públicas de la ciudad de Neuquén, que se enumeran a continuación:

**Escuela Primaria 183:** Maestros Neuquinos 1060

- 20 Estudiantes de 5to grado
- 3 Docentes de Primaria

**Escuela Primaria 125:** Roca 850

- 40 Estudiantes de 5to grado
- 3 Docentes de Primaria

**Escuela Primaria 180:** Tte. Gral. Pablo Ricchieri 999

- 30 Estudiantes de 4to y 5to grado
- 5 Docentes de Primaria

**Escuela Primaria 207:** Bahía Blanca 1100

- 40 Estudiantes de 5to grado
- 3 Docentes de Primaria

**Escuela Primaria 103:** Perticone 1841

- 30 Estudiantes de 5to grado
- 3 Docentes de Primaria

Las escuelas primarias públicas destinatarias se encuentran en diferentes barrios de la ciudad de Neuquén: Belgrano(2 escuelas), Confluencia, Área Centro Este, Área Centro Oeste.

Por otra parte, es destinataria de este proyecto la **Comunidad Internacional Prolog**, con la que compartiremos y socializaremos las experiencias realizadas y los recursos generados.

### 4. Resultados Esperados

Considerando los objetivos propuestos, esperamos lograr los siguientes recursos y acciones como resultado de este proyecto

- Diseño de dispositivos didácticos para favorecer aprendizajes sobre Programación Lógica, que facilitan la comprensión y el aprendizaje de conceptos y habilidades, tales como el razonamiento deductivo, la recursividad, la resolución de problemas y la creación de algoritmos.
- Creación de un repositorio digital las secuencias educativas desarrolladas que estará disponible en línea para su consulta y descarga por parte de docentes, estudiantes e investigadores interesados en el tema.
- Desarrollo de un prototipo de una herramienta para la generación aplicaciones basadas en Programación Lógica.
- Organización de ferias de ciencia donde se expongan y sociabilicen los resultados de las experiencias con la comunidad educativa de cada institución.
- Socialización de experiencias en la Comunidad Internacional Prolog.

## RA y RV aplicadas al resguardo del patrimonio cultural

Carina Fracchia<sup>1</sup> - Amadeo Laurín<sup>2</sup>

<sup>1</sup> Departamento de Computación Aplicada / Facultad de Informática  
Universidad Nacional del Comahue

<sup>2</sup> Asesor Museólogo Secretaría de Estado Cultura Río Negro y Subsecretaría de Cultura  
Municipalidad de Neuquén  
carina.fracchia@fi.uncoma.edu.ar  
amadeos.laurin@gmail.com

**Resumen.** Los avances continuos de la tecnología nos enfrentan diariamente al reto de articular e integrar adecuadamente el aprendizaje formal con las vivencias cotidianas e instancias socializadoras del entorno inmediato, percibidas como más vitales, profundas y reales.

Teniendo presente las posibilidades que ofrecen los museos como espacios de aprendizaje no formal, donde el público visitante puede estudiar, vincularse e interactuar con los objetos y demás elementos que allí se exhiben, nace en 2021 la iniciativa del proyecto de extensión “El museo va a las escuelas” y en 2022 “El museo vuelve a las escuelas”. En los mismos se trabaja con tecnologías visuales e interactivas, tales como la Realidad Aumentada (RA) y la Realidad Virtual (RV), buscando que la presentación y difusión del patrimonio sea comprensible para todo público, promoviendo la valorización de las sociedades del pasado por parte de la sociedad actual.

**Palabras clave:** Educación, Realidad Aumentada, Realidad Virtual, Museos, Patrimonio Cultural.

### 1 Introducción

El presente trabajo está enmarcado en los proyectos de extensión “El museo va a las escuelas” y “El Museo vuelve a las escuelas”, este último está en desarrollo en la actualidad. Participan docentes y estudiantes avanzados de la carrera de Licenciatura en Ciencias de la Computación (FAIF), profesores asesores de la Maestría en Enseñanza en Escenarios Digitales de la red AUSA (Asociación de Universidades Sur Andina), investigadores y trabajadores de Museos de la región, docentes y estudiantes de diferentes instituciones de los niveles primario y secundario de las provincias de Neuquén y Río Negro.

Con el término patrimonio se hace referencia al capital cultural de las sociedades contemporáneas, su importancia radica en la revalorización continua de las culturas, de las identidades, y los conocimientos que se buscan transmitir entre las generaciones. Esta construcción de un sentido de pertenencia, individual y colectivo,

ayuda a mantener la cohesión social y territorial. Desde un punto de vista económico y turístico, implica nuevos retos para su conservación, protección y valoración [1].

Son muchos los factores que entran en juego en el fomento de la sostenibilidad del patrimonio: “Registros e Inscripciones”, “Protección, Salvaguardia y Gestión” y “Transmisión y Movilización de Apoyos”.

En este último es donde nuestro proyecto se fundamenta, dado que abarca acciones dirigidas a sensibilizar al público en general sobre el valor y el sentido del patrimonio, así como sobre las amenazas potenciales, por medio de la señalética, los servicios de interpretación, las campañas de comunicación y las actividades educativas. Cubre también la participación del sector privado y de la sociedad civil en cuestiones relacionadas con el patrimonio. Se persigue la sensibilización mediante acciones encaminadas a promover el potencial educativo del patrimonio y su transmisión, buscando fomentar el conocimiento, el reconocimiento, el respeto y el fortalecimiento del patrimonio en la sociedad.

En este sentido, dos tecnologías que han sido estudiadas y exploradas en el marco de diferentes proyectos de investigación y extensión son la RA y RV, las mismas favorecen el desarrollo de escenarios digitales más visuales e interactivos, que pueden ser aplicados a una gran variedad de contextos, principalmente el educativo y el de patrimonio cultural. Mediante la RA se puede complementar el contexto real con información sintética generada a través de un dispositivo que debe contar con una cámara web. Esta adición de información permite fácilmente contribuir a ampliar lo presentado mediante diferentes medios, tales como texto, audio, imágenes, videos u objetos 3D [2]. En un desarrollo realizado para el Museo Gregorio Alvarez de nuestra ciudad se puede observar el potencial de la RA [3].

En el caso de la RV, si bien todo el escenario generado y presentado es sintético [4], el uso de imágenes o videos 360° constituye un valioso recurso para acercar los distintos museos a las instituciones educativas de los diferentes niveles. La experiencia realizada en 2022 con estudiantes de tercer año del secundario, donde se les suministró imágenes 360° del Museo Paraje Confluencia para la realización de un recorrido virtual, nos sorprendió tanto en su nivel de calidad como creatividad. Los productos resultantes mostraron cómo los estudiantes lograron trascender los muros físicos de la institución mencionada, creando nuevas salas de exhibición. Además se elaboraron paneles que permiten acceder a las producciones realizadas por los estudiantes, constituyendo una muestra itinerante de la institución que forma parte de la muestra permanente del museo, y que en breve recorrerá localidades del norte neuquino y diferentes instituciones educativas de la ciudad de Neuquén.

La experiencia lograda en estos años nos permitió observar la importancia de esta iniciativa con la cual no sólo se presenta el museo a profesores y estudiantes, sino que esta invitación se realiza a todo el entorno familiar, que es contagiado para visitar de manera presencial y vivenciar las instalaciones de los mismos.

En la actualidad el desafío es acercar las muestras permanentes de los museos "Museo Nacional de Bellas Artes", "Museo Histórico del Riego" y "Museo Provincial Carlos Ameghino". El primer desafío se plantea en la diversa tipología de Museos que pretendemos abarcar: Bellas Artes, Histórico Regional y Paleontológico. Para ello se prevé el desarrollo de modelos 3D, software RA y RV que permitan, desde un punto

de vista didáctico, aprovechar las posibilidades de interacción, interpretación e intermediación que ofrecen este tipo de recursos para atraer el interés del público hacia el patrimonio, introduciendo mejores formas de comunicarse con él durante la visita.

La firma de un convenio entre la Municipalidad de Neuquén y la Facultad de Informática permitió trabajar de manera conjunta para acercar los museos de nuestra región a las instituciones educativas, con el fin de que se pudiera ayudar a percibirlo como lo que es, una institución dinámica de puertas abiertas que no sólo guarda y conserva, sino que también produce, expone y vincula de manera permanente con la comunidad.

## 2 Resultados esperados

Dentro de los objetivos perseguidos se encuentran el diseño de recursos TIC que permitan acercar la muestra permanente de los museos de la región a las escuelas y sociedad en general; y contribuir al resguardo del patrimonio cultural alojado en los museos mediante el uso de recursos TIC, y las tecnologías Realidad Aumentada y Realidad Virtual.

Se prevé el desarrollo de variados recursos TIC, software RA y RV, visitas con el personal de los museos a las diferentes instituciones educativas, y la organización de una muestra general de los productos obtenidos, en la denominada “noche de los museos”, evento que se realizará el día Sábado 02 de Diciembre del corriente año, en los Museos dependientes de la Secretaría de Cultura Capacitación y Empleo - Municipalidad de Neuquén. En el mismo los proyectos de extensión “El museo va a las escuelas” y “El Museo vuelve a las escuelas” se presentarán en un espacio dentro del Museo Nacional de Bellas Artes. Además se continuará con la transferencia de resultados por medio de la presentación de artículos en congresos, jornadas, entre otros y su publicación en revistas nacionales e internacionales.

Cabe aclarar que los recursos y softwares desarrollados quedarán disponibles y públicos, para su uso y reutilización, según la licencia que se especifique en forma conjunta con los diferentes organismos intervinientes.

## Referencias

1. Organización de las Naciones Unidas para la Educación, la Ciencia y La Cultura (Unesco). (2012). Indicadores UNESCO de cultura para el desarrollo: manual metodológico.
2. Azuma, R. T. (2019). The road to ubiquitous consumer augmented reality systems. *Human Behavior and Emerging Technologies*, 1(1), 26-32.
3. <https://norbertolarrosa.github.io/>
4. Sherman WR, Craig AB (2003). *Understanding Virtual Reality: Interface, Application, and Design*. Amsterdam:

# Proyecto de Extensión: FaIComm: Charlando Sobre Informática

Equipo Control Z

Facultad de Informática, Universidad Nacional del Comahue  
<https://controlz.fi.uncoma.edu.ar>  
[controlz@fi.uncoma.edu.ar](mailto:controlz@fi.uncoma.edu.ar)

**Resumen** Teniendo presente la inmersión tecnológica que tenemos en nuestro día a día y aprovechando las posibilidades y beneficios de la radio, presentamos una alternativa de estrategia comunicativa, donde la comunidad de la facultad de informática fueran los protagonistas. En el presente artículo compartimos nuestro proyecto de extensión para crear vías de expresión y comunicación con distintos actores de la sociedad.

**Keywords:** Comunicación · Recursos audiovisuales · Actualidad informática · Noticias FaI.

## 1 Fundamentos de nuestro proyecto

En nuestra vida cotidiana utilizamos artefactos tecnológicos todo el tiempo. Estos atraviesan nuestro día a día ya sea en el trabajo como en el ocio, cambiando nuestras costumbres y hábitos. El tipo de tarea puede ser en apariencia simple, como enviar mensajes de texto, o compleja, como hacer uso de un sistema inteligente. Sin embargo, es necesario conocer los conceptos y el funcionamiento de estos artefactos para promover un debate sobre temas que nos interesan socialmente: respecto a su uso responsable, que preserve nuestra privacidad y seguridad, que se converse acerca de la manipulación ética de datos, que se compartan curiosidades y formas de utilización, que se conozca el potencial y los posibles efectos, que se enseñe y se estudie el software y hardware, entre otros.

En muchos ámbitos sociales también se encuentra influenciada la informática. Por ejemplo, en la currícula de las escuelas primarias y secundarias incluye contenido de seguridad informática, robótica e inteligencia artificial, entre otros; en el mundo profesional e industrial se realizan capacitaciones para el uso de herramientas informáticas, siendo hasta componentes esenciales para su trabajo diario.

La facultad no es ajena a esto, todo lo contrario: posee múltiples proyectos de investigación y extensión con amplio contacto y colaboración con los diferentes ámbitos sociales y laborales, aportando un intercambio rico en capacitación e información. Por ejemplo, proyectos con interacción con escuelas secundarias

## 2 Equipo Control Z

para promover el internet seguro, la programación y creación de aplicaciones<sup>1</sup>, la robótica<sup>2</sup>, la realidad aumentada para museos<sup>3</sup>, entre otros.

Por ello, integrantes de la comunidad de la Facultad de Informática creamos un espacio dinámico y entretenido donde conversar para conocer, entender y compartir el modo en que la tecnología nos pone a vivir. Comunicando las distintas actividades universitarias como así también promoviendo el debate acerca de los conceptos, de las nuevas tecnologías y el uso ético de las mismas.

## 2 Objetivos y Metodología

Es por esto, que este proyecto propone cuatro ejes para generar una vinculación de doble vía entre la facultad y las instituciones públicas con alcance social: Establecer vínculos, promover la palabra, producción y promoción de recursos audio-visuales, y la realización de talleres y conversatorios. Todos estos objetivos intentan fomentar la curiosidad tecnológica del público en general mientras se comunica y vincula la comunidad FaI y sus distintos proyectos.

### 2.1 Ejes metodológicos

Como metodología de trabajo se empleó un modelo constructivista, en el cual los integrantes de nuestro equipo participan activamente para la formación de esta propuesta de comunicación. El proyecto se compone de cuatro ejes bien diferenciados.

El primer eje, busca establecer vínculos promoviendo el intercambio de saberes tanto desde la facultad hacia las instituciones, como de estas hacia la facultad. Conocimientos más técnicos de, por ejemplo, conceptos, metodologías, tecnologías y su funcionamiento ayudan a las instituciones y sus integrantes a comprender las distintas noticias y novedades donde la informática puede encontrarse embebida en ellas. Además, conceptos y prácticas de comunicación y socialización promueven a la facultad a enriquecer sus interacciones con dicha institución, y con el público de estas, fomentando y compartiendo nuestros proyectos y producciones. La radiodifusión, los talleres y conversatorios es una forma de llevar a cabo este eje, promoviendo la vinculación e interacción a medida que se realizan actividades entre integrantes de las instituciones y de la facultad.

El segundo eje, consiste en compartir la palabra de nuestra facultad acerca de nuestros proyectos, propuestas y novedades, como así también de nuestros conocimientos. Esto, no suele ser agenda habitual de las instituciones con marcada influencia social, pero aún así, no deja de ser de interés para el público en general. Para ello, la elaboración de audiovisuales promueve la comunicación y el debate, permitiendo llevar nuestro mensaje por los diferentes medios de difusión y así remitir la palabra a una audiencia general.

<sup>1</sup> <https://incuba.fi.uncoma.edu.ar/webinar/appinventor/>

<sup>2</sup> <http://incuba.fi.uncoma.edu.ar/franklab>

<sup>3</sup> <https://www.fi.uncoma.edu.ar/index.php/prensa/feria-it-joven-en-tecnopolis/>

El tercer eje, es la producción y promoción de recursos audio-visuales como medio creativo y para compartir información y actualidad. Estos recursos pueden ser producidos tanto por estudiantes, docentes, no docentes y graduados de nuestra facultad como por otras personas de la comunidad en colaboración con nuestro proyecto. Por ejemplo, estudiantes de escuelas secundarias<sup>4</sup> profesionales y trabajadores de la radio, personas interesadas en compartir experiencias de la informática, etc. Promueven el debate y la comunicación de información a la sociedad en general. También, se de gran utilidad para complementar las actividades de los otros ejes.

Por último, el cuarto eje conlleva la realización de talleres y conversatorios por parte de los integrantes del proyecto para intercambiar conocimientos y promover tópicos que sean de su interés, ampliando la vinculación por medio de una interacción más personal y no solo por medio de recursos digitales.

### 3 Experiencias y resultados

Actualmente, la Facultad de Informática está desarrollando un programa de radio denominado de una hora “Control Z” emitido por la FM 103.7 Radio UNCo-CALF<sup>5</sup>. Su contenido se produce en un ambiente colaborativo recibiendo novedades e información de los distintos claustros de nuestra facultad.

Las personas participantes del proyecto, comparten el objetivo en común de producir y realizar cada emisión del programa, aprendiendo el proceso que se requiere para ello: la producción del guión y del contenido, la elaboración de recursos auditivos, la preparación y realización de entrevistas, la interacción con el personal de radio (operadores, directores de radio, productores, etc.), conducción del programa, la gestión de redes sociales y la consideración de la identidad gráfica, etc.

Es relevante mencionar que estas actividades y la interacción con el público general les propone una experiencia innovadora y satisfactoria para las personas participantes de la producción y la conducción del programa, particularmente estudiantes, incentivando sus aportes, su creatividad y su predisposición para estudiar e investigar temáticas de informática y de la universidad y para intercambiar su experiencias ante la sociedad.

La integración de las TICs como recurso de investigación y formación junto con la dinámica radial favoreció el vínculo con la comunidad universitaria y con otras comunidades como centros educativos para intercambios de experiencias escolares, y grupos de especialistas informáticos de la zona, entre otros.

---

<sup>4</sup> Ejemplos de spots creados en otros proyectos de la FaI: EESISeg de los cuales Control Z puede utilizar y colaborar <https://internetseguro.fi.uncoma.edu.ar/categoria-producto/tipo-de-recurso/spot-radio/>), visitado el día 30 de octubre del 2023.

<sup>5</sup> <https://universidadcalf.com/> visitado el 30 de octubre del 2023.

## 4 Equipo Control Z

**Integrantes del Equipo Control Z**

El equipo de Control Z está compuesto por integrantes de diferentes claustros de la Facultad de Informática de la Universidad Nacional del Comahue.

- Acosta, Ian Franco Matías. Integrante del claustro Estudiante de la Facultad de Informática.
- Avellaneda, Ángel. Integrante del claustro Estudiante de la Facultad de Informática.
- Bruno, Leo. Integrante del claustro Estudiante de la Facultad de Informática.
- Cañibano, Rodrigo. Integrante del claustro Docente de la Facultad de Informática.
- Galaz, Lara Virginia. Integrante del claustro Estudiante de la Facultad de Informática.
- Gimenez, Christian Nelson. Integrante del claustro Docente de la Facultad de Informática.
- Kogan, Pablo. Integrante del claustro Docente de la Facultad de Informática.
- Mazalu, Rafaela. Integrante del claustro Docente de la Facultad de Informática.
- Monserrat Vidal, María Elvira. Integrante del claustro Estudiante de la Facultad de Informática.
- Morán, Marina. Integrante del claustro Docente de la Facultad de Informática.
- Rodríguez, Jorge Pablo. Integrante del claustro Docente de la Facultad de Informática.
- Rotter, Valeria. Integrante del claustro No docente de la Facultad de Informática.
- Villarroel, Valentina. Integrante del claustro Estudiante de la Facultad de Informática.

# Conectados con la Escuela

## Proyecto de Extensión

Directoras: Ingrid Godoy, Susana Parra,  
Integrantes: Alejandro Antezana De La Rivera, Alejandro Jerson Mora  
Vazquez, Daniela Zacharias, Federico Nahuel Mamani, Ian Franco Matias  
Acosta, Ignacio Ciruzzi, Jonathan Maximiliano Cabrera, Jorge Pablo  
Rodriguez, Leo Bruno, Luis German Coralle, María Elvira Monserrat Vidal,  
Pablo Kogan, Rafael Ignacio Zurita, Rodrigo Esequiel Piersigilli,  
Ruben Ramirez, and Valeria Zoratto

Facultad de Informática  
Universidad Nacional del Comahue  
{ingrid.godoy, susana.parra}@fi.uncoma.edu.ar

**Keywords:** Educación en Ciencias de la Computación · Escuela Secundaria · Recursos Educativos Desenchufados.

## 1 Fundamentación

La iniciativa de extensión universitaria propone acompañar las acciones enfocadas en el mejoramiento de la enseñanza de las Ciencias de la Computación y la construcción de ciudadanía en las escuelas secundarias. En este marco se procura ampliar y profundizar las vinculaciones establecidas, entre un conjunto de escuelas, el Observatorio Electoral de la UNComa y la Facultad, para juntos construir dispositivos didácticos novedosos que ayuden a estudiantes de escuelas secundarias a conectarse con conceptos fundamentales de la disciplina y les permitan participar informadamente de los procesos democráticos.

En este sentido se proyecta concretar acciones de extensión en el campo de la enseñanza de la programación, la promoción del uso responsable de las tecnologías y la ciudadanía que resulten convocantes y significativas para las adolescencias. Las metodologías ágiles tienen la posibilidad de configurar experiencias especialmente atractivas como soporte metodológico para conectar a las adolescencias con las Ciencias de la Computación.

Los torneos de programación, el desarrollo de artefactos computacionales y la programación de drones y robots configuran escenarios didáctico-disciplinares para la enseñanza y el aprendizaje de la computación en los que convergen el trabajo colaborativo, el desarrollo de la creatividad y el pensamiento lógico. Tienen amplias posibilidades de acercar a la población estudiantil de la escuela secundaria a los conceptos fundamentales de las ciencias y en particular de la informática.

En este sentido la construcción de aplicaciones móviles, permite a los estudiantes involucrarse en la resolución de problemas auténticos y experiencias reales

que les ayudarán a adquirir habilidades de trabajo en grupo y a desarrollar su imaginación y creatividad.

Por otra parte, existe un consenso amplio acerca de la eficacia de los juegos como herramienta de aprendizaje. Los torneos de programación con estudiantes secundarios ayudan a mejorar la participación e interacción de los estudiantes ofreciendo un entorno de aprendizaje entretenido.

La programación de drones y robots es una estrategia ampliamente difundida y seleccionada con frecuencia como entorno metodológico para el armado de cursos introductorios al aprendizaje de las Ciencias de la Computación destinados a jóvenes sin formación previa en el campo disciplinar.

Este tipo de actividades ayudan a asociar el conocimiento disciplinar con aplicaciones que los y las estudiantes pueden construir y modificar. A través de estas experiencias se conectan con diversos conceptos relacionados con la computación, como algoritmos y programación e ingeniería de software.

La divulgación y formación en aspectos relacionados a la Internet Segura y la Seguridad Informática se vuelve cada vez más importante y necesaria a los efectos de que las personas que acceden a los diferentes servicios ofrecidos por Internet sepan cuáles son los riesgos y cómo protegerse de los mismos para minimizar los problemas derivados de la inseguridad.

El proyecto busca promover el uso responsable de las tecnologías y ubicar a los estudiantes secundarios en el rol de promotores a partir de la elaboración y difusión de piezas gráficas y audiovisuales que se integran a un repositorio digital de acceso público.

En el contexto de los cuarenta años de democracia, se propone trabajar sobre la construcción de ciudadanía acercando a las escuelas experiencias que permiten a las adolescencias experimentar diferentes sistemas de votación para formarse en el ejercicio ciudadano de la emisión del voto y para mejorar las posibilidades de participar en forma informada de los debates de la sociedad.

La infraestructura tecnológica en la institución puede ubicarse como uno de los principales obstáculos en la implementación de la educación en escenarios digitales. Las necesidades de docentes y estudiantes y la disponibilidad de recursos, ya sean de hardware, software o conectividad pueden constituir una fuente de frustración. En este sentido se programan acciones tendientes a lograr la infraestructura TIC requerida para enseñar, colaborando con la institución en la instalación de las piezas de software necesarias para enseñar los conceptos propuestos.

