

# LIBRO DE ACTAS

---

XXVI WORKSHOP DE INVESTIGADORES EN  
CIENCIAS DE LA COMPUTACIÓN

# WICC2024

Puerto Madryn, abril 2024  
UNPSJB - Red de Universidades  
de Carreras de Informática RedUNCI



Libro de Actas XXVI Workshop de Investigadores en Ciencias de la Computación/  
Compilación de Marcos Daniel Zárate; Francisco Ezequiel Paez; Romina Stickar. -  
1a ed. - Comodoro Rivadavia: Universitaria de la Patagonia -EDUPA, 2024.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-8352-57-2

1. Computación I. Zárate, Marcos Daniel, comp. II. Paez, Francisco Ezequiel, comp.  
III. Stickar, Romina, comp. IV. Título.  
CDD 600

ISBN 978-987-8352-57-2



# Autoridades

## **Universidad Nacional de la Patagonia San Juan Bosco**

### **Rectora**

ANTONIA LIDIA BLANCO

## **Facultad de Ingeniería**

### **Decana**

MARÍA ELIZABETH FLORES

## **Autoridades RedUNCI**

### **Coordinador Titular**

PESADO, PATRICIA (UNLP)

### **Coordinador Alterno**

ETEROVIC, JORGE (UNLAM)

# BDMD

---

XXVI WORKSHOP DE INVESTIGADORES EN  
CIENCIAS DE LA COMPUTACIÓN



**BASES DE DATOS Y MINERÍA DE DATOS**

# **Análisis de datos públicos abiertos e identificación de patrones para predicciones: Proyecto INCUCAI para el estudio de disponibilidad de Órganos y Tejidos en Argentina**

Roxana Martínez, Carlos Bateca, Hernán Zabala, Eugenio Clavijo, Pablo Vilaboa, Juan Iannini, German Colato, Matías Navone, Sebastián Tsiro, Nicolás Rizzo

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

{roxana.martinez, pablo.vilaboa, juan.garciaianini, matias.navone}@uai.edu.ar

{CarlosAlberto.BatecaCaicedo, HernanGuillermo.Zabala, Eugenio.Clavijo,  
GermanDario.Colato, SebastianMartinDemetrio.Tsiros,  
nicolasleonardo.rizzo}@alumnos.uai.edu.ar

## **RESUMEN**

En la actualidad, si bien, hay estudios y guías de buenas prácticas sobre cómo disponibilizar los datos públicos, hoy por hoy, no existen demasiados trabajos de investigación que se centren en el contenido de los conjuntos de datos (datasets) que se encuentran en sitios web gubernamentales, lo que hace que el análisis de la calidad del contenido de estos datasets quede en un segundo plano. Este proyecto, se centra en el análisis de contenido, validación, tratamiento de herramientas de verificación, cotejo en las estructuras de datos, y propuestas de guías de buenas prácticas. A través de este estudio, se apoya para lograr gobiernos más transparentes en cuestiones de rendición de cuentas, como así también, brindar acceso a la información pública, que permite llevar a cabo un mejor desarrollo de políticas públicas para las personas, pero para ello es necesario partir de una base de datos que sea legible, limpia y validada. Este trabajo, no solo trabaja en prototipos de software en validaciones de datos específicos, sino que, además, permite estudiar modelos predictivos (aprendizaje automático) que pueden brindar un mejor análisis para la toma de decisiones.

**Palabras clave:** Datos Abiertos, Gobierno Abierto, Métricas de Calidad de Datos,

Machine Learning, Predicciones a partir de Patrones.

## **CONTEXTO**

El presente trabajo es la última parte del proyecto denominado “Investigación y desarrollo de software para la validación de la calidad de datos abiertos e identificación de patrones para predicciones”, que tuvo inicio en el mes de marzo 2022 y con fecha de culminación en el mes de marzo 2025. Este proyecto pertenece a la línea de investigación de Ingeniería de Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI), el cual contribuye al desarrollo de las Tecnologías de la Información y las Comunicaciones (TIC) en Argentina y en el mundo, llevando adelante la investigación en diversas áreas a través de tratamiento y manipulación de datos en diversas temáticas, tales como: aspectos gubernamentales, gobierno de datos, economía, salud, tecnología, entre otras. El proyecto es financiado y evaluado por la Secretaría de Investigación de la Universidad, tiene una duración de 3 años, y cuenta con la participación de docentes y estudiantes de grado y posgrado en diversas carreras de la Facultad de Tecnología Informática.

## 1. INTRODUCCIÓN

Si bien Gobierno Abierto implica varios enfoques, este trabajo apunta en la implementación de nuevas técnicas de gestión de datos y en el estudio de la disponibilización de datos públicos a través de un diseño tecnológico, que se encuentra centrado en el ciudadano promedio para fomentar la inclusión social.

Este proyecto conlleva varias temáticas con enfoques sobre propuestas, pero en esta oportunidad, se presenta el estudio de la disponibilidad de datos sobre Órganos y Tejidos en la República Argentina. Este análisis se realiza a través del sitio web del Instituto Nacional Central Único Coordinador de Ablación e Implante, INCUCAI [1], y como eje central se espera analizar los datos históricos extraídos del sitio web público del INCUCAI, que permitan resaltar y descubrir tendencias que aporten información y concientización