La participación de la computación en las propuestas de enseñanza para la escuela secundaria aumentó con el nuevo diseño curricular para la escuela secundaria. La propuesta se estructura a partir de la concreción de actividades breves que se organizan en tres tramos: formación docente, actividad con estudiantes secundarios, evaluación de experiencias con docentes. Las temáticas a desarrollar son:

Internet Seguro: Estudiantes de la escuela secundaria trabajan en grupos en la elaboración de piezas destinadas a divulgar buenas prácticas acerca del uso responsable de internet. Los receptores de la divulgación son principalmente

estudiantes y docentes de escuelas primarias. Las piezas construidas de suman Repositorio Digital sobre Internet Seguro de la Facultad de Informática de la Universidad Nacional del Comahue, EESISeg — Repositorio

**Desarrollo de aplicaciones Móviles:** Los estudiantes se aproximan a los conceptos fundamentales de la programación a través de la construcción de aplicaciones móviles. Utilizando el entorno de desarrollo MIT appInventor, una herramienta simple y amigable que no requiere conocimientos previos sobre programación, los estudiantes producen aplicaciones para dispositivos Android.

**Torneos de Programación:** Se presenta este taller como una forma de promover la comprensión de conceptos relacionados a la programación. Los estudiantes organizados en pequeños grupos compiten contra el reloj en un juego donde además de la lógica y el empleo de estrategias para resolver problemas, es importante la colaboración y comunicación entre los integrantes de cada equipo.

**Sistemas de Votación - Construcción de Ciudadanía:** La adolescencia, estudiantes de las escuelas participantes, se ponen en contacto con los debates actuales acerca de los procesos electorales siendo parte de una recreación de una jornada electoral que les permite explorar los diferentes sistemas de votación: Boleta Partidaria, Boleta Única de Papel y Voto Electrónico.

**Infraestructura tecnológica para enseñar:** Se programan intervenciones tendientes a lograr la infraestructura TIC requerida para enseñar, colaborando con la institución en la instalación de las piezas de software necesarias para enseñar los conceptos propuestos.

## 2 Objetivo

Se trata de una iniciativa extensionista que tiene como objetivos generales, tendientes a colaborar y acompañar a las escuelas en los procesos de mejoramiento de la enseñanza de la computación: Contribuir al mejoramiento de la enseñanza de la computación en las escuelas secundarias. Contribuir a la formación docente de los profesores y profesoras de los espacios curriculares destinados a Informática de las escuelas secundarias. Contribuir al mejoramiento de la infraestructura TIC Escolar. Contribuir a la construcción de ciudadanía.

## 3 Población Destinataria

La población destinataria se localiza en las provincias de Neuquén y Río Negro e integran las comunidades educativas de un grupo de Escuelas Secundarias de la Región. Estudiantes y docentes de las escuelas secundarias. Comunidades Educativas.

## 4 Resultados Esperados

Desarrollar experiencias educativas que exponen conceptos de las Ciencias de la Computación a estudiantes secundarios. Desarrollar experiencias educativas

que permiten reflexionar sobre los procesos democráticos. Producir y almacenar en un repositorio digital contenido destinado al uso responsable de Internet. Desarrollar soluciones TIC

# Análisis forense de pericias médicos-legales asistidas por sistemas inteligentes: Preparación de Datos Proyecto de Extensión

Directores: Sandra Emilce Roger, German Alejandro Braun,  
Integrantes: Carlos Campos Fuentes, Manuel Latorre, Santiago Villarroel,  
Gabriel Omar Jerez, Cristian Martin Vincenzini, Fabio Marin Apablaza,  
Rafaela Jesus Maricela Mazalu, Romina Tealdi, Alejandro Jerson Mora  
Vazquez, Federico Solorza, Gabriela Noemi Aranda, Javier Ignacio Rey,  
Alejandro Daniel Cozzarin, Matias Ezequiel Oyarzun, and Luis Eduardo  
Garcia Carrillo

Facultad de Informática  
Universidad Nacional del Comahue  
{roger, german.braun}@fi.uncoma.edu.ar

**Keywords:** Pericias Medicos-Legales. Aprendizaje Profundo. Sistemas Inteligentes. Violencia. Vulnerabilidad. Objetividad.

## 1 Fundamentación

El análisis forense juega un papel fundamental en los casos jurídicos que involucran cuestiones médicas y de salud. Su importancia radica en que proporciona una evaluación imparcial basada en conocimientos médicos-científicos. Asimismo, las pericias judiciales forenses (PJF) colaboran en valorar daños, determinación de responsabilidades e interpretación de evidencia médica. Durante los 40 años de democracia en Argentina, el rol de las PJF ha adquirido una relevancia crítica en el sistema de justicia del país. La consolidación del Estado de Derecho y la garantía de los derechos humanos han sido pilares fundamentales de este período histórico, y las PJF han desempeñado un papel esencial para alcanzar estos objetivos, siendo clave en la investigación de crímenes y violaciones a los derechos humanos, contribuyendo a la búsqueda de la verdad, la identificación de responsables y la obtención de pruebas sólidas para asegurar juicios justos y transparentes. Además, en casos de violencia de género, las PJF han sido instrumentales para entender la magnitud del problema, comprender sus implicaciones psicosociales y, sobre todo, colaborar en brindar una respuesta adecuada y sensible a las víctimas. A partir del 2020, el Cuerpo Médico Forense del Poder Judicial de Neuquén (CMFPJN) ha recopilado una gran cantidad de datos (alrededor de 18000 imágenes y 1800 informes), sobre pericias judiciales realizadas. Estos datos crudos no están estructurados, estandarizados, ni digitalizados de manera adecuada. Particularmente, los informes presentan un vocabulario específico del dominio médico, aunque no normalizado. Actualmente, la

facilidad en la recolección y almacenamiento de los datos, junto con tecnologías e infraestructuras que permiten manipularlos, ha motivado la creación de sistemas de inteligencia de datos a grandes escalas. Un punto de partida para el desarrollo de estos sistemas es contar con volúmenes de datos con el formato adecuado para su procesamiento. Así, surge la necesidad de transformar y procesar los datos almacenados por el CMFPJN, para que puedan ser analizados e interpretados, a fin de producir información útil y valiosa, que permita la toma de decisiones respecto de las políticas gubernamentales. Este marco es el punto de partida de un trabajo de vinculación con el CMFPJN. La presente propuesta consiste en preparar los datos y almacenarlos adecuadamente, generando una base de datos y conocimiento con la información existente en las áreas PJF de Neuquén. En este sentido, se trabajará en forma colaborativa identificando y estableciendo estándares en cuanto al material fotográfico; analizando el vocabulario y terminología empleadas en los informes y realizando un proceso de modelado ontológico de la misma; seleccionando las herramientas que se utilizarán en el desarrollo y que faciliten la digitalización de todos los datos con los que se cuenta hasta el momento; y determinando los requerimientos tecnológicos de equipamientos que serán utilizados, no solo de almacenamiento sino también para el procesamiento futuro de estos datos. El equipo de trabajo de esta iniciativa cuenta con profesionales de diferentes disciplinas incluyendo a Médicos Forenses, Especialistas en Fotografía y profesionales de Ciencias de la Computación, logrando un grupo interdisciplinar la cual representa una poderosa alianza, ya que combina la experiencia y conocimientos de estos campos de conocimiento. La Inteligencia Artificial (IA) ha experimentado avances significativos en la última década, revolucionando diversos campos saber. Ha demostrado su capacidad para manipular, comprender, procesar, tanto lenguaje humano como imágenes. Con el avance tecnológico de los modelos de Aprendizajes Automáticos, se ha ampliado el espectro de aplicaciones de los modelos de aprendizaje automático llegando a límites impensados. El proyecto de investigación “Tecnologías Semánticas para el Desarrollo de Agentes Inteligentes-04/F020 (Res Cd: 0092/2021), que es codirigido por la directora de esta propuesta desarrolla líneas de investigación en los campos del modelado de los agentes inteligentes, modelado ontológico, procesamiento de lenguaje natural, entre otros, conocimientos afines que serán de utilidad, y que se continuará ampliando con las investigaciones que devengan con este trabajo. En este contexto, se espera que estudiantes de grado y de posgrado integrantes de esta iniciativa de extensión desarrollen sus tesis en estas temáticas, desarrollando y/o adaptando enfoques y técnicas de IA e implementando herramientas. Es interesante remarcar que esta propuesta persigue un impacto práctico y tangible al hacer posible la aplicación de estas herramientas y técnicas en casos concretos. Al implementar y probar los modelos desarrollados en situaciones reales, se podrá evaluar su rendimiento y utilidad en el dominio, lo que proporcionará una valiosa retroalimentación para mejorar la eficacia y precisión de éstos. Es interés de ambas partes participantes de esta propuesta la vinculación a largo plazo continuando la colaboración una vez que finalice el proceso de la preparación de los datos. Se plantea continuar con el uso de estos

datos en sistemas inteligentes, produciendo así nuevo conocimiento, definiendo, de esta manera, una nueva dimensión a considerar para resolver problemas reales y complejos, tal como ayudar en la asistencia de los requerimientos que surjan a los profesionales que trabajan en el cuerpo forense del poder judicial. Es fundamental fundar la toma de decisiones en los dictámenes forenses con la mayor evidencia y aval de pares posible, evitando, en la mayor medida, los sesgos en la evaluación forense (cognitivo, de proximidad, etc.) y respetando las normativas vigentes a nivel nacional, como los acuerdos a nivel internacional. Así, la Facultad de Informática, potenciará su lugar estratégico de asesor técnico y ético para quienes tienen responsabilidad de generar nuevas políticas asociadas a las problemáticas judicializadas de sectores vulnerables. Este proyecto representa el comienzo de una colaboración estrecha con los médicos forenses del Poder Judicial de la provincia de Neuquén. Se tiene la expectativa de que en fases posteriores se expanda este trabajo hacia las provincias de Córdoba y Buenos Aires, creando así una red de colaboración que permitirá abordar de manera más integral las cuestiones médicas-legales en casos forenses.

## **2 Objetivo**

Estructurar y digitalizar los datos disponibles por el cuerpo médico forense del Poder Judicial de la provincia del Neuquén relacionados a pericias médicas judiciales, los cuales incluyen fotos e informes judiciales. Establecer un sistema de clasificación experta normalizada de un conjunto de estas pericias

## **3 Población Destinataria**

La población destinataria se localiza en las provincias de Neuquén. Cuerpo Médico Forense del Poder Judicial. Personal interviniente en el proceso de peritaje. Indirectamente la comunidad que requieran peritaje.

## **4 Resultados Esperados**

Se esperar, por un lado, contar con un repositorio de datos digitalizados mediante el análisis, estandarización, limpieza y normalización de la información disponible por el Poder Judicial de Neuquén. Por otro lado, se espera contar con un conjunto de pericias forense-legales clasificado por un conjunto de peritos expertos.

## **ALERTAR: un sistema informático inteligente y resiliente para la vigilancia y detección temprana de severidad de pacientes en unidades de cuidados no intensivos**

Docentes: Javier Balladini<sup>1</sup>, Claudio Zanellato<sup>1</sup>, Marina Morán<sup>1</sup>, Rodrigo Cañibano<sup>1</sup>.

Graduados: Cristina Orlandi<sup>2</sup>, Darío Semenzato<sup>1</sup>.

Estudiantes: Mariano Conchillo<sup>1</sup>, Héctor Mella<sup>1</sup>, Manuel Latorre<sup>1</sup>, Adriano Lusso<sup>1</sup>, Dana Contreras<sup>1</sup>, Aldana Vega<sup>1</sup>, Lara Acuña<sup>1</sup>, Santino Castagno<sup>1</sup>.

No-docente: Cristian Ilabaca<sup>1</sup>.

<sup>1</sup> Facultad de Informática, Universidad Nacional del Comahue  
javier.balladini@fi.uncoma.edu.ar

<sup>2</sup> Hospital Área Programa General Roca

**Resumen.** El proyecto propone desarrollar la aplicación ALERTAR, destinada a la vigilancia y detección temprana de gravedad en unidades de cuidados no intensivos. Habiendo diseñado la arquitectura general del sistema en el marco de un proyecto de investigación, en este proyecto de extensión se espera completar su implementación y desplegarlo en el hospital de General Roca, provincia de Río Negro. ALERTAR optimiza recursos humanos adaptando los controles de enfermería según el riesgo de los pacientes, y notifica los cambios de gravedad. Se abordan cuestiones técnicas complejas para garantizar la seguridad de los datos y la resiliencia del sistema, contribuyendo al avance de soluciones informáticas en el campo de la salud. El impacto territorial es significativo al mejorar la atención médica, promover la equidad en el acceso a la salud (especialmente en áreas con escasos recursos humanos especializados) y fomentar la investigación interdisciplinaria y la formación de estudiantes de informática en el desarrollo de aplicaciones críticas. Los resultados esperados incluyen la mejora de la calidad de la atención médica, la optimización de recursos humanos y datos clínicos para investigaciones.

**Palabras claves:** alertas tempranas de severidad, vigilancia de enfermería, inteligencia artificial.

### **1 Contexto**

El trabajo de extensión aquí presentado corresponde al proyecto de extensión titulado “ALERTAR: un sistema informático inteligente y resiliente para la vigilancia y detección temprana de severidad de pacientes en unidades de cuidados no intensivos”, con inicio el 20/10/2023 y finalización el 19/10/2024, aprobado por Ordenanza 369/23 de la Universidad Nacional del Comahue (UNComa). El destinatario del

proyecto es el Hospital Área Programa General Roca, de la provincia de Río Negro. Cristina Orlandi, médica intensivista del hospital, y presidente de la Sociedad Argentina de Terapia Intensiva, es integrante del proyecto.

Debido a la complejidad del sistema ALERTAR, el mismo conforma una de las líneas del proyecto de investigación 04/F017 "Cómputo paralelo y distribuido: métricas de rendimiento, aplicaciones de big data e inteligencia artificial", financiado por la misma universidad, con inicio el 01/01/2021 y finalización el 31/12/2024, acreditado por el Ministerio de Educación de Argentina.

## **2 Descripción del problema**

El triage es un método que clasifica a los pacientes en diferentes niveles de severidad o riesgo de adquirir una determinada condición. Su aplicación en una sala de internación general permite mejorar el rendimiento del personal, al disminuir los controles en pacientes menos graves e incrementarlos en los más graves, y contribuir a la reducción de la mortalidad inesperada.

Si bien el triaje se puede realizar mediante cálculos manuales, este enfoque es propenso a errores y puede añadir una carga significativa de trabajo al personal de salud. En presencia de una historia clínica digital, surge la oportunidad de integrar el triaje en esa plataforma. Sin embargo, las historias clínicas digitales convencionales, cuyo objetivo principal es documental, podrían no ser adecuadas para implementar eficazmente una solución que, a partir del resultado del triaje, guíe o dirija el proceso de atención de pacientes en los múltiples aspectos requeridos. Incluso, si fuera posible, la solución no sería portable a otros sistemas de historias clínicas digitales. En consecuencia, una mejor alternativa es crear una aplicación independiente y específica, que pueda interoperar con cualquier historia clínica digital y otros sistemas de salud. Así, la solución podría ser fácilmente adoptada por cualquier institución.

## **3 Propuesta**

En este proyecto proponemos avanzar en el desarrollo de la aplicación informática ALERTAR, destinada a la vigilancia y emisión de alertas tempranas de gravedad en unidades de cuidados no intensivos, y ponerla en marcha en el hospital destinatario. ALERTAR es sucesora de nuestra aplicación COVINDEX [1], que implementa un modelo computacional de alerta temprana para COVID-19, que clasifica a los pacientes en 4 niveles de riesgo, siendo el riesgo una predicción de la gravedad que un paciente podría tener transcurridas 24 horas. ALERTAR es una generalización de COVINDEX, en el sentido de que está diseñada para permitir el soporte de múltiples enfermedades. Por cada paciente se calculará mínimamente el National Early Warning Score 2 [2], y posiblemente algún triaje más específico de la enfermedad o tipo de

enfermedad que lo afecta (COVID-19, enfermedad neurológica, etc.), cuando el mismo se encuentre disponible.

ALERTAR permitirá dirigir el proceso de vigilancia de pacientes mediante alertas visuales y sonoras que el personal de enfermería recibe en sus dispositivos móviles, indicando la proximidad o retraso de los controles, calculados según el nivel de riesgo del paciente. Además, ALERTAR permitirá una rápida respuesta del equipo médico al emitir alertas tempranas de cambios del nivel de riesgo en sus dispositivos móviles, y facilitar la evaluación de los pacientes por contar con un registro electrónico de datos clínicos específico del tipo de morbilidad.

Considerando que el tipo de servicio que intenta brindar ALERTAR es de carácter crítico por su incidencia en la vigilancia de los pacientes, la aplicación informática debe garantizar la seguridad de los datos y la resiliencia ante fallos de componentes y sus comunicaciones. Además, es importante que el alcance del sistema sea multihospitalario para favorecer la recolección de una mayor cantidad de datos clínicos, útiles para formular o calibrar sistemas de alerta temprana sobre la base de técnicas de aprendizaje automático. A fin de lograr una solución multihospitalaria, es necesario simplificar la administración del sistema, por lo que el único equipamiento requerido de los hospitales son dispositivos móviles (y no servidores estándares). Finalmente, otro aspecto relevante a considerar es la adaptabilidad del sistema para soportar nuevas enfermedades.

A continuación se presentan los objetivos específicos del proyecto.

### 3.1 Objetivos específicos

En el cuadro 1 se describen los objetivos específicos del proyecto y sus metas.

**Cuadro 1.** Objetivos específicos y sus metas

Descripción del objetivo	Meta
Terminar de implementar funcionalidad inconclusa.	Completar la implementación de interfaces de usuario de enfermedades, y el mecanismo de resiliencia ante fallos de dispositivos y comunicaciones.
Implementar nueva funcionalidad.	Definir e implementar aspectos de interoperabilidad con otros sistemas. Implementar el mecanismo de autenticación y seguridad de datos. Implementar el mecanismo de auditoría.
Detección y corrección de errores.	Disponer del sistema validado por el personal de salud.
Puesta en marcha del sistema de producción.	Sistema desplegado. Personal capacitado en el uso del sistema.

#### 4 Impacto territorial

El proyecto intenta tener un fuerte impacto en los siguientes aspectos:

- Optimización de la vigilancia de enfermería: Se busca aumentar la eficiencia del proceso de vigilancia de enfermería al ajustar los controles en pacientes según su riesgo, liberando recursos para casos más graves.
- Aplicabilidad en diversos contextos: ALERTAR no solo se enfoca en crisis como pandemias, sino que también mejora la atención médica diaria.
- Unificación de criterios de atención y equidad: La aplicación unifica criterios de atención, incrementando aún más su beneficio al aplicarse en áreas con escasez de recursos humanos especializados, y promoviendo la equidad en la atención.
- Aspectos técnicos y de investigación: Se abordan problemas técnicos tradicionales y otros complejos que requieren ser tratados en un proyecto complementario de investigación científica. Así, se contribuye al desarrollo de soluciones avanzadas en el campo de la informática y la salud.
- Colaboración interdisciplinaria para alcanzar soluciones efectivas: Expertos en informática, medicina y enfermería colaboran para abordar las necesidades reales del sector salud de manera efectiva.
- Formación de estudiantes: La participación de estudiantes permite formar profesionales en el campo de la tecnología aplicada a la salud y al desarrollo de aplicaciones críticas.

#### 5 Resultados esperados del uso de ALERTAR

Los resultados que se esperan obtener al utilizar el sistema ALERTAR comprenden:

- Mejora de la calidad de los cuidados de los pacientes y reducción de la mortalidad inesperada.
- Optimización de recursos humanos mediante la organización, coordinación y simplificación del trabajo de médicos y enfermeros.
- Información para la toma de decisiones organizacionales.
- Datos clínicos para investigaciones.
- Permitir auditoría médica.

#### Referencias

1. Rodrigo Cañibano, Santino Castagno, Mariano Conchillo, Guillermo Chiarotto, Claudia Rozas, Claudio Zanellato, Cristina Orlandi, Javier Balladini (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). <http://sedici.unlp.edu.ar/handle/10915/140661>
2. Royal College of Physicians. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. Updated report of a working party. London: RCP, 2017.

# Reduciendo Brechas Proyecto de Extensión

Leo Bruno, Ian Acosta, Valentina Villarroel, Jonathan Maximiliano Cabrera,  
Valeria Rotter y Jorge Rodríguez

Facultad de Informática, Universidad Nacional del Comahue  
{leo.bruno, ian.acosta, valentina.villarroel, jonathan.cabrera}  
@est.fi.uncoma.edu.ar  
{valerotter, j.rodrig}@fi.uncoma.edu.ar

**Resumen** En el presente artículos compartimos nuestro proyecto de extensión el cual se centra en ofrecer talleres a escuelas secundarias con el objetivo de fomentar el interés y la adopción de nuevas tecnologías. A través de estas iniciativas, hemos trabajado en estrecha colaboración con el cuerpo docente y los estudiantes para enriquecer su formación y fomentar un mayor compromiso con la educación y la innovación.

**Keywords:** Talleres · Vinculación con Escuelas Secundarias · Fomentar Nuevas tecnologías

## 1 Fundamentos de Nuestro Proyecto

La iniciativa de extensión universitaria propone acompañar las acciones desarrolladas en escenarios de tecnológicos que estén enfocadas en contribuir a fortalecer de la enseñanza de la computación en las escuelas primarias y secundarias, a mejorar la calidad de los cuidados y de vida de los pacientes en el sistema público de salud, a hacer más accesible el acceso al patrimonio cultural disponible en los museos y a promover una vinculación activa entre personas graduadas con la Facultad.

En este marco se procura ampliar y profundizar las vinculaciones establecidas, entre un conjunto de escuelas, hospitales y museos con la Facultad, para juntos construir dispositivos novedosos que ayudan a construir intervenciones tecnológicas que resultes superadoras y persistentes.

En este sentido se proyecta concretar acciones de extensión en el ámbito de la educación pública. La acción se despliega sobre escuelas primarias y secundarias de la región con tres tipos de actividades 1. Producción de recursos para la enseñanza de las Ciencias de la Computación, 2. Talleres en las Escuela y 3. Jornadas con estudiantes de escuelas secundarias y primarias.

En el marco de esta iniciativa se promueve la interacción fluida entre actividades de investigación, docencia y extensión, como estrategia para construir intervenciones más robustas, consistentes e integradas a la vida de la Facultad. Se impulsa la participación activa de la población destinataria en el diseño y

2 Authors Suppressed Due to Excessive Length

ejecución conjunta de las intervenciones que se realicen en en territorio como forma de avanzar en la constitución de comunidades con mayor autonomía, en este contexto se presta especial atención a poblaciones vulnerables con intención de ayudar a reducir brechas digitales de forma persistente.

## 2 Objetivos y Metodología

Contribuir de manera persistente a reducción de brechas digitales y mejoramiento de la situación actual a partir de intervenciones mediadas tecnológicamente en los dominios de la Educación Pública. Se busca mejorar la articulación de las actividades de Investigación, Docencia y Extensión para producir intervenciones más robustas y consistentes.

### 2.1 Ejes Metodológicos

Actualmente, la Facultad de Informática está desarrollando el proyecto “Reduciendo brechas” en donde se realizaron x talleres en el transcurso de x meses. Los talleres se desarrollan intentando compartir la experiencia universitaria y con el fin de acercar a lxs estudiantes a la misma, a través de jornadas breves en escuelas de nivel secundario, las cuales son llevadas adelante por estudiantes avanzados en las distintas carreras.

En los diversos talleres se requiere la participación de lxs estudiantes, quienes trabajan en conjunto en las actividades que luego son relacionadas con el concepto teórico correspondiente, siendo algunos de ellos software libre, ciclo de vida, claves simétricas, claves asimétricas y grooming. Esto ha permitido llevar varios tópicos distintos, relacionados con el ámbito de estudio, llevando a cabo actividades en donde se fomenta la actividad en grupo.

Todo el grupo participante trabaja de forma colaborativa para dar un pantallazo, tanto de lo que es la vida universitaria, como de distintos temas de la informática. Es importante destacar que el objetivo principal de este proyecto es despertar vocaciones en lxs estudiantes, por lo que se les habla sobre la experiencia de estudiar computación de manera profesional y se busca motivarlos y despertar su curiosidad. Esto hace que las jornadas sean más amenas y se genere un espacio en donde se puedan despejar las dudas acerca de las profesiones informáticas.

- Taller de Programación de drones. la programación de drones educativos es una forma de aprendizaje que permite la aproximación a los conceptos fundamentales de la programación, de la ingeniería de software y del software libre.

El objetivo de la actividad es completar el circuito con la menor cantidad de intentos posibles, promoviendo el trabajo colaborativo. El grupo se organiza en subgrupos de cuatro integrantes y se les asignan estaciones de trabajo separadas por unos metros. Cada subgrupo debe construir un programa para lleva el dron hasta la estación siguiente y luego pasar el control al próximo grupo.

- Internet Seguro: Estudiantes de la escuela secundaria trabajan en grupos en la elaboración de piezas destinadas a divulgar buenas prácticas acerca del uso responsable de internet. Los receptores de la divulgación son principalmente estudiantes y docentes de escuelas primarias.  
Las piezas construidas de suman Repositorio Digital sobre Internet Seguro de la Facultad de Informática de la Universidad Nacional del Comahue, EESISeg | Repositorio
- Construyendo Aplicaciones móviles en el aula. Los estudiantes se aproximan a los conceptos fundamentales de la programación a través de la construcción de aplicaciones móviles. Utilizando el entorno de desarrollo MIT appInventor, una herramienta simple y amigable que no requiere conocimientos previos sobre programación, los estudiantes producen aplicaciones para dispositivos Android.
- una Máquina de Turing en la Escuela. Se busca ofrecer un primer contacto con de prácticas y conceptos relacionados a las Máquinas de Turing y a nociones introductorias de computabilidad. Se trata de una experiencia lúdica basada en el trabajo en equipos conformados por 3 a 4 integrantes y la idea fundamental es presentarla como un juego con determinadas reglas. El objetivo principal de la actividad es entender el funcionamiento de una máquina que no es creada con tecnología pero que ejecuta instrucciones de la misma manera que una computadora actual.
- Taller de Seguridad Informática. El objetivo general de este taller es presentar ideas básicas sobre la Seguridad Informática a través de una experiencia conocida tomando como referencia el envío de mensajes encriptados a través de Internet.
- Taller de Ciencia de datos: Se busca promover una experiencia similar al análisis y procesamiento de datos; señalar las responsabilidades del equipo de desarrollo de los sistemas de Inteligencia Artificial cuando estos son utilizados para resolver cuestiones humanas y destacar la necesidad de una mirada crítica

### 3 Experiencias y Resultados

- Programate Itinerante. Esta actividad se realizó en colaboración con la Municipalidad de Mariano Moreno y el CPEM 37. Se trabajó con estudiantes de los dos últimos años de estudio y se realizaron talleres sobre internet seguro, programación de drones y desarrollo de ampliaciones móviles.  
Cantidad de Talleres: 6  
Cantidad de Estudiantes: más de 90 de los turnos mañana, tarde y vespertino del CPEM 37.
- Programate - Talleres. En el marco de Programate 2023, se desplegaron talleres máquinas de Turing, seguridad informática, ciencia de datos, programación de drones y desarrollo de ampliaciones móviles.  
Cantidad de Talleres: 12

## 4 Authors Suppressed Due to Excessive Length

Cantidad de Estudiantes: más de 180 de las escuelas ESRN 14 (Fernandez Oro - Río Negro), EPET 11 (Zapala - Neuquén), CPEM 37 (Mariano Moreno - Neuquén), CPEM 73 (El Cholar - Neuquén), CPEM 71 (Centenario - Neuquén) y AMEN (Neuquén - Neuquén).

- Visita a Escuelas. Durante 2023 se realizaron visitas a escuelas secundarias de la región con intención ofrecer contactos tempranos con las Ciencias de la Computación y difundir la oferta académica de la Facultad de Informática. Se desplegaron talleres máquinas de Turing, internet seguro, programación de drones y desarrollo de ampliaciones móviles.

Cantidad de Talleres: 20

Cantidad de Estudiantes: más de 450 de las escuelas CPEM 54 (Neuquén - Neuquén), CPEM 48 (Neuquén - Neuquén), CPEM 49 (Neuquén - Neuquén), CPEM 71 (Centenario - Neuquén), ESRN 17 (Cipolletti - Río Negro) y ESRN 137 (Sargento Vidal - Río Negro).

# Observatorio Electoral UNCo

## Participación ciudadana en defensa de la democracia

Pablo Kogan<sup>1</sup>, Daniela Zacharias<sup>2</sup>, Christian Gimenez<sup>1</sup> y Jorge Rodríguez<sup>1</sup>

<sup>1</sup> Facultad de Informática

<sup>2</sup> Centro Regional Bariloche

Universidad Nacional del Comahue

<https://observatorioelectoral.uncoma.edu.ar>

[observatorioelectoral@fi.uncoma.edu.ar](mailto:observatorioelectoral@fi.uncoma.edu.ar)

**Resumen** Con el objetivo de cuidar la democracia y mejorar los procesos electorales el Observatorio Electoral de la UNCo realiza el monitoreo y seguimiento de los comicios que se realizan en las Provincias de Neuquén y Río Negro. En el presente documento se describen los fundamentos de creación del Observatorio, los objetivos y las principales actividades realizadas.

**Keywords:** Observación Electoral · Participación Ciudadana · Sistemas de votación

### 1. Fundamentos

Los procesos electorales constituyen la expresión institucional de la voluntad popular en los sistemas democráticos. Nuestra sociedad ha ido ensayando diferentes sistemas para realizar estos procesos, tendiendo a incrementar su transparencia y equidad como forma de legitimar el poder de las autoridades electas. Los mecanismos de observación electoral tienden a aumentar dicha legitimidad y a garantizar el respeto por el derecho al voto de toda ciudadanía.

En su acepción más simple, la observación electoral consiste en realizar una serie de actividades coordinadas y metodológicamente coherentes durante las jornadas electorales para evaluar la calidad de los comicios de manera independiente a la autoridad electoral a cargo. Numerosas experiencias en otros países y gobiernos locales demuestran la efectividad de esta práctica para identificar y prevenir errores, distorsiones y manipulaciones en los procedimientos de logística electoral y de la votación, pudiendo evaluar la calidad global de los procesos del día de elecciones y proyectar, o verificar, los resultados oficiales. Para que los resultados de la observación se consideren legítimos, la actividad debe ser realizada por organismos u organizaciones aceptadas como imparciales por los diferentes actores del proceso electoral.

En la Argentina las universidades públicas son verdaderos pilares de la vida democrática del país por sus características inclusivas, por ser espacios donde se

2 Pablo Kogan, Daniela Zacharias, Christian Gimenez y Jorge Rodríguez

ejercita el diálogo en base al conocimiento científico y por el rol que han cumplido en la promoción de valores fundamentales como la defensa de los derechos humanos y la formación integral de las personas que participan de la comunidad universitaria.

El Observatorio Electoral, enmarcado como actividad de Extensión, contribuye en la promoción de estos valores, permitiendo e incentivando la participación de la ciudadanía en el cuidado de la democracia.

## 2. Objetivos

El Observatorio Electoral de la Universidad Nacional del Comahue (OEUNCo) es el ámbito institucional que es responsable del contacto permanente y seguimiento de los procesos electorales de las Provincias de Río Negro y Neuquén.

### Objetivos Específicos

- Promover la participación y formación de la ciudadanía y de la comunidad Universitaria en particular, en la mecánica de los procesos electorales.
- Ser un referente de consulta sobre la observación y evaluación de la calidad de los procesos electorales, antes, durante y después de los comicios, generando recomendaciones para la mejora de los mismos.
- Conformar un espacio de intercambio y diálogo entre la academia, el estado y la opinión pública.
- Generar, compilar, sistematizar y analizar información sobre distintos aspectos de los procesos electorales.

### 2.1. Historia y antecedentes

A inicios del año 2015 [6], el Centro Regional Universitario Bariloche de la Universidad Nacional del Comahue y la Defensoría del Pueblo de Bariloche, crean el primer observatorio electoral de la provincia de Río Negro: el Proyecto ObserBar. Al año siguiente, fue seleccionado en la convocatoria Compromiso Social Universitario, convirtiéndose en el Observatorio Electoral de Río Negro y extendiendo su cobertura a Bariloche y General Roca.

En 2021 se conforma el Observatorio de la Universidad Nacional del Comahue según Ordenanza CS 919/2021 [5] ampliando el ámbito de influencia a las provincias de Neuquén y Río Negro.

El OEUNCo forma parte del registro de entidades de Acompañamiento Cívico de la Cámara Nacional Electoral (CNE) [1]

## 3. Impacto

La observación electoral es crucial para la salud de las democracias, ya que contribuye a la transparencia, la confiabilidad de los procesos electorales y la promoción de la participación ciudadana, elementos esenciales para la estabilidad y el fortalecimiento de las instituciones democráticas. El impacto del OEUNCo abarca varias dimensiones:

### 3.1. Observación de procesos electorales

Con el fin de obtener indicadores acerca del desarrollo del proceso electoral mediante la observación en una muestra de mesas de votación, en las cuales se recoge información cualitativa y cuantitativa la red de observadores/as ha estado presente en distintos procesos electorales:

- Elecciones municipales: San Carlos de Bariloche (2015 y 2019), General Roca / Fiske (2023), Cinco Saltos (2023), Dina Huapi (2023), Neuquén elección Concejales y Referéndum Popular (2021) Cutral-Có y Plaza Huinul (2023)
- Elecciones provinciales de Río Negro (2015, 2019 y 2023),
- Elecciones provinciales de Neuquén (2023).
- Elecciones nacionales generales (Primera y segunda vuelta del 2015, 2019)
- Elecciones legislativas nacionales de Río Negro (2017 y 2021), de Neuquén (2021);
- Elecciones Universidad Nacional del Comahue (2022)
- Elecciones gremio ADUNC (2022); gremio APUNC (2022)

Asimismo, realizamos estudios [2] [3], electorales y capacitaciones para la comunidad. Todos estos estudios se encuentran disponibles en el sitio Web:

<https://observatorioelectoral.uncoma.edu.ar>

### 3.2. Red de Observadores/as

Desde los inicios y hasta la fecha, han participado de manera voluntaria más de 300 personas en distintas localidades de Neuquén (Neuquén capital, Plottier, Centenario, Zapala, Cutral-Co, Plaza-Huinul, San Martín, Senillosa, Junín de los Andes, Los Guañacos, Villa la Angostura) y de Río Negro (Bariloche, Viedma, Cipolletti, Allen, Fernández Oro, Roca, Cinco Saltos, Godoy, Chimpay, Chichinales, Cervantes, Choele Choel, Belisle, Sierra Grande, Dina Huapi, Valcheta, Mainque, Villa Regina, Beltrán, El Bolsón y Clemente Onelli). Estas personas al involucrarse en la observación del cumplimiento de los procedimientos electorales pudieron conocer y valorar el derecho de sufragio. Muchas siguieron participando luego como autoridades de mesa o fiscales.

### 3.3. Actividades de formación ciudadana en escuelas secundarias

Se trata de una experiencia educativa que busca que los/as estudiantes vivencien tempranamente el ejercicio del sufragio construyendo así los conocimientos para la participación en futuros procesos electorales[4].

Además, en esta experiencia, se les propone explorar diferentes sistemas de votación, generando criterios para poder participar de manera informada en discusiones ciudadanas acerca de las formas de votar.

Esta actividad se enfoca en capacitar a los/as estudiantes en todo el proceso electoral, respondiendo las preguntas: ¿qué categorías se eligen?, ¿cuándo se vota?, ¿qué hace falta llevar?, entre otros. En este sentido se les propone a

4 Pablo Kogan, Daniela Zacharias, Christian Gimenez y Jorge Rodríguez

las/os estudiantes el “Desafío Votar con tres sistemas el mismo día”, simulando todo el proceso de votación, con la oferta electoral real, urnas, sobres, boletas, cuartos oscuros, padrones.

Durante el año 2023 participaron de la actividad más de 300 estudiantes de las escuelas: CPEM 71 (Centenario), CPEM 43 (Cutral Co), ESRN 23 (Cinco Saltos), ESRN 45 (Bariloche) y ESRN 96 (Dina Huapi). Se espera que las visitas a las escuelas secundarias promuevan la participación democrática y contribuyan a la formación ciudadana en aspectos relacionados a los procesos electorales.

### 3.4. Contribuciones a procesos electorales

Algunos de los aportes sugeridos por el observatorio fueron tomados por los órganos electorales. Algunos ejemplos:

- Elecciones municipales de San Carlos de Bariloche. Diseño de Boleta Única de Papel. Incorporación de afiches con tabulación para mejorar conteo de votos en elecciones municipales de San Carlos de Bariloche.
- Elecciones Universidad Nacional del Comahue. Implementación de Boleta Única de Papel.
- Elecciones provinciales de Neuquén. Incorporación de Protocolo de Voto Aleatorio en sistema Boleta Única Electrónica. Actualización de librerías y buenas prácticas de codificación en sistema de Boleta Única Electrónica.

## Referencias

1. Cámara Nacional Electoral (CNE) Argentina: *Registro de Entidades de Acompañamiento Cívico*. <https://old.pjn.gov.ar/cne/acompacivico/publico.php>,
2. Kogan, Pablo, Silvia Soto, Claudio Vaucheret y Daniela Zacharias: *¿ Cómo aumentar la velocidad de publicación de resultados de una elección?* Memorias de las JAIIO, 8(12):69–79, 2022.
3. Kogan, Pablo, Silvia Soto y Claudio A Vaucheret: *Aportes al proceso de fiscalización electrónica: experiencia sobre uso de la boleta única electrónica en elecciones 2019, provincia de neuquén*. En *XIII Simposio de Informática en el Estado (SIE 2019)-JAIIO 48 (Salta)*, 2019.
4. Kogan, Pablo, Daniela Zacharias, Christian Gimenez y Jorge Rodríguez: *Desafío: Votar con tres sistemas el mismo día*. En *Simposio de Informática en el Estado (SIE 2023)-JAIIO 52 (Buenos Aires)*, 2023.
5. Universidad Nacional del Comahue, Consejo Superior: *Ordenanza CS N° 1386/13 Universidad Nacional del Comahue*. [https://despacho.uncoma.edu.ar/archivos/ord\\_1386\\_2013\\_46.pdf](https://despacho.uncoma.edu.ar/archivos/ord_1386_2013_46.pdf),
6. Zacharías, Daniela y Gilda Garibotti: *En pos de la mejora de la calidad electoral: la experiencia de la creación del Observatorio Electoral de Bariloche*. *Revista Política, Globalidad y Ciudadanía*, 2(3):36–45, 2016.

ACTAS DE LAS

JORNADAS DE LA FACULTAD DE INFORMÁTICA



NOV 2023  
NEUQUÉN

FACULTAD DE INFORMÁTICA  
UNIVERSIDAD NACIONAL DEL COMAHUE

*Trabajos de estudiantes de grado y posgrado*

# Observaciones al rol de la infraestructura tecnológica escolar en la educación en Ciencias de la Computación

Jonathan Maximiliano Cabrera

Facultad de Informática, Universidad Nacional del Comahue  
Neuquén, Argentina  
jonathan.cabrera@est.fi.uncoma.edu.ar

**Abstract.** La finalización de la emergencia sanitaria por covid-19 nos ha mostrado la necesidad de estar preparados ante posibles situaciones que impidan la asistencia presencial a las escuelas. En este contexto, la dimensión tecnológica se ha vuelto más evidente y relevante para todos. Por que las tecnologías forman parte de la vida cotidiana de la mayoría de la población y las nuevas generaciones las más involucradas en los cambios, se consideró a la educación pública en observación para la relevación de datos e interpretarlos con parámetros internacionales que estiman una educación de calidad. En efecto se evidencia un ostentable margen de inferioridad en proporción a los promedios de recomendación. Por consiguiente, es pertinente un refortalecimiento continuo y creciente del presupuesto al área Informática de la educación, favoreciendo a la comunidad educativa en general.

**Keywords:** relación estudiante-computadora, TIC, laboratorios informáticos, conectividad en escuelas, políticas educativas, infraestructura tecnológica escolar.

## 1 Introducción

En este artículo se estudia el rol de la infraestructura tecnológica en la educación, con especial énfasis en la enseñanza y el aprendizaje de las Ciencias de la Computación (como la programación, el diseño, la robótica, etc.). El objetivo es analizar el número de computadoras, laboratorios y la calidad de la conexión a internet en las escuelas de la ciudad de Neuquén y compararlos con los estándares internacionales. Así se pretende determinar si los establecimientos educativos evaluados cumplen con el mínimo recomendado para una educación de calidad en esta área.

## 2 Metodología: recopilación de la información

Los dispositivos informáticos tienen una gran importancia en los ámbitos más abarcadores de las ciencias educacionales. Por ello, destaco su importancia como

objeto de estudio (una ciencia) y como herramienta de utilidad para las demás áreas del aprendizaje. Asimismo, una computadora es un recurso importante para el estudio y para otras áreas en las escuelas. De esta manera se puede considerar como un medio didáctico o un recurso educativo, según la definición de [1] [3]. Por otro lado, para Pere Marqués [2](p.1) el primero es un material elaborado con la intención de facilitar los procesos de enseñanza y aprendizaje, mientras que el segundo es un material que se utiliza con finalidades didácticas o para facilitar el desarrollo de las actividades formativas en un contexto educativo determinado. De este modo, la utilización de una computadora permite optimizar y potenciar los mecanismos de enseñanza. Así, el docente puede propiciar distintas formas de acceder al conocimiento.

Partiendo desde la fundamentación del párrafo anterior, la investigación se basó en datos obtenidos por diferentes escuelas públicas de la ciudad de Neuquén<sup>1</sup>. Se recopiló información sobre la cantidad de computadoras por escuela, salas de computación y la conectividad a internet que tenía cada una, es decir, los recursos tecnológicos disponibles, en proporción a la matrícula de cada establecimiento.

El objetivo fue analizar el acceso y el uso de las tecnologías digitales en el ámbito educativo, por ejemplo, cómo se integran las TIC en las prácticas pedagógicas, a partir de la información de las escuelas en conjunto. Para realizar el relevamiento, en cuanto al equipamiento de computadoras se analizaron los datos distinguiendo los turnos mañana y tarde. De este modo, el sondeo arrojó que la matrícula del turno mañana tiene un promedio de 185,60, mientras que la del turno tarde es 179,00. La docencia matutina tiene una media de 20,20 y los docentes vespertinos un 19,80. Del mismo modo, la relación entre el número de máquinas y el número de estudiantes en el turno mañana promedia 0,16; en tanto en el turno tarde tiene una media de 1,18. De esta manera la sección turno mañana promedia un 10,25 y la sección turno tarde un 10,00 de término medio. En cuanto a servidores el promedio es 1,50 y en cantidad de máquinas 31,50 a correspondencia, del total censado. En esa misma línea la cantidad de laboratorios por sección turno mañana tiene una media de 0,22, así como el turno tarde tiene un 0,22 de promedio. Por otro lado, en relación a la conectividad, el sondeo se dividió en parte administrativa y parte pedagógica en relación a las velocidades de descarga, carga y latencia de la conexión a Internet. En los lugares donde se realizaron los tests, se observó que la velocidad promedio de descarga de la conexión a internet es de 7 megabits<sup>2</sup> por segundo, y que en la sección administrativa es 8,283461538; mientras que en la sección pedagógica 3,411153846. Sin embargo, la velocidad promedio de subida de la conexión a internet es de 8 megabits por segundo al momento de testear, dando un 17,68230769 la de administración y 3,192692308 la pedagógica. En cuanto al tiempo promedio que

---

<sup>1</sup> Relevamiento realizado en el año 2022 por Jorge Rodríguez (j.rodrig@fi.uncoma.edu.ar) y Pablo Kogan (pablo.kogan@fi.uncoma.edu.ar), docentes de la Facultad de Informática de la Universidad Nacional del Comahue

<sup>2</sup> unidad de medida de información que se utiliza para medir la velocidad de transmisión de datos en una red. Un megabit equivale a 1.000.000 de bits.

tarda la conexión a internet en comunicarse con otro equipo en la red, es de 6 milisegundos<sup>3</sup>, se encontró que la parte de administración es de 17,68230769 en comparación al 28,03846154 de la parte educativa. En general, la descarga promedió un 5,847307692 Mbps mientras que la velocidad de subida, que permite enviar información a internet, promedió 7,655769231 Mbps. Respecto a la medida de latencia, el promedio arrojó 22.86038462 ms<sup>4</sup> de retraso producido entre el envío y la recepción de información. Antes de continuar, es propicio considerar para el siguiente relevamiento, los dos modos<sup>5</sup> más usados para la prestación presencial de servicios de computación a los miembros de una comunidad educativa pública. Estos son: (1) el traslado de alumnos a las salas donde se encuentran los computadores, y (2) la llegada del carrito informático de computadoras al aula o sala. Esto se debe a que estas modalidades permiten optimizar el uso de los recursos tecnológicos disponibles y facilitar el acceso y la integración de las TIC en el proceso educativo. Para evaluar la velocidad de descarga, carga y ping<sup>6</sup> de la conexión a internet en la sección de laboratorios, se realizó un test en relación con la cantidad de equipos conectados a la red local. Se subdividió el test en tres partes: (1) hasta 10 equipos, (2) hasta 20 equipos y (3) más de 20 equipos. Los resultados fueron los siguientes: con un promedio de 7 Mbps de descarga, una sala con hasta 10 equipos conectados arrojaba 4,45 Mbps, con hasta 20 dispositivos electrónicos conectados 2,78 Mbps y con más de 20, 1,66 Mbps; con un promedio de 8 Mbps de carga si había 10 equipos conectados se marcaba 5,89 Mbps, con 20 dispositivos vinculados 0,05 Mbps y con más de 20 el test media 0,54 Mbps. En cuanto al tiempo que demora el rastreador de paquetes en redes, que medía 6 ms en promedio, al tener 10, 20 y más de 20 equipos conectados se medían 20,99 ms, 29,55 ms y 41 ms, respectivamente.

### 3 Resultados: análisis con parámetros

Para analizar e interpretar los resultados obtenidos, se consideraron los siguientes factores: accesibilidad a computadoras que estén actualizadas y en buen estado, accesibilidad a softwares relevantes, las signaturas, la orientación educativa y el nivel de educación. En primer lugar, según estudios internacionales, se sugiere que la relación ideal entre el número de estudiantes y computadoras es de 1:1 (una a uno) en el nivel secundario y superior, y al menos 1:3 (una por cada tres) en el nivel primario. Sin embargo, los resultados del primer análisis no satisfacen lo esperado en ambos turnos. Por otro lado, se observa que la disposición de cada sección al acceso diario de un computador es de 40 minutos diarios para la enseñanza de signaturas destinadas a la ciencia de la computación y para el

<sup>3</sup> unidad de medida de tiempo que equivale a la milésima parte de un segundo.

<sup>4</sup> Abreviatura de milisegundo: Unidad de medición del retardo de datos de un computador a otro.

<sup>5</sup> Rodriguez, J. (comunicación personal, 10 de agosto de 2023).

<sup>6</sup> Abreviatura de “packet internet groper”, herramienta de diagnóstico que permite verificar el estado de una conexión de un host local con al menos un equipo remoto en una red TCP/IP.

aprendizaje de otras áreas. Además, no se observó una diferencia marcada en los promedios comparando el turno tarde y el turno mañana; aunque se favorece ligeramente al último. En el segundo análisis, relacionado con la conectividad en las escuelas, se evidencia que los espacios áulicos tienen una conexión más deficiente en comparación con los espacios destinados a la administración de las escuelas. En consecuencia, se corroboran los estándares internacionales de anteaño que aseveran que la mayoría de los países de América Latina no cumplen con las condiciones básicas de conectividad en las escuelas. Obviamente los laboratorios o salas de computación se utilizan como soporte para las áreas académicas de formación y aprendizaje de programas de informática. También son ambientes inclusivos y estructurados para el desarrollo de habilidades tecnológicas. El último análisis interpreta que, en un aula con más de 20 equipos, el promedio de descarga es 1.655 con un mínimo de 1.38, cuando la recomendación internacional para 2021 era de 2 a 4.

## 4 Conclusiones

En general, se recomienda que haya suficientes computadoras para que cada estudiante tenga acceso a una durante las actividades en clase. Además, la utilización de dispositivos informáticos permite optimizar y potenciar los mecanismos de enseñanza y su aprovechamiento propicia en la actualidad los ámbitos más abarcativos de las ciencias educacionales. Asimismo, la utilización y reutilización de softwares representa un abanico de oportunidades, favoreciendo propuestas más eficientes que ahorren costo, tiempo y esfuerzos cuando el docente sabe aprovechar su uso. Hasta la fecha de publicación no se hallaron variantes notorias en cuanto a conexión a redes, como tampoco hubo nuevas incorporaciones de insumos informáticos en las escuelas censadas<sup>7</sup>. Cabe destacar que las desigualdades económicas tienen su correlato en cuanto a conexión y dispositivos electrónicos en los hogares. Si bien existen políticas no educativas que tienden a reducir la brecha digital, no bastan y tampoco satisfacen. Por ello, el haber y disponer en las escuelas, es fundamental para reducir la brecha de desigualdad que existe en cuanto al acceso a computadores y redes de los estudiantes y docentes, pudiendo ser el único momento en el día a su disposición. A modo de conclusión, es oportuno un fortalecimiento constante y observable del presupuesto en el área computacional de las escuelas, a quien les compete. Con esto se demostraría una concordancia acorde a la prioridad que es, debe seguir siendo y debería de aumentar este sector fundamental para la educación. Además, es pertinente visibilizar el tema en agenda, ya que la problemática de la infraestructura tiene realmente sus consecuencias en las posibilidades de enseñar y que quienes estudian aprendan sobre Ciencias de la Computación. De esta manera, se podrán alcanzar los estándares de recomendación internacional para que el área educacional de la informática de Argentina esté a nivel de competencia de mercado.

---

<sup>7</sup> Relevamiento año 2023: Datos obtenidos por el autor.

## **Agradecimientos**

Agradecimiento a los docentes de la Universidad Nacional de Comahue que aportaron con sus textos e información.

## **Referencias**

1. Litwin, E.: De caminos, puentes y atajos: el lugar de la tecnología en la enseñanza. In: Conferencia inaugural II Congreso Iberoamericano de Educación y Nuevas Tecnologías. Buenos Aires. vol. 30 (2005)
2. Pere Marquès, G.: Los medios didácticos y los recursos educativos. Disponible en: [peremarques.pangea.org/medios.htm](http://peremarques.pangea.org/medios.htm). [consultado 26 abril 2014] (2000)
3. Zabala, A.: La práctica educativa: cómo enseñar. Graó (2000)

# Herramienta de diseño de aplicaciones Web para la consulta de programas Prolog

Santiago A. Villarroel, Christian N. Gimenez,  
Jorge P. Rodríguez, and Laura A. Cecchi

Grupo de Investigación en Lenguaje e Inteligencia Artificial  
Facultad de Informática  
Universidad Nacional del Comahue  
{santiago.villarroel, christian.gimenez, j.rodri,  
lcecchi}@fi.uncoma.edu.ar

**Palabras Clave:** Enseñanza de la Programación Lógica . Prolog . Aplicaciones Web

## 1 Objetivos

Los objetivos particulares que nos hemos planteamos en esta línea de investigación, que está enmarcada en las actividades del GILIA son dos. En primer lugar, desde el punto de vista disciplinar, buscamos *desarrollar una herramienta Web que permita diseñar y generar automáticamente una aplicación Web para consultar un programa lógico en Prolog.*

En segundo lugar, nos proponemos *fortalecer los vínculos con otros grupos de investigación internacionales cuyos investigadores participan del Prolog Education Committee.* Ya hemos establecido un marco de colaboración con estos grupos, y nuestro objetivo es profundizar este intercambio de conocimientos y experiencias en la enseñanza de Prolog en todos los niveles educativos.

## 2 Línea de Trabajo

En la actualidad, existe un consenso internacional sobre la importancia del Pensamiento Computacional (PC) [8] y del Pensamiento Lógico (PL) en el desarrollo de habilidades para encontrar soluciones a problemas de propósito general. Así, muchos países han comenzado a incorporar el PC en los planes de estudio desde la escuela primaria [2, 7].

Existen diversos recursos didácticos que colaboran en la introducción del PC a los niños a través de la Programación Imperativa. La mayoría de estos proyectos emplean lenguajes de programación visual, mediante los cuales los estudiantes pueden crear programas utilizando elementos gráficos e interfaces de arrastrar y soltar. Ejemplos de estas herramientas son Scratch <sup>1</sup>, Alice<sup>2</sup> y Open Roberta<sup>3</sup>.

<sup>1</sup> <https://scratch.mit.edu/>

<sup>2</sup> <http://www.alice.org/>

<sup>3</sup> <https://lab.open-roberta.org/>

Asimismo, existe una serie de recursos que permite a los estudiantes diseñar sus propias aplicaciones o páginas Web programando con Programación Imperativa como MIT App Inventor <sup>4</sup>.

El paradigma de la Programación Lógica (LP) no tiene la suficiente representación en los cursos de introducción a las Ciencias de la Computación (CC) y el desarrollo de las habilidades de PC y PL de los niños. Sin embargo, dicho paradigma es una opción viable y deseable para niños de escuela primaria, incluso para aquellos sin conocimientos previos de programación. Se han realizado investigaciones para introducir las CC y desarrollar habilidades de PC a través de la LP para estudiantes de nivel secundario [1, 9, 10]. Del mismo modo, creemos que es posible lograr la introducción fructífera de esta ciencia y el desarrollo de dichas habilidades mencionadas en niños escolarizados de nivel primario.

Para que la enseñanza del LP sea viable, es necesario contar con los recursos didácticos adecuados. Por ejemplo, una implementación visual de Prolog y herramientas Web. El uso de lenguajes basados en bloques [4, 6] evita la necesidad de enseñar y manejar la sintaxis textual de Prolog, mientras que un entorno Web permite el trabajo colaborativo y elimina los inconvenientes asociados con la instalación de software.

En este sentido, es crucial realizar esfuerzos continuos para desarrollar nuevos entornos tecnológicos que permitan consolidar la LP en la escuela primaria, aumentando y fortaleciendo las iniciativas ya existentes. Actualmente, se están desarrollando recursos educativos con el fin de involucrar a los niños, elaborar lecciones personalizadas, y fomentar la participación en el aula. Por ejemplo, en [5] los autores introducen un entorno Web interactivo basado en Ciao Playground<sup>5</sup> para guiar a los estudiantes con ejercicios específicos durante el proceso de aprendizaje de la programación mediante Prolog.

En [3] se presenta una serie de actividades para enseñar la LP en las escuelas primarias, las cuales consisten en juegos divididos en etapas. En las mismas, se plantea un problema y los estudiantes deben analizar pistas para llegar a una solución, familiarizándose con conceptos de este paradigma, tales como constantes, variables, y relaciones.

En estas actividades se identifica la necesidad de un recurso que permita que los niños interactúen con programas lógicos desarrollados en las experiencias educativas, y pongan en práctica conceptos de la PL aprendidos durante las mismas. Esta interacción, idealmente, debe ser amena y permitir cierto grado de creatividad por parte de los estudiantes.

En este contexto, la solución propuesta consiste en el diseño y la implementación de un recurso didáctico visual y Web que permita la creación de una aplicación Web personalizada por el usuario, para consultar un programa Prolog ya construido.

En el diseño e implementación de este recurso, se priorizará que la interfaz de usuario sea altamente intuitiva, de manera que su uso no requiera conocimien-

---

<sup>4</sup> <https://appinventor.mit.edu/>

<sup>5</sup> <https://ciao-lang.org/playground/> visitado el día 10 de octubre del 2023

tos previos. Además, se trabajará sobre una implementación de Prolog Web, evitando las posibles complicaciones a la hora de instalar software.

Se espera que con esta herramienta se alcance un rango mayor de estudiantes en el ámbito de la enseñanza de la Programación Lógica, brindando un entorno de diseño intuitivo que permite a cualquier usuario, especialmente niños, construir aplicaciones Web. Se busca democratizar el desarrollo de software empoderando a todas las personas, particularmente a niños y jóvenes en edad escolar, y apoyando la transición desde el rol de consumidores de tecnología hacia el de creadores de la misma.

### 3 Contexto

Este trabajo se enmarca en los proyectos desarrollados en el contexto del Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA):

- Proyecto de Investigación “Tecnologías Semánticas para el desarrollo de Agentes Inteligentes” (04/F020). Departamento de Teoría de la Computación. Facultad de Informática. Universidad Nacional del Comahue;
- Proyecto de Investigación “Modelos Formales, Agentes Inteligentes y Aplicaciones para la Enseñanza de las Ciencias de la Computación” (04/F022). Departamento de Teoría de la Computación. Facultad de Informática. Universidad Nacional del Comahue;
- Proyecto de Extensión “Programación lógica en la escuela primaria” (Ordenanza 0369/2023), Departamento de Teoría de la Computación. Facultad de Informática. Universidad Nacional del Comahue.

Esta línea de investigación se desarrolla en forma colaborativa con investigadores del IMDEA Software Institute<sup>6</sup>, Madrid, España. Asimismo, es parte de las iniciativas internacionales que se llevan adelante en el marco de las actividades del Prolog Education Committee<sup>7</sup>, del que autores de este trabajo forman parte.

### References

1. Beux, S., Briola, D., Corradi, A., Delzanno, G., Ferrando, A., Frassetto, F., Guerini, G., Mascardi, V., Oreggia, M., Pozzi, F., et al.: Computational thinking for beginners: A successful experience using Prolog. In: CILC. pp. 31–45 (2015)
2. Bocconi, S., Chiocciariello, A., Kampylis, P., Dagienè, V., Wastiau, P., Engelhardt, K., Earp, J., Horvath, M.A., Jasutè, E., Malagoli, C., et al.: Reviewing Computational Thinking in Compulsory Education. Tech. rep., Joint Research Centre (Seville site) (2022)
3. Cecchi, L.A., Rodríguez, J.P., Dahl, V.: Logic programming at elementary school: Why, what and how should we teach logic programming to children? In: Warren, D.S., Dahl, V., Eiter, T., Hermenegildo, M., Kowalski, R., Rossi, F. (eds.) Prolog - The Next 50 Years. No. 13900 in LNCS, Springer (July 2023)

<sup>6</sup> <https://software.imdea.org/>

<sup>7</sup> <https://prologyear.logicprogramming.org/Education.html>

4. Laboratorio Kagawa: Homepage Blockly for Prolog , <https://guppy.eng.kagawa-u.ac.jp/kagawa/Members/Sano/prolog.html>
5. Morales, J.F., Abreu, S., Hermenegildo, M.V.: Teaching Prolog with Active Logic Documents . In: Warren, D.S., Dahl, V., Eiter, T., Hermenegildo, M., Kowalski, R., Rossi, F. (eds.) Prolog - The Next 50 Years. No. 13900 in LNCS, Springer (July 2023)
6. Niklas Holtz: Homepage blockly prolog - universität oldenburg, <http://www.programmierkurs-java.de/blocklyprolog/>
7. Ottestad, G., Gudmundsdottir, G.B.: Information and communication technology policy in primary and secondary education in Europe. Second handbook of information technology in primary and secondary education pp. 1–21 (2018)
8. Wing, J.M.: Computational thinking. *Communications of the ACM* **49**(3), 33–35 (2006)
9. Yuen, T.T., Reyes, M., Zhang, Y.: Introducing computer science to high school students through logic programming. *Theory and Practice of Logic Programming* **19**(2), 204–228 (2019)
10. Zhang, Y., Wang, J., Bolduc, F., Murray, W.G.: LP based integration of computing and science education in middle schools. In: *Proceedings of the ACM Conference on Global Computing Education*. pp. 44–50 (2019)

# Clasificación de usuarios que preguntan en CQA

Adriana Mabel Cárdena, Gabriela N. Aranda, Valeria Zoratto

Tesis de Licenciatura en Ciencias de la Computación  
Facultad de Informática, Universidad Nacional del Comahue  
Neuquén, Argentina

**Abstract.** Las Community Question Answering (CQA) se han convertido en una forma popular de búsqueda de información en línea, donde los usuarios pueden interactuar e intercambiar conocimientos en forma de preguntas y respuestas. Particularmente, los foros de discusión cuentan con información que suele ser considerada muy valiosa por usuarios que acceden en busca de soluciones a problemas recurrentes. Por este motivo, en los últimos años se ha generado un especial interés en desarrollar técnicas de recuperación que permitan realizar el análisis de la información que se extrae de los hilos de discusión, para poder reutilizarla rápidamente y por medio de tecnología que dé soporte a búsquedas automáticas o semi-automáticas. Pero además del interés en las soluciones a problemas reales, también ha crecido el interés en investigar sobre las redes sociales que conforman los usuarios intervinientes en las CQA, y los roles que en éstas ocupan. A partir de revisiones bibliográficas preliminares, detectamos que no estaba muy desarrollada la investigación sobre el comportamiento de los usuarios que realizan preguntas en dichos tipos de comunidades. Por ese motivo, en esta tesis realizamos un estudio del comportamiento de los usuarios desde la cantidad y calidad de su participación en foros de discusión, y proponemos un modelo de clasificación de perfiles enfocado en dichas características. Dicho modelo ha sido aplicado en dos conjuntos de datos de usuarios obtenidos de foros de discusión reales con resultados prometedores.

**Keywords:** Foros de discusión · Recuperación de información · Roles de usuarios

## 1 Introducción

Los foros de discusión técnicos se han transformado en bitácoras actualizadas de conocimiento sobre temas específicos, que están en la mira de varios proyectos enfocados en aplicar estrategias de Recuperación de Información (RI). Además, los usuarios que participan en dichos foros comparten conocimientos y experiencias que son de gran utilidad cuando un problema surge repetidamente.

En un trabajo anterior [3], avanzamos en la detección de los roles de usuarios participantes en foros de discusión técnicos. En dicho trabajo se definieron las características de los distintos roles, la forma de clasificarlos y se presentó una validación preliminar del modelo propuesto en un caso de estudio. A partir de

dicho trabajo y de la revisión de trabajos relacionados sobre roles de usuarios, pudimos observar que había pocos estudios enfocados en el comportamiento de los usuarios que realizan preguntas. Por dicho motivo, el trabajo de esta tesis se enfocó principalmente en analizar trabajos que abordaran temáticas relacionadas al comportamiento de los usuarios y de la calidad de sus contribuciones, con el fin de proponer un modelo de clasificación de los usuarios que preguntan en foros de discusión.

Esta tesis fue desarrollada por la estudiante Adriana Cárdena entre 2020 y 2023, dirigida por la Dra. Gabriela Aranda y la Lic. Valeria Zoratto de FAI-UNCOMA. La misma se ha enmarcado en los Proyectos de investigación “Reuso de conocimiento en Foros de Discusión Técnicos - Parte II”, subproyecto del Programa 04/F009 “Desarrollo de Software Basado en Reuso - Parte II” (Periodo 2017-2021) y en el Proyecto 04/F018 “Reuso de Información en Comunidades Virtuales” (Periodo 2022-2025), desarrollados por el Grupo de Investigación GI-ISCo de la Facultad de Informática, Universidad Nacional del Comahue.

## 2 Contenido de la tesis

En base a la revisión de varios trabajos relacionados [7][2][6][4] [5][1], cuyo reporte se presenta en el Capítulo 2, se desarrolló una estrategia para la recuperación y procesamiento de archivos provenientes de hilos de discusión reales de la Web. Dicha estrategia se enfocó en la recuperación de información asociada a usuarios y sus correspondientes contribuciones (posts) en foros de discusión, como así también en la definición de los algoritmos y fórmulas a aplicar adecuados para el posterior procesamiento de sets de datos reales, con la finalidad de definir características de calidad y comportamiento que permitan la clasificación de los usuarios que preguntan.

Para ello se establecieron tres preguntas de investigación, basadas en la participación de los usuarios y la percepción de la comunidad sobre la calidad de las contribuciones realizadas por ellos:

- ¿Cómo varía a lo largo del tiempo la cantidad de preguntas que hace un usuario?
- ¿Cómo varía a lo largo del tiempo la cantidad de respuestas aceptadas por el usuario que pregunta?
- ¿Cómo varía a lo largo del tiempo la percepción de la comunidad con respecto a las preguntas realizadas por un usuario?

Para responder dichas preguntas, se diseñó e implementó una componente software en lenguaje Java que permite el preprocesamiento y recuperación de información concerniente a distintos conjuntos de usuarios y de todos los posts escritos por ellos. Dicha herramienta realiza el filtrado inicial de datos, la agrupación y cálculo de características tales como cantidad y tipo de posts realizados en forma total y por periodos de 12 meses considerando la fecha de registración de un usuario en el sitio, así como contabilización de marcas de favoritos, puntaje balance entre votos positivos y negativos, detección de marca de respuesta

aceptada, entre otros datos considerados de utilidad. La salida de datos proporcionada por esta herramienta fue utilizada para proponer un modelo de clasificación de usuarios que preguntan, en un conjunto de cuatro perfiles, tal como fue desarrollado en el Capítulo 3.

Para evaluar el modelo de clasificación propuesto, se utilizó la base de datos del foro StackOverflow del 03/06/2019. Partiendo de dicha base, se obtuvieron dos sets de datos, el primero conformado por usuarios que fueron dados de alta en el mismo mes y año de creación del foro (Agosto 2008) y el segundo conformado por usuarios dados de alta tres años después (año 2011). En ambos casos los usuarios incluidos debían tener su último acceso al foro durante 2018, para poder analizar los cambios de comportamiento en el tiempo. También se analizaron los post creados por dichos usuarios durante su historia completa en el sitio.

En base al modelo de clasificación propuesto, se procedió a analizar a los usuarios de ambos casos de estudio de dos maneras diferentes:

- Mediante un perfil global, determinado por la actividad de cada usuario durante toda su vida activa en el foro. Para ello se consideró como variable el tiempo de permanencia en estado activo de cada usuario, ya sea realizando preguntas o respuestas, a fin de detectar si los perfiles de cada usuario podían influir de alguna manera en la permanencia o abandono de la comunidad.
- A continuación, se repitió el análisis asignando un perfil por ventanas de tiempo consecutivas de 12 meses, a fin de detectar cambios entre perfiles con el paso del tiempo. En este caso, se realizó el cálculo de la probabilidad de que un usuario pase de un perfil determinado a otro en la ventana de tiempo siguiente, o que pase a ser inactivo.

### 3 Conclusiones y trabajo futuro

De acuerdo a los resultados obtenidos, se consideró que el modelo propuesto es de fácil aplicación en sets de datos reales. Además, se detectó que los resultados son consistentes con trabajos relacionados y que existen coincidencias entre los datos observados en ambos sets de datos, aunque el estudio debería replicarse con sets de datos más amplios y actualizados para demostrar las tendencias observadas.

El trabajo futuro se enfocará en ampliar los aspectos a considerar para mejorar el modelo de clasificación, tanto para incluir nuevos perfiles como para redefinir las métricas utilizadas para la clasificación, considerando:

- Temas sobre los que se pregunta. Los cambios constantes en las tendencias pueden motivar cambios de comportamiento, como aumento o disminución de la cantidad de preguntas realizadas, que nuestro modelo no ha tenido en cuenta y para el que pueden plantearse futuras extensiones.
- Otros factores de calidad de la pregunta. En este trabajo tomamos como medida de calidad de la pregunta la percepción de la comunidad mediante votos positivos y negativos y marcas de favorito. Para un trabajo futuro se podría considerar si una pregunta ha sido borrada o editada por un moderador, que

podría ser indicador de baja calidad en cuanto a redacción, utilidad o por ser una pregunta repetida.

- Factor de interés de la pregunta. En este trabajo solo se tuvo en cuenta la percepción de calidad de la pregunta, sin embargo, las marcas de favorito podrían considerarse como una medida de interés o utilidad, junto con otros aspectos no tenidos en cuenta en el trabajo actual, como la cantidad de respuestas y comentarios que se generan a partir de una pregunta.

## Referencias

1. Baltadzhieva, A., Chrupala, G.: Question quality in community question answering forums: a survey. *SIGKDD Explor.* **17**(1), 8–13 (2015). <https://doi.org/10.1145/2830544.2830547>, <https://doi.org/10.1145/2830544.2830547>
2. Furtado, A., Andrade, N., Oliveira, N., Brasileiro, F.V.: Contributor profiles, their dynamics, and their importance in five q&a sites. In: Bruckman, A.S., Counts, S., Lampe, C., Terveen, L.G. (eds.) *Computer Supported Cooperative Work, CSCW 2013*, San Antonio, TX, USA, February 23-27, 2013. pp. 1237–1252. ACM (2013). <https://doi.org/10.1145/2441776.2441916>, <https://doi.org/10.1145/2441776.2441916>
3. Martínez Carod, N., Aranda, G., Zoratto, V., Murray, C.: Una propuesta para clasificación de roles de usuarios en foros de discusión técnicos. In: *XXV Congreso Argentino de Ciencias de la Computación (CACIC 2019)*. pp. 836 – 845. UniRío Editora, Río Cuarto, Córdoba - Argentina (2019)
4. Shoji, Y., Fujita, S., Tajima, A., Tanaka, K.: Who stays longer in community QA media? - user behavior analysis in cqa -. In: Liu, T., Scollon, C.N., Zhu, W. (eds.) *Social Informatics - 7th International Conference, SocInfo 2015*, Beijing, China, December 9-12, 2015, Proceedings. *Lecture Notes in Computer Science*, vol. 9471, pp. 33–48. Springer (2015). [https://doi.org/10.1007/978-3-319-27433-1\\_3](https://doi.org/10.1007/978-3-319-27433-1_3), [https://doi.org/10.1007/978-3-319-27433-1\\_3](https://doi.org/10.1007/978-3-319-27433-1_3)
5. Slag, R., de Waard, M., Bacchelli, A.: One-day flies on stackoverflow - why the vast majority of stackoverflow users only posts once. In: Penta, M.D., Pinzger, M., Robbes, R. (eds.) *12th IEEE/ACM Working Conference on Mining Software Repositories, MSR 2015*, Florence, Italy, May 16-17, 2015. pp. 458–461. IEEE Computer Society (2015). <https://doi.org/10.1109/MSR.2015.63>, <https://doi.org/10.1109/MSR.2015.63>
6. Squire, M., Funkhouser, C.: "a bit of code": How the stack overflow community creates quality postings. In: *47th Hawaii International Conference on System Sciences, HICSS 2014*, Waikoloa, HI, USA, January 6-9, 2014. pp. 1425–1434. IEEE Computer Society (2014). <https://doi.org/10.1109/HICSS.2014.185>, <https://doi.org/10.1109/HICSS.2014.185>
7. Wang, S., Lo, D., Jiang, L.: An empirical study on developer interactions in stackoverflow. In: Shin, S.Y., Maldonado, J.C. (eds.) *Proceedings of the 28th Annual ACM Symposium on Applied Computing, SAC '13*, Coimbra, Portugal, March 18-22, 2013. pp. 1019–1024. ACM (2013). <https://doi.org/10.1145/2480362.2480557>, <https://doi.org/10.1145/2480362.2480557>

# Desarrollo de un entorno basado en programación en bloques para robots educativos

Juan Martín de la Fuente

Jorge P. Rodríguez    Laura A. Cecchi

Grupo de Investigación en Lenguaje e Inteligencia Artificial  
Facultad de Informática  
Universidad Nacional del Comahue  
{juan.delafuente, j.rodrig, lcecchi}@fi.uncoma.edu.ar

TESIS DE LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN

**Palabras Clave:** Robótica Educativa . Programación Visual . Programación por Bloques . IDE . Frankestito

La incorporación de la enseñanza de las Ciencias de la Computación en la educación obligatoria ha pasado de ser una necesidad reconocida, a una realidad impulsada por el amplio consenso existente alrededor de las ventajas que proveen los aprendizajes en este campo[2,5,8,9,14,15]. A nivel nacional, se han establecido los Núcleos de Aprendizajes Prioritarios por parte del Ministerio de Educación, Ciencia y Tecnología en la búsqueda de integrar la programación y la robótica en todos los estadios de la educación [4]. En igual medida en la provincia de Neuquén, el ministerio de Educación, mediante el programa *Educación Digital* brinda capacitación y recursos a los docentes para utilizar en las aulas.

En este sentido, en los últimos años se han desarrollado herramientas que buscan facilitar el acceso a la programación, recurriendo a gramáticas gráficas que permiten mediante la manipulación de figuras generar un programa ejecutable, reduciendo la complejidad sintáctica propia de los lenguajes de programación *textual*. Entre estos lenguajes de programación visual, se han destacado en mayor medida los *lenguajes de programación por bloques*, cuyas sentencias se representan con bloques gráficos encastrables, que permiten la construcción de programas complejos mediante la interconexión de los mismos [13].

Por otra parte, las experiencias educativas que involucran la programación de robots han demostrado ser un mecanismo efectivo para la enseñanza de la programación y de otros conceptos afines a las Ciencias de la Computación que pueden ser difíciles de percibir [1,3,7,10].

En el ámbito regional, la Facultad de Informática de la Universidad Nacional del Comahue desarrolló una familia de robots educativos cuyo nombre de fantasía es *Frankestito*[6], que son de bajo costo, con licencias de código abierto y que cuentan con la capacidad de movimiento, visión y comunicación WiFi. Estos robots han probado ser un vehículo efectivo para la enseñanza de este campo disciplinar en actividades que involucran tanto alumnos como docentes de la escuela media desde el año 2013.

Para facilitar la adopción y utilización de estos robots educativos en diversos establecimientos se diseñó una Plataforma Web de Robótica Educativa llamada *FrankLab* [12], que establece mecanismos para desarrollar *Laboratorios* y diversos *Entornos de desarrollo* para programar robots compatibles con Frankestito. *FrankLab* incorpora también el primer Entorno de Desarrollo Integrado (IDE) propio para controlar los robots mediante un lenguaje de programación textual desde un navegador web, facilitando la portabilidad de la herramienta [12].

Sin embargo hasta el momento esta herramienta, que ha probado ser un dispositivo efectivo en la enseñanza de la programación y se ha establecido como un referente regional, encuentra limitaciones en el público objetivo y en el rango etario abarcable, al contar únicamente con lenguajes de programación textuales y con entornos de programación complejos en su utilización. El universo potencial de estudiantes alcanzables por estas experiencias y por los robots Frankestito se ve limitado debido a las dificultades inherentes de este tipo de programación, como la imposibilidad de realizar el seguimiento o la traza del código en ejecución, la necesidad de incorporar conocimiento propio del lenguaje a la vez que se comprenden los fundamentos de la programación y las dificultades propias de utilizar una sintaxis específica con palabras claves en un lenguaje no cotidiano.

Existen desarrollos que extienden lenguajes y gramáticas de programación por bloques con foco en la utilización de robots para hacer frente a las dificultades presentadas por los lenguajes de programación textuales, tales como LEGO WeDo, PicoBoard y OpenRoberta. Sin embargo, éstos no se adaptan a las necesidades de Frankestito, dado que cuentan con una funcionalidad más limitada, están discontinuadas o no presentan la conectividad adecuada.

En este trabajo se propone la definición, el diseño, y la implementación de un lenguaje de programación gráfico y de una herramienta Web, basada en modelo cliente-servidor, que permita la programación de robots compatibles con Frankestito a través de este lenguaje de bloques encastrables.

Para ello se extiende la gramática definida por *Blockly* [11], incorporando elementos que permiten la generación de instrucciones propias de los robots Frankestito. En este sentido, se agregan los bloques gráficos necesarios y su correspondiente interpretación a lenguaje textual. En la Figura 1 se presenta un programa construido con los bloques del lenguaje definido.

Además la herramienta incorpora ciertas características para facilitar su utilización, como conocer qué bloque de código se está ejecutando en cada momento, la posibilidad de guardar o exportar el programa logrado y permitir la traducción del lenguaje de bloques a lenguajes textuales.

Como resultado se ha desarrollado un entorno de programación por bloques, fácil de utilizar, integrado a la Plataforma de Robótica Educativa *FrankLab* que permite a los usuarios manipular los robots físicos mediante bloques encastrables, observar la visión del robot, guardar y compartir los programas creados y obtener una traducción a un lenguaje textual compatible con *Frankestito* para acompañar el proceso de crecimiento y aprendizaje personal de los alumnos. En la Figura 2 se muestra una vista general de la interfaz web.

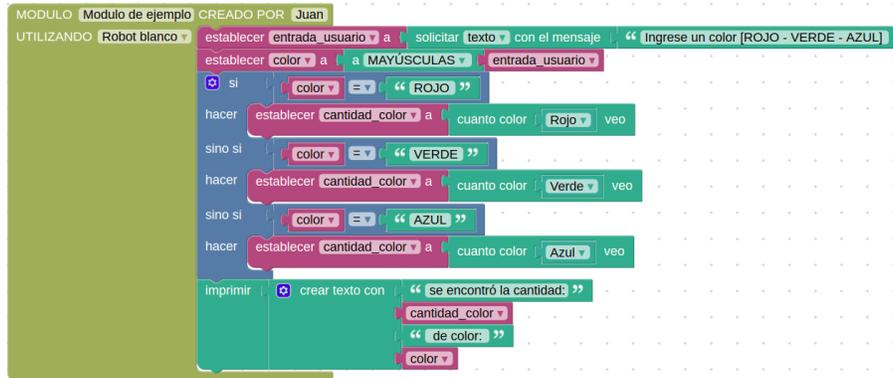


Figura 1. Ejemplo de uso del lenguaje de bloques compatible con Frankestito

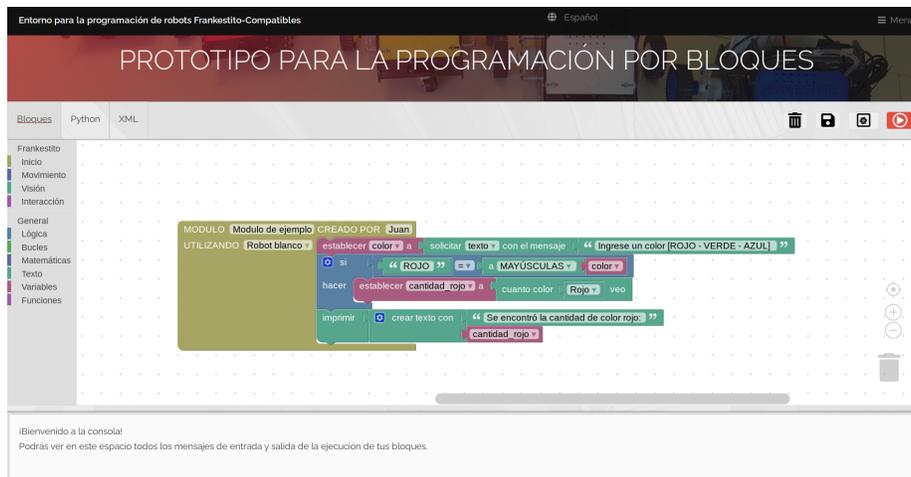


Figura 2. Vista general de la interfaz web

Gracias a su diseño modular, y su licencia *open source* se puede seguir expandiendo su funcionalidad y desarrollar trabajos futuros relacionados con la misma, incorporando más funcionalidades que faciliten el trabajo colaborativo, que favorezcan la incorporación de conocimientos sobre Inteligencia Artificial o que reduzcan los límites de su utilización.

## Referencias

1. Benitti, F.B.V.: Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education* **58**(3), 978–988 (2012)
2. Committee, K..C.S.F.S.: K-12 computer science framework (2016). <https://doi.org/10.1145/3079760>
3. Conde, M., Rodríguez-Sedano, F.J., Fernández-Llamas, C., Gonçalves, J., Lima, J., García-Peñalvo, F.J.: Fostering steam through challenge-based learning, robotics, and physical devices: A systematic mapping literature review. *Computer Applications in Engineering Education* **29**(1), 46–65 (2021)
4. Consejo Federal de Educación de la República Argentina: Núcleos de Aprendizaje Prioritarios para Educación Digital, Programación y Robótica (NAP). Resolución N°343/18 (2018)
5. Consejo Provincial de Educación de la Provincia de Neuquén: Diseño Curricular Jurisdiccional de los tres primeros años de la Escuela Secundaria Neuquina. Resolución N°1463/18 (2018)
6. Grosclaude, E., Zurita, R., Castillo, R.d., Lechner, M., Riquelme, J.: Designing a myro-compatible robot for education as copyleft hardware. In: XX Congreso Argentino de Ciencias de la Computación (2014)
7. Kert, S.B., Erkoç, M.F., Yeni, S.: The effect of robotics on six graders' academic achievement, computational thinking skills and conceptual knowledge levels. *Thinking Skills and Creativity* **38**, 100714 (2020)
8. McGarr, O., Johnston, K.: Curricular responses to computer science provision in schools: current provision and alternative possibilities. *The Curriculum Journal* **31**(4), 745–756 (2020)
9. Ottestad, G., Gudmundsdottir, G.B.: Information and communication technology policy in primary and secondary education in europe. *Second handbook of information technology in primary and secondary education* pp. 1343–1362 (2018)
10. Papert, S.: *Mindstorms: Computers, children, and powerful ideas*. Basic books (1980). <https://doi.org/10.5555/1095592>
11. Pasternak, E., Fenichel, R., Marshall, A.N.: Tips for creating a block language with blockly. In: 2017 IEEE Blocks and Beyond Workshop (B B). pp. 21–24 (Oct 2017). <https://doi.org/10.1109/BLOCKS.2017.8120404>
12. Picucci, M.: *Ambiente Web para la Programación del Robot Frankestito*. Tesis de grado - licenciatura en ciencias de la computación, Facultad de Informática - Universidad Nacional del Comahue (2021)
13. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., Kafai, Y.: Scratch: Programming for all. *Commun. ACM* **52**(11), 60–67 (Nov 2009). <https://doi.org/10.1145/1592761.1592779>
14. Smith, M.: *Computer science for all*. The White House (2016)
15. Tissenbaum, M., Ottenbreit-Leftwich, A.: A vision of k-12 computer science education for 2030. *Communications of the ACM* **63**(5), 42–44 (2020)

# Modelado Bottom-up de Variedad de Dominio en Sistemas Big Data

Líam Osycka

Tesis de Licenciatura en Sistemas de Información  
liam.osycka@est.fi.uncoma.edu.ar

**Abstract.** Para identificar variedades presentes dentro de un dominio, este trabajo de Tesis propone al proceso *B-VIP* (Bottom-up Variety Identification Process), el cual parte de las actividades del ciclo vida de *big data* y documenta las variedades encontradas en cada etapa haciendo uso de hojas de datos funcionales (*datasheets*), almacenadas en una base de conocimiento. Para dar soporte al proceso, se diseña e implementa la herramienta *CoVaMaT*, que permite nutrir esta base de conocimiento del dominio con el almacenamiento de las *datasheets* confeccionadas y los casos trabajados. Todo esto puede ser recuperado en situaciones futuras similares para un posible reuso.

*B-VIP* fue aplicado en dos casos de estudio, uno dentro del dominio de aguas superficiales (foco en ríos) con el objetivo de encontrar causas de variación de la temperatura del agua; y otro en conjunto con INTA Alto Valle sobre aguas subterráneas (foco en acuíferos libres), donde se analizaron las causas e incidencias de cambios en el nivel freático.

**Keywords:** Variedad en Sistemas Big Data · Reusabilidad · Analítica de Datos

## 1 Motivación

Big data se refiere a sistemas que manipulan grandes volúmenes de datos y que cumplen con las propiedades conocidas como las *5v's*: volumen, velocidad, variedad, veracidad y variabilidad. Particularmente, la variedad es la que resalta la heterogeneidad presente en los datos, es decir, cómo y de qué manera son diferentes entre sí. Esta diversidad puede estudiarse desde el punto de vista de la estructura que poseen, el contenido que representan, los tipos de fuentes que se manejan y del procesamiento que reciben los datos. Estos tipos de variedad pueden extenderse si se los vincula al dominio en el que se los encuentran mediante el agregado de semántica, es decir, qué tipos de estructura, de contenido, de fuentes y de procesamiento suelen utilizarse bajo un dominio en particular. A esto denominamos *variedad de contexto* y es utilizada para permitir la reutilización de características identificadas dentro de dominios similares. Entonces, respondiendo a la pregunta de investigación:

RQ: *¿Cómo puede identificarse la variedad de la información de dominio de manera de incorporar reusabilidad en el desarrollo de Sistemas Big Data*

(SBDs)?

esta Tesis tomó como partida una estructura de etapas asociadas al desarrollo de SBDs, instanciada en artefactos software producidos durante esas etapas, e incorporó el modelado de variedad de dominio partiendo de un proceso bottom-up. Esto significa que las características variantes fueron identificadas a partir de un proceso de análisis de datos, donde las variedades serán inferencias a ser corroboradas por expertos de dominio.

El modelo y el caso de estudio fueron asistidos por la Lic. Ayelén Montenegro (INTA Alto Valle), en el rol de experta de dominio. La Tesis fue dirigida por la Dra. Alejandra Cechich.

## 2 Una breve vista a los capítulos de la Tesis

Primeramente, se recopilaron propuestas relevantes en los tópicos referidos a variedad (sección 2.1) y reusabilidad (sección 2.2.1) en big data. Luego, a partir del trabajo realizado en el contexto de investigación del proyecto en el que se enmarca la Tesis (sección 2.3), se propuso el proceso *B-VIP* (Bottom-up Variety Identification Process) (sección 3.1). El concepto *bottom-up* circunscribe la propuesta a un análisis de variedad exploratorio: a partir de los datos se identifican distintos tipos de variedad (de fuentes, contenido, procedimiento y contexto), en las distintas actividades del ciclo de vida de un SBD. El proceso propuesto consiste en dos etapas: (1) documentación de la variedad encontrada mediante datasheets, adaptadas al caso de análisis de datos; y (2) reutilización de la variedad encontrada.

Las dos etapas del proceso fueron descritas en términos de un caso de estudio dentro del dominio hidrológico (sección 3.2.1), particularmente sobre datos de cuerpos de aguas superficiales (ríos). El proceso propuesto y este caso de estudio dieron lugar a dos publicaciones:

Memorias del XXVII Congreso Argentino de Ciencias de la Computación (CACIC), pag. 367–376.: L. Osycka, A. Buccella, and A. Cechich, “Identificación de variedad contextual en modelado de sistemas big data”.

Una versión extendida fue publicada como capítulo en el libro 27th Argentine Congress, CACIC 2021, Salta, Argentina, October 4-8, 2021, Revised Selected Papers. Communications in Computer and Information Science 1584, Springer 2022, ISBN 978-3-031-05902-5, pages 124-138. L. Osycka, A. Buccella, and A. Cechich, “Data Variety Modeling: A Case of Contextual Diversity Identification from a Bottom-up Perspective”.

Luego, un segundo caso de estudio, realizado en conjunto con el INTA Alto Valle, permitió mostrar la viabilidad de la propuesta en otra problemática del

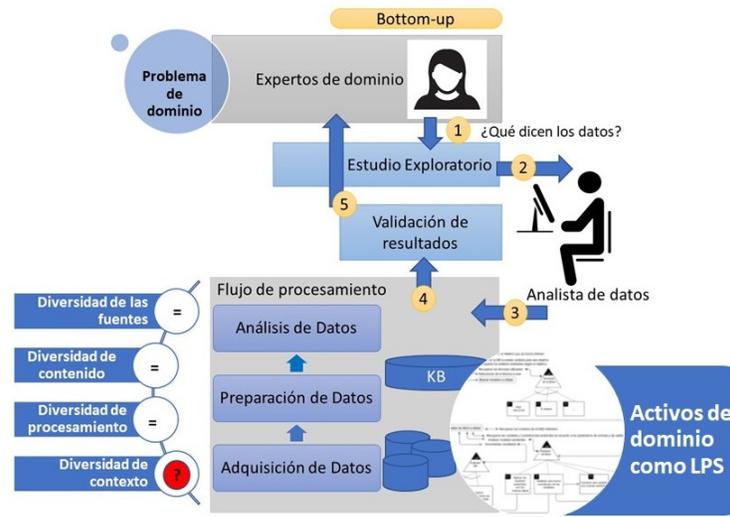


Fig. 1. Enfoque bottom-up para identificación de variedad

dominio hidrológico (sección 3.2.2): la caracterización de variedad en la fluctuación del nivel freático, dentro del dominio de aguas subterráneas.

Considerando que una herramienta de soporte es esencial para gestionar la variedad en la práctica, se diseñó e implementó una primera versión de CoVaMaT (Context-based Variety Management Tool), cuya arquitectura lógica fue modelada con el lenguaje ArchiMate (sección 4.1). CoVaMaT tiene la propiedad de exponer el conocimiento adquirido dentro de un dominio de una forma declarativa a través de documentos JSON y puede servir de soporte tanto para B-VIP, como para un posible proceso de identificación top-down (T-VIP). La arquitectura de la herramienta a nivel de aplicación, incluyendo decisiones de diseño e infraestructura para el desarrollo del prototipo, fue presentada en la sección 4.2. El capítulo 4 también describe el modelo de datos propuesto para el almacenamiento de elementos y relaciones, así como el desarrollo y ejecución de un ejemplo.

CoVaMaT fue utilizada en el caso de aguas superficiales (ríos), mostrando cómo se puede almacenar la variedad identificada luego de la aplicación de B-VIP. El diseño de la herramienta dio lugar a otra publicación en:

Proceedings of Cloud Computing, Big Data & Emerging Topics - 11th Conference, JCC-BD&ET 2023, pages 57-74: Lám Osycka, Alejandra Cechich, Agustina Buccella, Ayelén Montenegro, Angel Muñoz, “CoVaMaT: Functionality for Variety Reuse Through a Supporting Tool”.

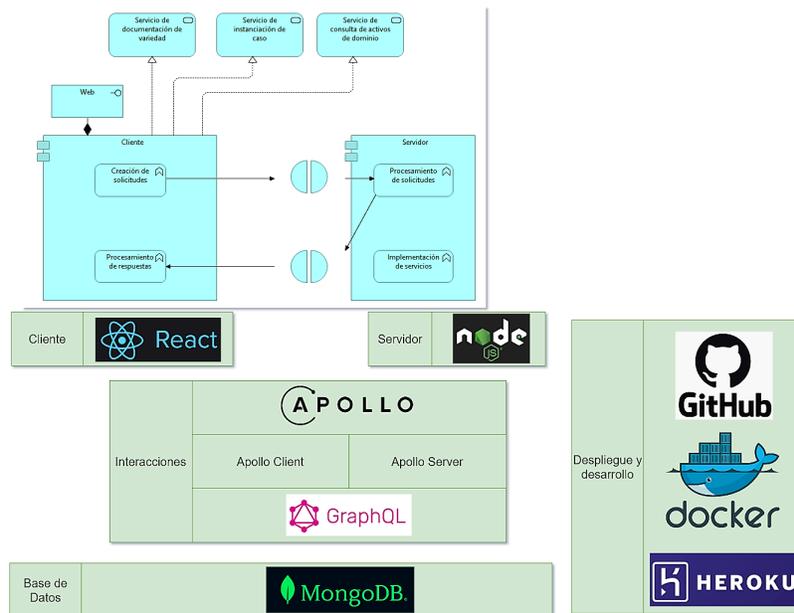


Fig. 2. Plataforma específica de aplicación y tecnologías seleccionadas para CoVaMaT

### 3 Conclusiones y Trabajos Futuros

Identificar variedad en diversos niveles (fuentes, contenido, procedimiento, contexto) requiere de habilidades distintas pero complementarias. Por ejemplo, la detección de variedad en procesamiento está ligada a tareas de analítica de datos, con decisiones sobre tipo de técnicas a aplicar en los diversos análisis, que muchas veces son dependientes del tipo y cantidad de datos disponibles; mientras tanto, la variedad en el contexto se relaciona con la semántica del entorno que en el que el fenómeno a estudiar está inmerso, relacionado con un dominio de aplicación que es inherentemente complejo (como el caso del dominio hidrológico trabajado).

La representación de documentación sobre casos de variedad no es el caso tradicional de construir software que recupere datos en una base de datos. Eso lleva a que, desde su concepción, el sistema se conciba como declarativo, con la capacidad de guardar conocimiento. Las decisiones sobre cómo guardar ese conocimiento, indudablemente impactarán en la facilidad para recuperarlo. De ahí que la construcción de esta herramienta se haya detenido en analizar alternativas al modelo lógico, de manera que se refuercen atributos de calidad para su extensibilidad, facilidad de modificación e interoperabilidad, entre otros.

Estas consideraciones dan lugar a trabajo futuro, por ejemplo: (1) uso de estándares de dominio, (2) mejora en la función de detección de similitud de casos y (3) mejoras en el prototipo *CoVaMaT*.

# Clasificación de posts en foros de discusión

Jaime Daniel Saldivia Urra, Gabriela N. Aranda, Valeria Zoratto

Tesis de Licenciatura en Ciencias de la Computación  
Facultad de Informática, Universidad Nacional del Comahue  
Neuquén, Argentina

**Abstract.** Los foros de discusión en línea son plataformas populares donde los usuarios buscan respuestas, comparten opiniones y discuten problemas. Dentro de dichos foros hay hilos de discusión, que son secuencias de mensajes escritos por diferentes usuarios, cada uno con un propósito único. Estos hilos son de interés tanto para sus autores como para otros usuarios de la Web con inquietudes similares; lo que ha dado lugar a una rama de investigación sobre la recuperación y clasificación de información en este tipo de contenidos. Un tema relevante en esta área es la identificación de los mensajes que plantean preguntas y aquellos que ofrecen soluciones. También se considera interesante detectar otros tipos de mensajes que enriquecen la discusión (como los que solicitan detalles adicionales o comparten experiencias) y los mensajes irrelevantes o "basura". Para abordar la clasificación de mensajes en hilos de discusión, trabajos relacionados han propuesto varios modelos que utilizan diversas características, como el contenido del mensaje, la estructura del hilo y la información de los usuarios, y se ha investigado la importancia relativa de dichas características. Esta tesis se enfoca en definir un conjunto de características para clasificar mensajes en foros de discusión según su rol en el hilo. En la misma se compara y extienden modelos existentes, y se evalúa la propuesta en conjuntos de datos reales, obteniendo resultados prometedores.

**Keywords:** Foros de discusión · Recuperación de información · Reuso de información.

## 1 Introducción

Los foros de discusión técnicos se han transformado en bitácoras actualizadas de conocimiento sobre temas específicos, que están en la mira de varios proyectos de investigación enfocados en aplicar estrategias de recuperación de información (RI). Los usuarios que participan en dichos foros comparten conocimientos y experiencias que son de gran utilidad cuando un problema surge repetidamente. Para facilitar la tarea de reutilización de dicho conocimiento, es necesario detectar los hilos de discusión que contienen las soluciones más adecuadas para un problema particular, analizando los debates que suceden en uno o varios foros de discusión a la vez.

Además, una vez detectados los hilos de discusión que pueden estar relacionados al problema planteado, es necesario detectar cuál o cuáles de los mensajes

(posts) que contienen dichos hilos presentan las posibles soluciones. Esto no es un problema menor, ya que un hilo no es más que es una cadena de mensajes que parten de una pregunta que da inicio a la discusión.

De acuerdo a la clasificación propuesta por Bathia y otros [1], además de los mensajes que plantean la pregunta principal y de los mensajes que aportan soluciones o dan pistas para resolverlo, a lo largo de una cadena de mensajes se pueden encontrar otros tipos de mensajes, como los que piden más detalles sobre el problema, los que repiten una pregunta anterior, los que dan más detalles del contexto, los que reportan buenas o malas experiencias aplicando las soluciones propuestas, así como otros que no aportan más que confusión o ruido a la discusión, comúnmente llamados mensaje basura.

De acuerdo a la revisión bibliográfica realizada, existen varios trabajos enfocados en identificar el rol de un post dentro del hilo de discusión, empleando características de distintos tipos (basadas en el contenido de los posts, estructura de la secuencia, comportamiento de los usuarios participantes, análisis de opinión, etc).

Siguiendo con esta línea de trabajo, esta tesis se enfocó en la definición de un conjunto de características que permite clasificar mensajes de foros de discusión de acuerdo a su rol en el hilo. Para ello, se analizaron y compararon distintos modelos provenientes del análisis de literatura relacionada y se propuso extender uno de ellos, con el objetivo de mejorar la clasificación obtenida por dicho modelo. Luego se evaluó su performance en dos sets de datos provenientes de hilos de discusión reales obteniéndose resultados prometedores.

Esta tesis fue desarrollada por el estudiante Jaime Saldivia, entre 2020 y 2023, dirigido por la Dra. Gabriela Aranda y la Lic. Valeria Zoratto de FAI-UNCOMA. La misma se ha enmarcado en los Proyectos de investigación “Reuso de conocimiento en Foros de Discusión Técnicos - Parte II”, subproyecto del Programa 04/F009 “Desarrollo de Software Basado en Reuso - Parte II” (Periodo 2017-2021) y en el Proyecto 04/F018 “Reuso de Información en Comunidades Virtuales” (Periodo 2022-2025), desarrollados por el Grupo de Investigación GI-ISCo de la Facultad de Informática, Universidad Nacional del Comahue.

## 2 Contenido de la tesis

En el Capítulo 2 se presentaron y evaluaron varios trabajos previos relacionados [3][4][2][1][5], considerando las elecciones de modelado del problema, los atributos de información incluidos en los modelos propuestos, el modo de cálculo de las métricas, los métodos de clasificación utilizados, el rendimiento de los clasificadores y el modo de evaluación de los resultados obtenidos. A continuación se definió como estrategia de investigación, utilizar el trabajo de Bhatia y otros [1] como modelo base y extenderlo incorporando un conjunto de características adicionales para cada categoría, basado en observaciones de los restantes trabajos relacionados estudiados. La elección de dicho modelo se basó en que el trabajo de Bhatia y otros considera un conjunto más amplio de características, calculables en base a atributos disponibles en todos los foros de discusión, y que los autores

dejaron disponibles de forma abierta y pública los sets de datos utilizados, con toda la información necesaria de los hilos manualmente clasificados.

A continuación, en el Capítulo 3, se detallada la propuesta de mejora sobre el modelo propuesto por Bhatia y otros, al que se le agregaron 6 características tomadas de trabajos relacionados y 2 características propias. Luego, se diseñó e implementó un componente software en lenguaje Java para el cálculo de las características del modelo extendido propuesto. Dicha componente incluye módulos para la realización de operaciones de recuperación y limpieza inicial del texto de los archivos XML del set de datos, cálculo de las características para cada hilo del conjunto de posts y salida de resultados con formato adecuado para la herramienta de clasificación. Dicho componente software permite obtener una salida de valores para el conjunto de características de todos los hilos de un set de datos en un archivo de formato arff. Esta salida de datos se utiliza en una segunda etapa para la clasificación de propiamente dicha con la herramienta Weka. A fin de validar el correcto funcionamiento de dicho componente, se tomaron los sets de datos originales (utilizados en el trabajo de Bhatia y otros [1], y se verificaron los resultados de la clasificación para que sean lo más cercanos posibles a los obtenidos por dichos autores, considerando como limitaciones que no contábamos con las fórmulas detalladas de todas las características originales y que no contábamos con acceso a las librerías originales utilizadas para el cálculo de Similitud Coseno y de Sentiment Score.

Finalmente, en el Capítulo 4, el componente de preprocesamiento y cálculo de características se utilizó para preparar un archivo arff para cada set de datos, que luego fueron utilizados como entrada para la herramienta Weka, aplicando el clasificador Logistic Regression utilizado en el trabajo original. Los resultados de la clasificación con el conjunto inicial de 20 características de Bhatia y otros se tomó como caso base, y se evaluaron 15 combinaciones incluyendo las 8 características agregadas como parte de nuestra propuesta de extensión. Como resultado se pudo observar que en ambos casos de estudio casi todas las combinaciones lograron mejor performance que el caso base. Además, respecto al rendimiento de las características en forma individual, se pudo observar que 3 de las características agregadas aparecen entre las 15 más relevantes en ambos casos de estudio. En el sitio <https://github.com/forumadvisor/corpus-posts-classification> se encuentran publicados los set de datos calculados.

### 3 Conclusiones y trabajo futuro

Los resultados obtenidos con la aplicación del modelo extendido en ambos sets de datos utilizados, han demostrado mejorar en casi la mayoría de las pruebas al caso base (igualandola en los restantes). Por dicho motivo consideramos interesante extender el trabajo realizado en las siguientes líneas de investigación a futuro:

- Minimizar el conjunto de características del modelo, quitando las características de rendimiento solapado, a fin de lograr un conjunto de características independientes entre sí.

- Eliminar o modificar el cálculo de características que parecen tener un desempeño igual: Por ejemplo, las características PostLength y PostLength-Stemmed, que calculan cantidad de palabras antes y después de aplicar Stemming dan resultados iguales para todos los posts analizados, por lo que creemos interesante la creación y evaluación de alternativas a dichas características.
- Utilizar el modelo sobre un conjunto de hilos de un foro que disponga de calificaciones de usuarios y de su contenido (reputación, votos positivos y negativos, marcas de pregunta favorita) y evaluar si contar con características de este tipo pueden mejorar la clasificación.

## Referencias

1. Bhatia, S., Biyani, P., Mitra, P.: Identifying the role of individual user messages in an online discussion and its use in thread retrieval. *Journal of the Association for Information Science and Technology* **67**(2), 276–288 (2016)
2. Catherine, R., Singh, A., Gangadharaiah, R., Raghu, D., Visweswariah, K.: Does similarity matter? the case of answer extraction from technical discussion forums. In: *Proceedings of COLING 2012: Posters*. pp. 175–184. The COLING 2012 Organizing Committee, Mumbai, India (2012), <https://aclanthology.org/C12-2018>
3. Feng, D., Shaw, E., Kim, J., Hovy, E.: Learning to detect conversation focus of threaded discussions. In: *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*. pp. 208–215. Association for Computational Linguistics (2006)
4. Hong, L., Davison, B.D.: A classification-based approach to question answering in discussion boards. In: *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. pp. 171–178. ACM (2009)
5. Obasa, A.I., Salim, N., Khan, A.: Hybridization of bag-of-words and forum metadata for web forum question post detection. *Indian Journal of Science and Technology* **8**(32), 1–12 (2016)

# Un enfoque para la unificación de UML 2.4.1, ORM 2 y EER

Federico Solorza<sup>1</sup>

Giuliano Marinelli<sup>1,2</sup>    Laura A. Cecchi<sup>1</sup>

<sup>1</sup> Grupo de Investigación en Lenguaje e Inteligencia Artificial  
Facultad de Informática - Universidad Nacional del Comahue

<sup>2</sup> CONICET

`federico.solorza@est.fi.uncoma.edu.ar`  
`{giuliano.marinelli,lcecchi}@fi.uncoma.edu.ar`

TESIS DE LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN

**Palabras Clave:** Representación de conocimiento . Lenguajes visuales . Ontologías . Razonamiento

En Ciencias de la Computación, se entiende por ontología a una representación formal del conocimiento, mediante la determinación de conceptos y relaciones entre ellos que le es más fiel al dominio que describe. UML 2.4.1 (de ahora en más diremos simplemente UML), ORM 2, y las variantes de EER han probado ser útiles entre sus dominios de aplicación. Sin embargo, distintos desafíos deben atenderse para lograr integrar modelos representados en dichos lenguajes. Esto se debe a la virtud por la cual se modela un dominio con determinados requerimientos con un lenguaje en particular. El uso de ontologías es importante en diversas áreas en lo que respecta a las Ciencias de la Computación: la Web Semántica, representación del conocimiento (siendo esencial en aplicaciones como lenguaje natural, integración de datos, etc), y en general en la Inteligencia Artificial; algunos ejemplos más particulares, como medio para interoperar diferentes bases de datos; empleadas para estandarizar concepciones de dominios ampliamente utilizados por organizaciones, verificación en la integración de nuevo conocimiento, permitir el uso de razonadores debido a su representación en lógica descriptiva (DL), entre otras. Los conceptos y relaciones de una ontología son pertinentes y convencionales al punto tal que permiten inferir conocimiento implícito entre ellos. Esto es posible debido a que OWL tiene su equivalente en DL, cuyos razonadores son capaces de hacer inferencias a partir de una base de creencias. Esta cualidad, le permite al sistema razonar sobre una ontología y por lo tanto ser capaz de responder consultas. Mediante un lenguaje que subsuma las capacidades de los tres mencionados anteriormente, se podría habilitar razonar a través de los modelos que realizamos en diferentes lenguajes para un mismo dominio. El metamodelo KF [3] fue desarrollado para permitir representar las primitivas de UML, ORM 2 y las variantes de EER. En un primer momento, el metamodelo KF fue representado utilizando UML en conjunto con reglas de

traducción junto con definiciones de sus primitivas en FOL [2]. No obstante, y por un lado, UML no es capaz de representar la totalidad de las capacidades de EER, ni tampoco la diversidad de restricciones que encontramos en ORM 2; por su parte, FOL es muy complicado para servir de vector de comunicación entre distintos individuos que desean intercambiar conocimiento. En la Fig-

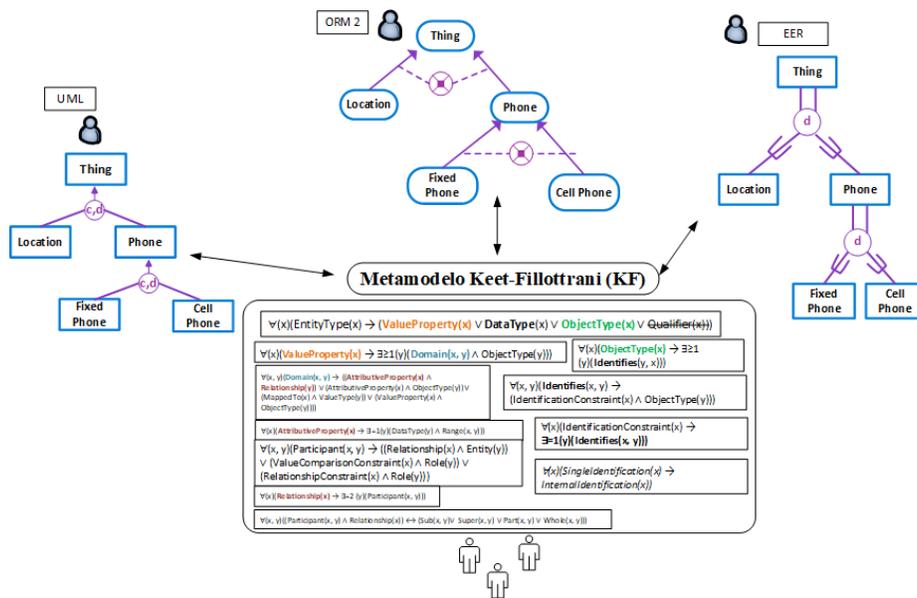


Figura 1. Heterogeneidad en los lenguajes mediados por el metamodelo KF

ra 1 se aprecia al metamodelo KF como un lugar donde convergen las diferentes maneras de representar un dominio con la heterogeneidad que es propia en el desarrollo de un sistema complejo, requiriendo interoperabilidad e integración entre otras funcionalidades. Esto motivó el desarrollo de un lenguaje visual para permitir emplear el metamodelo KF con menor esfuerzo, reduciendo, significativamente, la necesidad de agregar elementos coloquiales para que el modelo sea interpretado de la forma pretendida (i.e. en la determinidad que el modelador supo comprender). El uso de los símbolos que componen el lenguaje visual para el metamodelo KF intensifica su uso, debido a que se requiere una menor carga cognitiva para entenderlo y aprovecharlo.

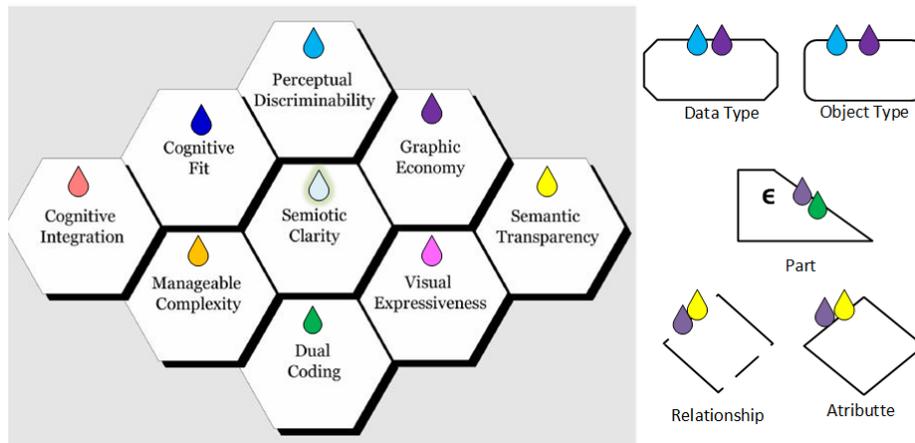
Esta tesis tiene como motivación principal el desarrollo de un lenguaje visual unificador, basado en el metamodelo KF, integrando los modelos UML 2.4.1, ORM 2 y ER/EER. El nuevo lenguaje visual se inspira en ideas provenientes de los principales referentes en el tema, como por ejemplo Peirce [6], Moody [5], Colin Ware [9,8] y Bertin Jacques [1], entre otros. En este sentido, éste está compuesto de diversos símbolos estratégicamente diseñados, para que su signi-

ficado semiótico se canalice de manera determinada y en concordancia con los tres lenguajes de modelado conceptual (UML, ORM 2 y EER). En general, el símbolo en sí mismo es una representación y encapsula ideas cuya interpretación puede venir definida por la manera en que se lo utilizó. Debido a esto, el lenguaje visual lo sustituye pero con un estilo similar a algunos símbolos propios de los mencionados lenguajes de modelado conceptual y agrega otros nuevos para articular una coherencia semiótica entre los tres lenguajes. La motivación para esta elección, es la densidad simbólica del lenguaje y su uso en el tiempo. En este sentido, en ORM 2 existen símbolos para propiedades que son poco utilizadas, como por ejemplo restricción “Strong Intransitivity”, por lo que se decidió reutilizar el símbolo base. Otra razón principal, en la elección del símbolo, es su reiterada utilización en el tiempo y va en línea con la meta de disminuir la carga cognitiva para *comprender* el significado del nuevo símbolo. Por ejemplo, en el lenguaje visual un rectángulo con bordes curvos sirve para representar la primitiva del metamodelo KF ObjectType.

La utilización de símbolos semióticamente cargados para cada una de las primitivas del metamodelo KF, permitirán que diversos profesionales puedan desarrollar sus propias ontologías en el lenguaje que le es más afín e integrarlas con otros de manera más directa y convencional. Incluso, puede advertir de posibles inconsistencias entre los términos conceptuales utilizados. Asimismo, se colabora en el entendimiento al crear un concepto. Por ejemplo, un atributo es esencialmente una relación; así lo que en un modelo, representado por ejemplo con UML, es entendido como entidad con un atributo, mediante el lenguaje visual quedará en claro, *que la entidad se relaciona con otra mediante una relación, que es justamente el atributo*. La semiótica como herramienta para desarrollar ontologías, es de una enorme influencia ya que habilita canales de comunicación estandarizados y convencionales, que encapsulan conocimiento en una unidad con una importante densidad de información[7]. Muchas de las características en el área de los símbolos para los lenguajes de modelado que se han propuesto [4] son pragmáticas, aunque *también* se alinean con el uso de la semiótica para poder fundamentar las decisiones de diseño. En la Figura 2 se puede apreciar un conjunto de características deseables y algunos símbolos del lenguaje visual definido que las cumplen. La definición y diseño de un lenguaje visual para el metamodelo KF permite el análisis de un modelo ontológico de forma más clara, considerando que el diagrama vuelve explícito el conocimiento que algunos de los lenguajes de modelado conceptual representan en forma implícita o bien no lo pueden representar.

Por otra parte, en esta tesis se avanzó con la formalización parcial en DL del metamodelo. Este resultado abre la posibilidad, como trabajo futuro, de implementar sobre una herramienta de Ingeniería Ontológica un editor de este lenguaje, que luego conecte con un razonador, para verificar en forma automática inconsistencias en el modelo ontológico.

Esta idea es novedosa dado que hasta donde conocemos, no existe ningún lenguaje visual para el metamodelo KF, ni tampoco un lenguaje visual que unifique constructores de diferentes lenguajes de modelado conceptual. La relevancia de



**Figura 2.** A la izquierda se muestran las características deseables según Moody [4] y a la derecha algunos símbolos del lenguaje visual.

contar con un lenguaje visual para el metamodelo KF radica en su aptitud para identificar posibles patrones, habilitar una forma muy potente de percepción e incluso facilitar la intervención de los modeladores al convertir constructores que no tienen equivalencia semántica entre los lenguajes de modelado conceptual. Se espera que a partir de los resultados alcanzados en esta tesis, herramientas de Ingeniería Ontológica que soporten la interoperabilidad entre los tres lenguajes mencionados, se beneficien del diseño visual y que los modeladores puedan intervenir de modo más amigable en el proceso de conversión entre lenguajes.

## Referencias

1. Bertin, J.: *Semiology of Graphics: Diagrams, Networks, Maps*. JSTOR (1985)
2. Fillotrani, P.R., Keet, C.M.: *Kf metamodel formalization* (2014). <https://doi.org/10.48550/ARXIV.1412.6545>, <https://arxiv.org/abs/1412.6545>
3. Keet, C.M., Fillotrani, P.R.: *An ontology-driven unifying metamodel of uml class diagrams, eer, and orm2*. *Data & Knowledge Engineering* (2015)
4. Moody, D.: *What makes a good diagram? improving the cognitive effectiveness of diagrams in is development*. In: *Advances in information systems development*. Springer (2007)
5. Moody, D.: *The physics of notations: toward a scientific basis for constructing visual notations in software engineering*. *IEEE Transactions on software engineering* (2009)
6. Peirce, C.S.: *The Essential Peirce, volume 2*. Harvard University Press (1992)
7. Sowa, J.F.: *Ontology, metadata, and semiotics*. In: *International conference on conceptual structures*. pp. 55–81. Springer (2000)
8. Ware, C.: *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc. (2004)
9. Ware, C.: *Visual Thinking: For Design*. Morgan Kaufmann Publishers Inc. (2008)

# Big Data en el Dominio Hidrológico: Caracterización del Contexto de Predicción de Turbidez

Gastón Vidart

Tesis de Licenciatura en Sistemas de Información  
gaston.vidart@est.fi.uncoma.edu.ar

**Abstract.** El objetivo general de esta Tesis fue caracterizar el contexto de predicción de turbidez del agua en sistemas Big Data. Para ello, se propuso un modelo de factores que puedan influenciar el análisis. El modelo se estructura como una taxonomía de variables de contexto y es instanciado en un caso de estudio, donde se analizan cuatro factores: climatología, calidad del agua, conformación de suelos y uso de los mismos. El estudio se enfocó en el Canal Principal de Riego del Alto Valle, con el objetivo de analizar relaciones entre turbidez y crecimiento de malezas. La remoción de malezas tiene un gran impacto operativo y económico, pero a éste se le suma que, si no se tratan, el suministro de agua de riego puede verse afectado y con ello también los cultivos del Alto Valle.

**Keywords:** Analítica de Datos · Big Data · Taxonomías

## 1 Motivación

El surgimiento del Big Data Geoespacial (BDG) proviene de la ubicuidad de los sistemas y sensores que generan y recolectan gran cantidad y variedad de información geográfica sobre las actividades de la vida diaria. El objetivo de BDG es investigar patrones complejos y desarrollar nuevos sistemas de soporte a las decisiones, basándose en la ubicación de los datos, proporcionando así un valor sin precedente para las ciencias, la ingeniería y los negocios.

En particular, la hidrología es la ciencia que estudia las aguas terrestres, incluyendo el origen, movimiento y distribución en el planeta, junto a sus propiedades físicas y químicas, incluyendo su interacción con el medio ambiente. Este dominio por su estrecha relación con otras ciencias geográficas como la climatología, hidrografía y la geología, genera una gran variedad de datos georreferenciados aplicables a BDG.

Las decisiones respectivas al uso apropiado para cuerpos de agua debe ser informada mediante el uso de modelos predictivos de calidad del recurso hídrico. Deben realizarse estudios para analizar el beneficio de alcanzar un uso designado en un determinado cuerpo de agua y qué costos e implicaciones requiere. La calidad del agua es medida por los cambios en parámetros químicos, ecológicos y espaciales, de los cuales además de estudiar sus valores, hay que ver sus interdependencias. Por ejemplo, la presencia y cantidad de industrias en una zona

ha mostrado afectar parámetros como el total de sólidos disueltos en el agua. Entre estos parámetros, analizar y predecir la *turbidez* de un cuerpo de agua es de gran utilidad; sin embargo no es tarea sencilla. Trabajos relacionados subrayan la importancia de analizar posibles mejoras a la predicción de turbidez, considerando variaciones en las muestras de datos, tipo de red neuronal usada y configuración. También se consideran variables de contexto, por ejemplo, incorporando modelos de erosión de suelos, a partir de datos obtenidos en casos particulares.

El modelo y el caso de estudio fueron asistidos por la Lic. Ayelén Montenegro (INTA Alto Valle), en el rol de experta de dominio. La Tesis fue dirigida por la Dra. Alejandra Cechich.

## 2 Una breve vista a los capítulos de la Tesis

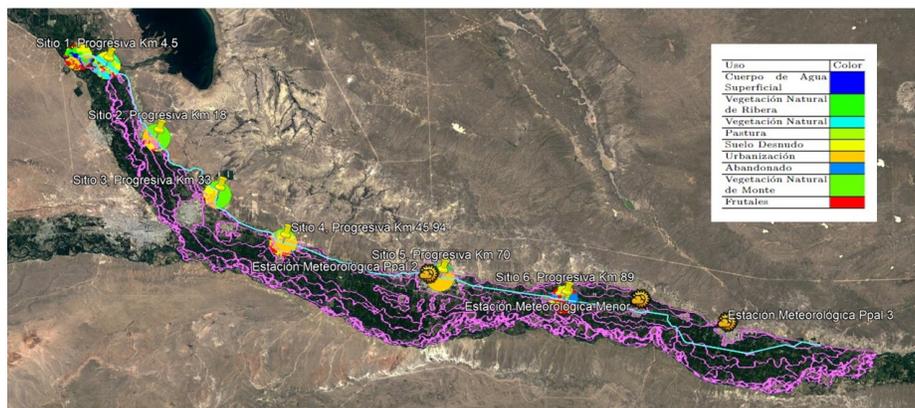
Para elaborar una clasificación del contexto (usos, variables y modelos de predicción aplicados a la turbidez del agua), se procedió al análisis de trabajos relacionados con dos aspectos principales del estado actual del arte: (1) el análisis predictivo sobre calidad del agua y (2) el análisis de turbidez en cuerpos de agua. El estudio realizado se centró en la identificación de variables que caractericen el análisis de turbidez, así como en las técnicas aplicadas. El análisis de trabajos relacionados fue detallado en el capítulo 2 de la Tesis.

Luego, con la información analizada y la contribución de expertos del dominio, se procedió a elaborar un modelo de factores que pueden influenciar el análisis de turbidez y/o su predicción, principalmente relacionados a la caracterización del entorno o contexto. En nuestro caso, las variables significativas se refirieron a diversos elementos y fenómenos que pueden caracterizar el entorno. El modelo taxonómico fue descripto en el capítulo 3 de la Tesis. Adicionalmente, los contenidos de este capítulo fueron publicados en el:

14<sup>o</sup> Congreso Argentino de AgroInformática (CAI-2022), parte de las 51<sup>o</sup> JAIIO: G. Vidart, A. Cechich, A. Buccella, A. Montenegro, “Análisis de Turbidez basado en Caracterización de Contextos”.

Para instanciar el modelo de contexto, se plantearon dos objetivos: (1) realizar un análisis de correlación de funciones ajustado a la naturaleza de los datos, buscando identificar relaciones entre ellos; y (2) agrupar los diferentes sitios de tomas de muestras en base a la similitud de las correlaciones encontradas, de acuerdo a los parámetros de la taxonomía propuesta. El primer objetivo buscaba determinar cuáles son las variables predictoras de las condiciones de crecimiento de malezas en el contexto del CPRAV; y, por otra parte, el segundo objetivo buscaba unir sitios similares, para mitigar los efectos de la escasez de datos en los sitios individuales.

Para la explotación de los datos, se enfocó en las actividades que son parte del ciclo de procesamiento de un sistema Big Data. La particularidad en el caso de estudio seleccionado fue que, aunque se apliquen las mismas actividades, el conjunto de datos disponible corresponde a las características de un repositorio pequeño o *small data set*, con lo que algunas de las técnicas fueron seleccionadas especialmente para el análisis de este tipo de repositorios. El caso seleccionado para instanciación, así como su alcance y modelo de datos usado, fueron descritos en las secciones 3.2 y 4.1 de la Tesis.



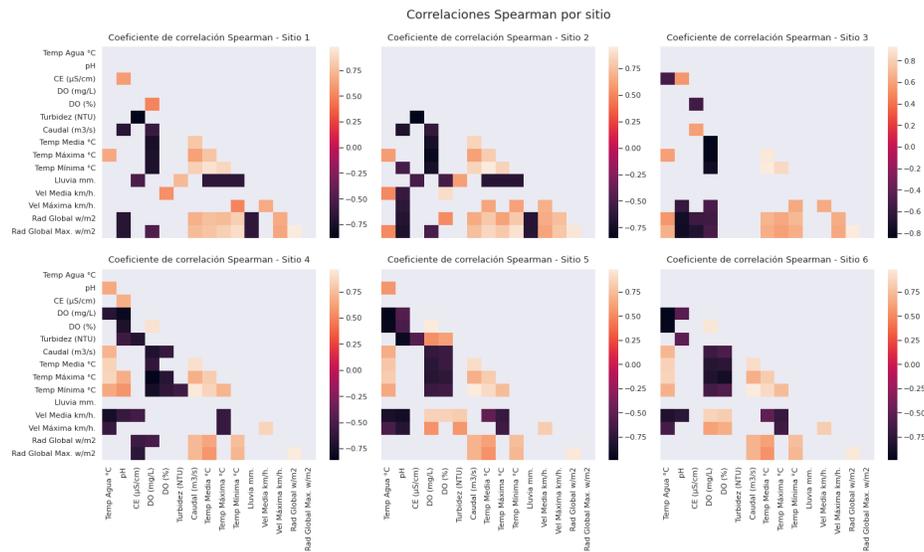
**Fig. 1.** Suelos y usos a lo largo del CPRAV

Previo a la explotación, se procedió a identificar los datos relevantes con respecto a la categorización presentada en la taxonomía. Con asistencia de la experta de dominio, los datos requeridos se dividieron en cuatro fuentes: *climáticos*, *calidad del agua*, *características de suelos y usos de suelos de la cuenca del CPRAV* (sección 4.2); y luego, previo al procesamiento, fueron ingresados a un lago de datos (sección 4.3).

Posteriormente, se realizó un proceso de preparación que incluyó tres etapas: *Conversión de archivos fuente*, *Limpieza de datos* y *Preprocesamiento de datos*. Éstas se realizaron sobre los cuatro archivos asociados a las diferentes fuentes. Las tres etapas citadas fueron detalladas en la sección 4.4. Finalmente, el análisis de los datos se realizó en dos sentidos: un análisis de correlación ajustado a su naturaleza (Small Datasets); y un análisis de clustering para encontrar agrupamientos de los diferentes sitios a lo largo del CPRAV. Ambos análisis fueron detallados en la sección 4.5.

### 3 Conclusiones y Trabajos Futuros

Cada análisis dio como resultado consideraciones interesantes para el caso de estudio. Por ejemplo, en el sitio 5 se encontró asociación con el viento, indicando



**Fig. 2.** Correlaciones de Spearman por sitio

que el viento aporta partículas al canal, aumentando el nivel de turbidez. Si bien el viento es un factor influyente en la turbidez, en el sitio 5 no se esperaba que sea relevante, ya que el suelo es principalmente urbano (85.9%) y no se esperaba que los sedimentos puedan alcanzar el canal. Respecto a la clase textural de este sitio (0.5), se encuentra que el suelo está compuesto por sedimentos finos, por lo cual la relación de turbidez con el viento puede ser una posibilidad, debido a su capacidad de transportarse por el aire, incitando a realizar mayores análisis. En este sitio el coeficiente de Kendall encontró una correlación negativa entre pH y turbidez (cuanto más alcalina, menos turbidez y viceversa). En relación a las características del sitio, la variable puede estar condicionada por el uso urbano con la posibilidad de algún nivel de perturbación química; por ejemplo: mediante la contaminación por intercambio gaseoso de  $\text{CO}_2$  con el agua, donde a mayor  $\text{CO}_2$  más ácida es el agua; o por el aumento de residuos arrojados al CPRAV.

Como trabajos futuros, podrían usarse otras técnicas de machine learning, para aumentar el conocimiento que se pueda obtener de los datos. Aquí, no sólo se plantea la predicción de turbidez con redes neuronales, sino también realizar análisis de series de tiempo que puedan predecir la turbidez a futuro y contemplen la periodicidad de los datos. De esta manera, se pueden planificar con mayor precisión las tareas de contingencia para contener la turbidez en el canal o para remover las malezas, reduciendo los riesgos de ruptura del CPRAV, la contaminación del agua de riego y el costo de las operaciones de mantenimiento realizadas sobre el mismo.

# Modelos de generación de comentarios de código basados en transformers

Cristian Martin Vincenzini<sup>1</sup>[0009-0002-1843-7983] and  
Sandra Roger<sup>1</sup>[0009-0008-9212-7514]

G.I.L.I.A. Facultad de Informática. Universidad Nacional del Comahue  
cristian.vincenzini@est.fi.uncoma.edu.ar, roger@fi.uncoma.edu.ar

**Abstract.** En este trabajo se entrena un conjunto de modelos de generación de lenguaje natural para crear comentarios en inglés a partir de código fuente. Para ello, se hace uso de una técnica de transferencia de aprendizaje sobre una arquitectura denominada *transformer*, lo que permitió optimizar modelos ya entrenados con nuevos conjuntos de datos. Posteriormente se analizaron los resultados obtenidos.

**Keywords:** Generación de Lenguaje Natural · Transformers · Machine Learning.

## 1 Introducción

En el ámbito de la programación de software, la escritura de código fuente de calidad es fundamental para garantizar la legibilidad, mantenibilidad y comprensión efectiva de los programas [1, 2]. Los comentarios de código son fragmentos de texto que se incluyen en el código fuente de un programa, pero que no se ejecutan ni afectan el funcionamiento del mismo. Su principal propósito es proporcionar documentación sobre el código escrito, facilitando la comprensión tanto para el programador original como para otros desarrolladores que puedan leer o modificar el código en el futuro [3, 4]. Sin embargo, en muchos casos estos comentarios no reflejan con exactitud el código que describen, pueden ser insuficientes, estar desactualizados, o en última instancia no existir en absoluto. En cualquiera de estos escenarios el código escrito resulta difícil de entender, corregir y mantener en el tiempo. Con la intención de mitigar los problemas mencionados, en este trabajo se aborda el desafío de asistir al programador para generar -de manera automática- comentarios claros y relevantes que permitan describir un fragmento de código a su elección.

El enfoque para cumplir con el objetivo propuesto consistió en entrenar un modelo de aprendizaje automático con el cual el desarrollador pueda interactuar, enviando código y recibiendo la descripción del mismo. Para la selección del modelo inicialmente se probó uno basado en redes neuronales recurrentes, pero debido a sus limitaciones se optó por una nueva arquitectura en auge denominada *transformer* [5].

Para entrenar la arquitectura se crearon dos conjuntos de datos, uno con código escrito en PROLOG y el otro con funciones escritas en GO. Los modelos

2 C. Vincenzini et al.

generados fueron luego evaluados utilizando un conjunto de métricas automáticas y se estableció un modelo de referencia.

Los resultados obtenidos en esta investigación revelan que el modelo especializado en el lenguaje GO superó en rendimiento al modelo de referencia. En cuanto al modelo de PROLOG, aunque los resultados fueron inferiores en las métricas establecidas, se introdujo posteriormente una métrica adicional de similitud semántica que demostró que las descripciones generadas mantienen una alta semejanza con las referencias humanas.

## 2 Metodología

El modelo de lenguaje utilizado 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* (60, 220 y 770 millones de parámetros respectivamente). Estos modelos fueron entrenados para diferentes tareas en varios lenguajes de programación 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. Esta estrategia 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 este trabajo tomamos la arquitectura T5 ya entrenada con los datos de CodeTrans y realizamos *fine-tuning* sobre cada uno de ellos, utilizando conjunto de datos propios para realizar diversos experimentos. En particular, tomamos los modelos realizados por transferencia de aprendizaje (TF) y multi-tarea (MT) para los tres tamaños considerados: *small*, *base* y *large*.

Utilizamos dos corpus para desarrollar los experimentos, uno para el lenguaje GO y el otro para PROLOG. Ambos contienen 120 líneas de entre las cuales se seleccionaron de forma aleatoria 20 líneas para formar un conjunto de datos de test, el resto se utilizó para el entrenamiento de los nuevos modelos. Los datos consisten en tuplas que referencian pares de funciones -en algún lenguaje de programación- y el comentario en inglés que describe la funcionalidad de la misma.

## 3 Resultados y discusión

Se llevaron a cabo experimentos para evaluar las tareas de *Transfer Learning* (TF) y de *Multi-Task Learning* con *Fine-Tune* (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: pequeño, base y grande (*small*, *base* y *large*, respectivamente). Para evaluar estas tareas se utilizó la medida BLEU-4 [7]. Los experimentos relacionados al lenguaje GO se

## Modelos de generación de comentarios de código basados en transformers 3

**Table 1.** Resultados de la evaluación de nuestro *dataset test* en todas las tareas para el lenguaje GO. Se utilizó BLEU-4 tanto para evaluar los modelos de CodeTrans[6] como los propios.

	Nuestra salida	Code-Trans
go-tf-s	<b>6,13</b>	3,66
go-tf-base	14,44	<b>15,04</b>
go-tf-large	<b>21,20</b>	17,11
go-mt-tf-s	13,31	<b>18,18</b>
go-mt-tf-base	12,92	<b>15,21</b>
go-mt-tf-large	<b>19,12</b>	14,18

**Table 2.** Resultados de la evaluación de todas las tareas para el lenguaje de programación PROLOG.

MEDIDA	PRO-TF-S	PRO-TF-B	PRO-TF-L
ROUGE-L	22,19	25,61	<b>27,69</b>
BLEU-1	22,73	22,10	27,04
BLEU-2	12,01	12,06	17,52
BLEU-3	6,78	6,19	11,57
BLEU-4	3,32	3,08	<b>7,76</b>
	PRO-MT-TF-S	PRO-MT-TF-B	PRO-MT-TF-L
ROUGE-L	26,79	<b>28,06</b>	24,10
BLEU-1	26,45	24,97	24,51
BLEU-2	15,15	13,54	12,42
BLEU-3	9,69	8,08	6,53
BLEU-4	5,71	4,49	0

realizaron para comparar cómo se comportaba nuestro *fine-tune*, construido con nuestro corpus, con respecto a los modelos de CodeTrans de [6] utilizando nuestro *dataset-test*. La Tabla 1 muestra los resultados. El desarrollo del *Fine-tune* sobre nuestro corpus brindó buenos resultados en los tamaños *large* de ambas tareas.

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. Cabe destacar que CodeTrans no cuenta con este lenguaje. 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. En esta oportunidad se utilizaron dos métricas: el BLEU y el ROUGE [8]. 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.

La Tabla 2 muestra la evaluación realizada para el lenguaje de programación PROLOG en las dos tareas y tamaños (*S*, *B*, *L*). Como se puede apreciar los valores del BLEU-4 son de un dígito en su totalidad, siendo el mejor resultado para la tarea de *TF* de un tamaño *large*.

Las medidas ROUGE-L arrojan valores más interesante. Por un lado señala que el mejor resultado para la tarea *TF* es la de tamaño *large* (coincidente con BLEU), pero también indica que la tarea *MT-TF* de tamaño *base* es el mejor

4 C. Vincenzini et al.

**Table 3.** Ejemplo de una pregunta del test. La referencia humana fue tomada del recurso donde se extrajo el programa.

Programa	<code>reverse([], Z, Z).</code> <code>reverse([H T], Z, Acc) :- reverse(T, Z,</code> <code>[H Acc]).</code>				
Referencia Humana	reverses a list of any length.	BLEU-1	BLEU-2	BLEU-3	BLEU-4
pro-tf-s	reverses the array of elements in the array z.	20,22	0	0	0
pro-tf-b	reverses the order of the elements in the list.	38,46	17,9	0	0
pro-tf-l	reverses an array.	16,67	0	0	0
pro-mt-tf-s	reverses the reverse of the given array.	28,22	0	0	0
pro-mt-tf-b	reverses a list of numbers.	66,67	63,25	58,48	50,81
pro-mt-tf-l	reverses the order of the elements in the array.	20,22	0	0	0

modelo de todos. Estos valores podrían deberse a las características propias de los comentarios para este tipo de lenguajes. Como lenguaje declarativo, podría pensarse que los comentarios 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.

La Tabla 3 muestra la evaluación de una única línea de PROLOG en dos tareas y utilizando distintos tamaños. También se observa su puntuación BLEU-n en relación a su referencia humana y su programa asociado. Como se puede apreciar en dicho ejemplo, las métricas podrían no ser del todo adecuadas ya que indican valores muy bajos para traducciones que podrían ser utilizadas.

## References

1. Spinellis, Diomidis: Reading, Writing, and Code: The key to writing readable code is developing good coding style. Queue 1.7 (2003): 84-89. <https://doi.org/10.1145/957717.957782>
2. Lawrie, Dawn, et al.: What's in a Name? A Study of Identifiers. 14th IEEE international conference on program comprehension (ICPC'06). IEEE, 2006.
3. Forward, Andrew, and Timothy C. Lethbridge: The relevance of software documentation, tools and technologies: a survey. Proceedings of the 2002 ACM symposium on Document engineering. 2002. <https://doi.org/10.1145/585058.585065>
4. Shinyama, Yusuke, Yoshitaka Arahori, and Katsuhiko Gondow: Analyzing code comments to boost program comprehension. 2018 25th Asia-Pacific Software Engineering Conference (APSEC). IEEE, 2018. <https://doi.org/10.1109/APSEC.2018.00047>
5. Vaswani, Ashish, et al.: Attention is all you need. Advances in neural information processing systems 30 (2017).
6. Elnaggar, Ahmed, et al.: 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. Papineni, Kishore, et al.: Bleu: a method for automatic evaluation of machine translation. Proceedings of the 40th annual meeting of the Association for Computational Linguistics. 2002.
8. Lin, Chin-Yew: Rouge: A package for automatic evaluation of summaries. Text summarization branches out. 2004.

# Especificación de Procesos para Preparación de Datos en Sistemas Big Data

Sebastián Yaupe

Tesis de Licenciatura en Sistemas de Información  
sebastian.yaupe@est.fi.uncoma.edu.ar

**Abstract.** Para extraer información de los datos en sistemas big data, los mismos deben atravesar una serie de escenarios con diferentes funciones, como ingestión, filtrado, transformación, etc. Considerando la necesidad de entender estos escenarios, existen antecedentes en el modelado de la preparación de los datos para depósitos de datos, que describen actividades relevantes.

El objetivo general de esta Tesis se enfocó en incorporar modelado de workflows de procesos en actividades para preparación de datos utilizando BPMN; pero a diferencia de trabajos relacionados, se propone extender a particularidades propias de sistemas big data, como actividades en lagos de datos. El modelo resultante fue instanciado en un caso de estudio en el dominio geoespacial (análisis de turbidez en el Canal Principal de Riego del Alto Valle), para mostrar su aplicabilidad y la especialización requerida en la instanciación específica.

**Keywords:** Procesos ETL/ELT · Modelado de Procesos · BPMN

## 1 Motivación

Big data puede definirse como *la capacidad de manejar un gran volumen de datos a una velocidad adecuada al contexto en el que se utilizan*. Dado el valor de las observaciones que se pueden obtener del análisis en sistemas big data (SBDs), los datos se han convertido en el combustible de crecimiento e innovación en las organizaciones; por lo tanto, poseer una arquitectura subyacente que soporte estos requerimientos de crecimiento es de gran importancia. En la práctica, el crecimiento de SBDs está obligando a la comunidad de Ingeniería de Software a replantearse un cambio de paradigma, donde el éxito se asocia especialmente a la correcta identificación y explotación de los datos (y a su correspondiente calidad) con respecto a características de los SBDs.

Considerando la necesidad de modelar funciones o capacidades de etapas específicas del proceso de big data, existen antecedentes en el modelado de la preparación de los datos para depósitos de datos, que describen actividades relevantes. Por ejemplo, para facilitar el modelado y gestión del proceso ETL (Extracción-Transformación-Carga), diversos autores presentan modelos conceptuales y frameworks de modelado.

En la literatura se utilizan varios términos para sistemas que implementan procesos: sistemas de workflows, Business Process Management Suite (BPMS), motor de workflow y Process Engine. La importancia de la gestión de flujos de trabajo (workflows) viene de la posibilidad de representar explícitamente las estructuras de los procesos en los modelos y en la actuación controlada sobre los procesos de acuerdo a esos modelos.

Considerando estas propuestas, este trabajo de tesis se enfocó en incorporar modelado de workflows de procesos en actividades para ETL utilizando BPMN, extendiendo a particularidades propias de sistemas big data, como actividades en ELT (Extracción-Carga-Limpieza).

La Tesis fue dirigida por la Dra. Alejandra Cechich.

## 2 Una breve vista a los capítulos de la Tesis

Para analizar el estado del arte, en el capítulo 2, se presentaron las partes de un proceso big data en forma general, para luego hacer foco en el subproceso ETL. Luego, se agregaron conceptos de lagos de datos, streaming de datos e internet de las cosas, para extender el concepto de ETL a ELT. Luego, se realizó un estudio de las diferentes investigaciones enfocadas en el modelado de procesos ETL. Además de la descripción de las investigaciones, éstas fueron clasificadas según sea el enfoque (conceptual o lógico).

Para poder analizar las actividades ETL se procedió a la investigación de trabajos relacionados y de modelado de escenarios ETL o de diseño de depósito de datos. De ellos, se prestó principal atención al propósito de los modelos propuestos. Algunos se enfocaban en el diseño un modelo conceptual del proceso ETL, otros en el diseño lógico. Otros definían notaciones gráficas que permitían representar las diferentes entidades involucradas en el proceso ETL, con el objetivo de poder definir la mayor cantidad de escenarios posibles. También, se encontraban algunos estudios enfocados en modelar el mapeo de atributos entre esquemas de datos participantes en el proceso ETL. Otro aspecto que se tuvo en cuenta al analizar el estado del arte, fue el lenguaje que utilizaban para plantear el modelo y cómo adaptaban los elementos que el mismo les proveía al contexto de procesos ETL. Dentro de los lenguajes de modelado utilizados se encuentran UML y BPMN.

De los modelos conceptuales, se analizaron los elementos utilizados para representar los escenarios ETL. En general, se encontraron elementos capaces de representar entidades de datos como bases de datos, esquemas, tablas de hecho, dimensiones; y otros elementos capaces de representar funciones sobre distintas entidades de datos. Entonces, de la descripción en los trabajos relacionados del capítulo 2, además de definiciones conceptuales del proceso ETL/ELT introducidas en las secciones 2.1.1 y 2.1.2, se conceptualizaron una serie de características y actividades que sentaron las bases al modelo introducido en el capítulo 3.

En este capítulo, se desarrolló un modelo para la construcción de escenarios de procesamiento de datos para big data. Dicho modelo se expresó en notación BPMN y está compuesto de 4 niveles. El primero es el más abstracto de todos y denota el proceso a modelar. El segundo nivel muestra los procesos contemplados dentro de la ingestión y procesamiento de datos, siendo éstos el proceso ETL y ELT (sección 3.1). El tercer nivel expande los procesos ETL y ELT, mostrando el flujo de datos a lo largo de las etapas, de las cuales se indican los diferentes subprocesos abarcados (sección 3.2). Por último, se especifica el nivel 4, donde se describen los procesos contenidos en la etapa de extracción, limpieza de datos, transformación y carga (sección 3.3). Este nivel describe los eventos que disparan los procesos, las tareas necesarias para llevar a cabo cada uno, los datos participantes, permitiendo de esta manera mostrar el flujo de datos desde su repositorio de origen hasta el destino.

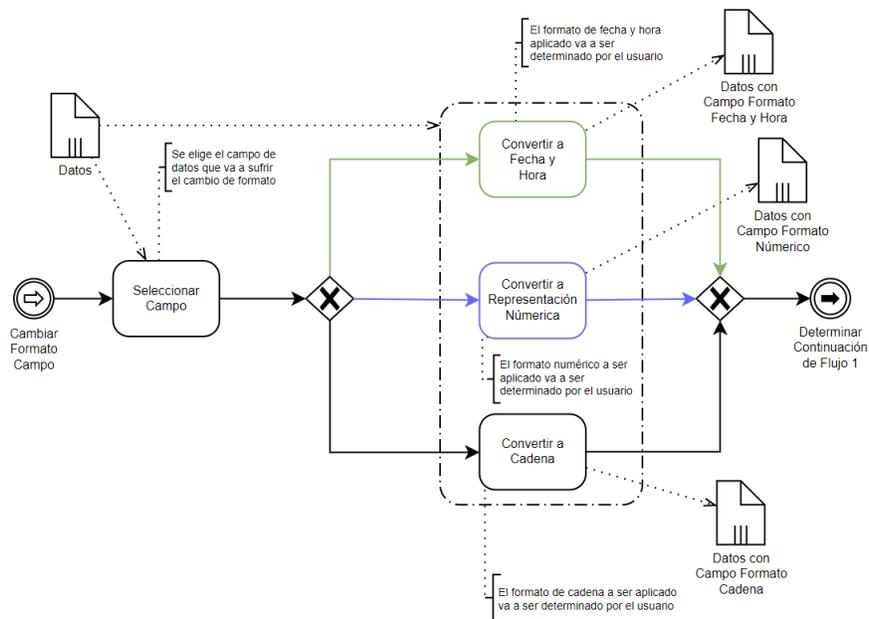


Fig. 1. Modelo Cambiar Formato Campo

Finalmente, en el Capítulo 4 se instanció el modelo propuesto con el caso de estudio *Big Data en el Dominio Hidrológico: Caracterización del Contexto de Predicción de Turbidez* del Licenciado Gastón Vidart, donde se implementa un proceso de Extracción, Carga y Transformación (ETL) para poder analizar la Turbidez del Canal Principal de Riego del Alto Valle (CPRAV). Se describió cómo los modelos de los niveles 2, 3 y 4 se aplican al caso de estudio. En el

caso del nivel 4, en la fase de *Extracción*, se instanció el proceso de ‘extraer asistido por tecnología’ (sección 4.2.1). En la fase de *Limpieza* se instanciaron los procesos ‘Filtrar’, ‘Cambiar Formato Campo’, ‘Modificar Campo’ y ‘Manejar Nulos’ (sección 4.2.2). Por último, en la fase de *Transformación*, se instanciaron los procesos de ‘Generar Resumen’, ‘Mezclar Datos’, ‘Agregar Campo’ y ‘Derivar’ (sección 4.2.3).

### 3 Conclusiones y Trabajos Futuros

Este trabajo resaltó la importancia del procesamiento de los datos para extraer información que sea de valor para la toma de decisiones dentro de una organización. Como también, que el modelado del procesamiento de datos puede ayudar a la hora de identificar los tipos de procesamiento, las entidades de datos involucradas y los requerimientos de negocio al que cada proceso responde. Por último, el hecho de mantener modelado todo el procesamiento de datos, ayuda al mantenimiento del mismo. Hay modelos que poseen un enfoque conceptual, donde se representan las entidades participantes de un proceso; en cambio, otros poseen un enfoque lógico, donde se utiliza un lenguaje formal para detallar el proceso. Al observar trabajos donde se utilizaba el lenguaje UML, se pueden ver las ventajas del lenguaje BPMN a la hora de poder expresar diferentes situaciones en contexto de procesamiento de datos.

Dentro de los trabajos a futuro, se considera la ampliación o adaptación del modelo para que las tareas de limpieza y transformación no sean secuenciales; ya que los modelos, tanto ETL y ELT, están planteados para que los datos siempre sufran una tarea de limpieza antes de ser sometidos a tareas de transformación. Por lo tanto, se podrían considerar casos donde los datos extraídos puedan sufrir tareas de transformación, sin la necesidad de atravesar tareas de limpieza, considerando una serie de condiciones de calidad en los datos.

Otro trabajo futuro se puede encontrar en los procesos de extracción en tiempo real, donde los procesos replican los cambios detectados en los repositorios origen en la copia de datos fuente que se encuentran en un área de almacenamiento intermedio o en el lago de datos. Por lo tanto, para el caso ETL se podría investigar una manera de replicar los cambios directamente en el depósito de datos (DW), con el objetivo de hacer un uso más efectivo del espacio de almacenamiento al no tener que mantener una replica de los datos.

También, se puede considerar para trabajos futuros, modificar o adaptar el modelo de *extracción de streaming data* ya que contiene una tarea que representa los cambios que sufren los datos antes de ser guardados en el repositorio destino. Por lo tanto, se podría mejorar el modelo considerando la reutilización de algunas actividades de limpieza o de transformación definidas en ‘Limpiar’ y ‘Transformar’.

# Modelando Ontologías con Patrones en un Ambiente Gráfico Web de Ingeniería Ontológica

Tomás Quiñonez

Christian Gimenez      Laura Cecchi

Grupo de Investigación en Lenguaje e Inteligencia Artificial  
Facultad de Informática  
Universidad Nacional del Comahue  
tomas.quinonez@est.fi.uncoma.edu.ar  
{christian.gimenez, lcecchi}@fi.uncoma.edu.ar

TESIS DE LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN

**Palabras Clave:** Ingeniería de Software basada en Conocimiento . Patrones de Diseño Ontológico . Ontologías . Lógicas Descriptivas

La Ingeniería Ontológica estudia los métodos y metodologías que orientan a los modeladores en las etapas de diseño, desarrollo, implementación, mantenimiento, uso y publicación de ontologías. Existen diversos enfoques para definir el concepto de ontología. En Ciencias de la Computación, en el contexto de [5], se entiende por ontología a una representación formal del conocimiento, mediante la determinación de conceptos y relaciones entre ellos que le es más fiel al dominio que describe. Otro enfoque dado en [7], define a una ontología como equivalente a una base de conocimiento en Lógica Descriptiva [1], lo que implica que los modeladores deben adquirir un profundo entendimiento de esta categoría de lógicas, lo que añade un nivel de complejidad a las tareas inherentes a la Ingeniería Ontológica.

Actualmente, existen diversas herramientas para el desarrollo de ontologías. Sin embargo, estas herramientas presentan dos falencias claramente identificadas. Por un lado, revisiones de la literatura muestran que escasas herramientas proveen soporte al usuario para modelado de ontologías durante el proceso de su desarrollo[8]. Por otro lado, los ambientes actualmente disponibles presentan un débil soporte para la reusabilidad de ontologías. La Ingeniería Ontológica enfrenta un desafío significativo en lo que respecta a este aspecto, siendo que los dos enfoques más importante son: la característica de poder diseñar componentes que sean portátiles y su interoperabilidad dentro de la Web Semántica. Las ontologías que se reutilizan suelen ser de gran envergadura y extremadamente complejas y de difícil comprensión para un modelador.

Una de las formas de atacar estas falencias y complejidades es incorporar, a herramientas ya existentes, metodologías que a través del uso de patrones, junto con paradigmas de modelado gráfico, podrían ayudar a los modeladores con limitado conocimiento en ontologías y lógica. En este sentido, el uso de patrones de

diseño ontológico (“*Ontology Design Patterns*”, de ahora en más ODPs) [6, 4, 2] se presenta como una posible solución. Los ODP representan soluciones de modelado destinadas a abordar problemas ontológicos recurrentes en el diseño. En esencia, un ODP encapsula un problema de modelado que resurge en múltiples ocasiones a lo largo del desarrollo de ontologías y ofrece una solución establecida y probada para resolver dicho problema en diversos contextos [3]. Así, los ODP son importantes ya que promueven el reuso de soluciones a problemas recurrentes de modelado. Dentro de los ODP, existe un tipo particular de patrones denominados “*Patrones De Diseño Ontológico de Contenido*” (de ahora en más CODP, por sus siglas en inglés). Los CODP pueden ser vistos como pequeñas ontologías bien definidas, usadas como modelos o *templates*, para ser comparadas con respecto a la ontología definida para el dominio en cuestión [4]. Estos patrones se centran en resolver problemas de diseño con respecto a clases y propiedades de un dominio en particular, y por lo tanto resolver problemas de contenido. Esto implica una ayuda al usuario durante la etapa de diseño, al reutilizar estos patrones como guías para determinar si la ontología que está diseñando o reusando, está siendo correctamente desarrollada. Sin embargo, los enfoques basados en CODP requieren un conjunto de patrones aceptados por la comunidad y metodologías adecuadas. Además, las herramientas actuales no brindan un soporte apropiado para el modelado basado en patrones y las estructuras lógicas suelen ser poco amigables.

En este contexto, se diseñó una metodología para el modelado de ontologías basado en patrones, que sugiere CODP durante el proceso de desarrollo de una ontología. Estos patrones se obtienen de un catálogo ampliamente aceptado por la comunidad de expertos. La metodología propuesta para la recomendación automática de patrones se divide en dos etapas: Preprocesamiento de Patrones y Análisis de Ontologías. En la etapa de Preprocesamiento de Patrones, se extrae información relevante de los patrones, incluyendo nombres de elementos y axiomas. También se permite la inclusión de sinónimos para los nombres. Luego, se crean dos grupos de consultas: uno relacionado con nombres y otro con axiomas. Estas consultas se ejecutan en el razonador para detectar coincidencias en la ontología del usuario. En la segunda etapa, el Analizador de Ontologías, se utilizan los resultados de las consultas junto con la ontología del usuario. Primero, se realiza una comparación de nombres para encontrar coincidencias y se registran en una Tabla Diccionario. Luego, la tarea de Comparación de Axiomas compara los axiomas de los patrones con los de la ontología, realizando ajustes según los sinónimos de la Tabla Diccionario. Finalmente, se realiza un postprocesamiento de los resultados, evaluando si un patrón debe sugerirse. Esto se basa en dos criterios: la presencia de clases del patrón en la ontología y la cantidad de nombres de relaciones y consultas afirmativas de los axiomas del patrón en la ontología. Si ambos criterios superan umbrales establecidos, se sugiere el patrón. El resultado de esta etapa es una lista de patrones sugeridos que ocurren total o parcialmente en la ontología del usuario.

Asimismo, se diseñó e implementó un prototipo de una herramienta basada en dicha metodología, denominado ODPR-API. La herramienta tiene como obje-

tivo agilizar el proceso de desarrollo de ontologías al permitir a los modeladores reutilizar y combinar CODP, lo que hace que las ontologías sean más modulares y reutilizables. ODP-APR desempeña un papel fundamental en el ámbito de las herramientas de modelado de ontologías basados en CODP, ya que se ha desarrollado como un aplicativo independiente de cualquier herramienta de modelado de ontologías existente, facilitando su integración con éstas y proveyendo una forma simple para la suscripción de nuevos CODP para ser detectados.

## Referencias Bibliográficas

- [1] F. Baader, I. Horrocks, C. Lutz, and U. Sattler. *An Introduction to Description Logic*. Cambridge University Press, 2017 (cit. on p. 1).
- [2] Eva Blomqvist, Pascal Hitzler, Krzysztof Janowicz, Adila Krisnadhi, Tom Narock, and Monika Solanki. “Considerations regarding Ontology Design Patterns.” In: *Semantic Web 7.1* (2016), pp. 1–7 (cit. on p. 2).
- [3] R de A Falbo, Giancarlo Guizzardi, Aldo Gangemi, and Valentina Presutti. “Ontology patterns: clarifying concepts and terminology”. In: *Proceedings of the 4th Workshop on Ontology and Semantic Web Patterns*. 2013 (cit. on p. 2).
- [4] Aldo Gangemi and Valentina Presutti. “Ontology design patterns”. In: *Handbook on ontologies*. Springer, 2009, pp. 221–243 (cit. on p. 2).
- [5] Nicola Guarino, Daniel Oberle, and Steffen Staab. “What is an ontology?” In: *Handbook on ontologies*. Ed. by Steffen Staab and Rudi Studer. Springer, 2009, pp. 1–17 (cit. on p. 1).
- [6] Pascal Hitzler, Aldo Gangemi, and Krzysztof Janowicz. *Ontology engineering with ontology design patterns: foundations and applications*. Vol. 25. IOS Press, 2016 (cit. on p. 2).
- [7] Ian Horrocks, Oliver Kutz, and Ulrike Sattler. “The Even More Irresistible SROIQ”. In: *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning*. Lake District, UK: AAAI Press, 2006, pp. 57–67. ISBN: 978-1-57735-271-6. URL: <http://dl.acm.org/citation.cfm?id=3029947.3029959> (cit. on p. 1).
- [8] M. Vigo, S. Bail, C. Jay, and R. Stevens. “Overcoming the Pitfalls of Ontology Authoring: Strategies and Implications for Tool Design”. In: *International Journal of Human Computer Studies* (2014) (cit. on p. 1).

# Ingeniería de reglas de asociaciones en Grafos de Conocimiento: Identificación de mejoras para lenguajes visuales

Christian N. Gimenez

Germán A. Braun    Laura A. Cecchi

Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA)  
Facultad de Informática, Universidad Nacional del Comahue  
{christian.gimenez, lcecchi, german.braun}@fi.uncoma.edu.ar

TESIS DE MAGISTER EN CIENCIAS DE LA COMPUTACIÓN  
(EN DESARROLLO)

**Palabras Clave:** Web Semántica . Grafos de conocimiento . Reglas de asociaciones . *Mapping rules* . Lenguajes visuales

Las ontologías y grafos de conocimientos son ampliamente usados para la Web Semántica y la Inteligencia Artificial [5]. Estas formas de representación de conocimiento se utilizan en la medicina<sup>1</sup> [12,11], la genética<sup>2</sup>, la industria del petróleo [18], la biología marina [19], la ciberseguridad [8], etc. Sin embargo, los sistemas informáticos actuales almacenan la información en bases de datos relacionales, en formatos basados en XML o JSON, en archivos de datos de tipo CSV o TSV, entre otros. Para que estos datos puedan ser usados en la Web Semántica, deben utilizarse motores de asociación, que se conectan con estas fuentes y producen grafos de conocimientos con la estructura requerida por los usuarios. Estos motores, en su mayoría, han prevalecido aquellos que también se alimentan de reglas de asociación (*mapping rules*) para definir cómo generar los elementos de los grafos de conocimiento y qué información de la fuente utilizar [4].

La creación de estas reglas es un proceso realizado habitualmente por ingenieros ontológicos que requiere de diversos conocimientos: el dominio, la estructura de la fuente de datos, la información que los diferentes *stackholders* desean consultar, la ontología y el vocabulario destino, etc. Por consiguiente, el diseño de estas reglas debe poder ser fácil de comunicar entre los ingenieros y a las distintas personas involucradas en el proceso y a los futuros usuarios.

Para realizar estos diseños, se puede utilizar un lenguaje visual ya que facilita la comprensión del modelo. En la ingeniería de software, lo habitual es utilizar lenguajes de este tipo, por ejemplo: UML [15], BPMN [14] y MER [3]. Si bien,

<sup>1</sup> Véase <https://www.snomed.org> y <https://bioportal.bioontology.org/ontologies/SNOMEDCT> para más información. Visitado el día 28 de septiembre del 2023.

<sup>2</sup> Véase <https://geneontology.org>. Visitado el día 28 de septiembre del 2023.

existen lenguajes visuales para representar reglas de asociación, estos no son estándares y son implementados por herramientas donde la sintaxis y semántica dependen fuertemente de la utilidad de las mismas. Cabe mencionar que, los lenguajes de modelado de la ingeniería de software, no son utilizados por estos lenguajes visuales, por ejemplo, para representar la fuente de datos.

Por ello, se propone una metodología para identificar mejoras sobre un lenguaje visual para la ingeniería de asociaciones. Se hace énfasis en ampliar la efectividad cognitiva y la expresividad de los mismos para comunicar las reglas de forma visual.

Esta metodología comienza con la descripción del lenguaje visual, describiendo su sintaxis visual y la semántica asociada. Para ello, se requiere relevar y analizar artículos que describan el lenguaje o la herramienta que lo utiliza y sus casos de estudio. Además, se utilizan ejemplos para interpretar la asociación entre cada sentencia del diagrama y su representación en las reglas de asociación.

Luego, se estudia la descripción realizada con respecto a los principios de *Physics of Notations* (PoN) planteados en [13]. Estos, describen criterios que persiguen ampliar la expresividad del lenguaje, reducir ambigüedades, gestionar la complejidad de los diagramas de gran tamaño y proponer dialectos para diferentes audiencias (p. ej.: personas con o sin experiencia en el uso del lenguaje) y medios (p. ej.: digital, papel, impresión, etc.).

Así, una vez analizado el lenguaje, se sugieren posibles modificaciones para que éste cumpla en mayor medida con los principios de PoN. Asimismo, se plantea la forma en que cada opción posible a modificar afecte, no solo al criterio que se hace foco, sino también a los otros, proveyendo recomendaciones para evitar impactos negativos.

En conclusión, se provee un procedimiento que busca describir a un lenguaje, analizarlo según los principios de PoN y, en base a este resultado, proponer mejoras y consideraciones para ampliar la expresividad y la comprensión de los diagramas producidos con el lenguaje. Considerando también el objetivo particular del lenguaje mismo, y a su vez la audiencia a la cual se desea presentar los diagramas.

Para comprender el funcionamiento del procedimiento propuesto se aplica esta metodología a lenguajes visuales existentes, particularmente Map-on [16], SQuaRE [1,2], Juma [10,9], KR2RML/Karma [17] y MapVOWL/RMLEditor [6].

Se espera que la utilización de esta metodología permita la creación de lenguajes más comprensibles y que facilite la comunicación de las reglas de asociación entre los ingenieros y los diferentes *stackholders*. Asimismo, esta investigación promueve mejoras en herramientas gráficas como crowd<sup>3</sup> para brindar soporte *Ontology-Based Data Access* (OBDA) con el fin de popular esquemas ontológicos con instancias por medio del uso de diseños visuales de reglas de asociación fáciles de entender.

---

<sup>3</sup> <https://crowd.fi.uncoma.edu.ar> visitado el día 30 de octubre del 2023.

## Referencias

1. Bak, J., Blinkiewicz, M.: Square: A visual tool for creating R2RML mappings. In: Kawamura, T., Paulheim, H. (eds.) Proceedings of the ISWC 2016 Posters & Demonstrations Track co-located with 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan, October 19, 2016. CEUR Workshop Proceedings, vol. 1690. CEUR-WS.org (2016), <http://ceur-ws.org/Vol-1690/paper62.pdf>
2. Bak, J., Blinkiewicz, M., Lawrynowicz, A.: User-friendly visual creation of R2RML mappings in square. In: Ivanova et al. [7], pp. 139–150, <http://ceur-ws.org/Vol-1947/paper13.pdf>
3. Chen, P.: Entity-Relationship Modeling: Historical Events, Future Trends, and Lessons Learned, pp. 296–310. Springer (2002), [http://bit.csc.lsu.edu/~chen/pdf/Chen\\_Pioneers.pdf](http://bit.csc.lsu.edu/~chen/pdf/Chen_Pioneers.pdf)
4. Dimou, A.: R2RML and RML comparison for RDF generation, their rules validation and inconsistency resolution. CoRR **abs/2005.06293** (2020), <https://arxiv.org/abs/2005.06293>
5. Harmelen, F.V., Lifschitz, V., Porter, B. (eds.): Handbook of Knowledge Representation. Elsevier (2008)
6. Heyvaert, P., Dimou, A., De Meester, B., Seymoens, T., Herregodts, A.L., Verborgh, R., Schuurman, D., Mannens, E.: Specification and implementation of mapping rule visualization and editing: Mapvowl and the rmleditor. Journal of Web Semantics **49**, 31 – 50 (2018). <https://doi.org/https://doi.org/10.1016/j.websem.2017.12.003>, <http://www.sciencedirect.com/science/article/pii/S1570826817300707>
7. Ivanova, V., Lambrix, P., Lohmann, S., Pesquita, C. (eds.): Proceedings of the Third International Workshop on Visualization and Interaction for Ontologies and Linked Data co-located with the 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 22, 2017, CEUR Workshop Proceedings, vol. 1947. CEUR-WS.org (2017), <http://ceur-ws.org/Vol-1947>
8. Jia, Y., Qi, Y., Shang, H., Jiang, R., Li, A.: A practical approach to constructing a knowledge graph for cybersecurity. Engineering **4** (02 2018). <https://doi.org/10.1016/j.eng.2018.01.004>
9. Junior, A.C., Debruyne, C., O’Sullivan, D.: Juma: An editor that uses a block metaphor to facilitate the creation and editing of R2RML mappings. In: Blomqvist, E., Hose, K., Paulheim, H., Lawrynowicz, A., Ciravegna, F., Hartig, O. (eds.) The Semantic Web: ESWC 2017 Satellite Events - ESWC 2017 Satellite Events, Portorož, Slovenia, May 28 - June 1, 2017, Revised Selected qPapers. Lecture Notes in Computer Science, vol. 10577, pp. 87–92. Springer (2017). [https://doi.org/10.1007/978-3-319-70407-4\\_17](https://doi.org/10.1007/978-3-319-70407-4_17), [https://doi.org/10.1007/978-3-319-70407-4\\_17](https://doi.org/10.1007/978-3-319-70407-4_17)
10. Junior, A.C., Debruyne, C., O’Sullivan, D.: Using a block metaphor for representing R2RML mappings. In: Ivanova et al. [7], pp. 1–12, <http://ceur-ws.org/Vol-1947/paper01.pdf>
11. Liu, Y., Hildebrandt, M., Joblin, M., Ringsquandl, M., Raissouni, R., Tresp, V.: Neural multi-hop reasoning with logical rules on biomedical knowledge graphs. In: Verborgh, R., Hose, K., Paulheim, H., Champin, P.A., Maleshkova, M., Corcho, O., Ristoski, P., Alam, M. (eds.) The Semantic Web. pp. 375–391. Springer International Publishing, Cham (2021)
12. Mann, M., Ilievski, F., Rostami, M., Aastha, Shbita, B.: Open drug knowledge graph. In: Chaves-Fraga, D., Dimou, A., Heyvaert, P., Priyatna, F., Sequeda, J.

- (eds.) Proceedings of the 2nd International Workshop on Knowledge Graph Construction co-located with 18th Extended Semantic Web Conference (ESWC 2021) (2021), <https://ceur-ws.org/Vol-2873/>
13. Moody, D.: The "physics" of notations: Toward a scientific basis for constructing visual notations in software engineering. *IEEE Trans. Softw. Eng.* **35**(6), 756–779 (Nov 2009). <https://doi.org/10.1109/TSE.2009.67>, <https://doi.org/10.1109/TSE.2009.67>
  14. Object Management Group: Business process model and notation (BPMN). Tech. rep., Object Management Group (2011), <http://www.omg.org/spec/BPMN/2.0/>, versión 2.0
  15. Object Management Group: OMG unified modeling language (OMG UML). Tech. rep., Object Management Group (2017), versión 2.5.1
  16. Sicilia, Á., Nemirovski, G., Nolle, A.: Map-on: A web-based editor for visual ontology mapping. *Semantic Web* **8**(6), 969–980 (2017). <https://doi.org/10.3233/SW-160246>, <https://doi.org/10.3233/SW-160246>
  17. Slepicka, J., Yin, C., Szekely, P.A., Knoblock, C.A.: KR2RML: an alternative interpretation of R2RML for heterogenous sources. In: Hartig, O., Sequeda, J.F., Hogan, A. (eds.) Proceedings of the 6th International Workshop on Consuming Linked Data co-located with 14th International Semantic Web Conference (ISWC 2105), Bethlehem, Pennsylvania, USA, October 12th, 2015. CEUR Workshop Proceedings, vol. 1426. CEUR-WS.org (2015), <http://ceur-ws.org/Vol-1426/paper-08.pdf>
  18. Soyulu, A., Kharlamov, E., Zheleznyakov, D., Jiménez-Ruiz, E., Giese, M., Skjæveland, M.G., Hovland, D., Schlatte, R., Brandt, S., Lie, H., Horrocks, I.: Optiquevqs: A visual query system over ontologies for industry. *Semantic Web* **9**(5), 627–660 (2018). <https://doi.org/10.3233/SW-180293>, <https://doi.org/10.3233/SW-180293>
  19. Zárate, M.D., Braun, G.A., Fillottrani, P.R., Delrieux, C., Lewis, M.: *BiGe-Onto*: An ontology-based system for managing biodiversity and biogeography data. *Appl. Ontology* **15**(4), 411–437 (2020). <https://doi.org/10.3233/AO-200228>, <https://doi.org/10.3233/AO-200228>

ACTAS DE LAS  
**JORNADAS DE LA FACULTAD DE INFORMÁTICA**



NOV 2023  
NEUQUÉN

FACULTAD DE INFORMÁTICA  
UNIVERSIDAD NACIONAL DEL COMAHUE

ISBN 978-987-604-653-4

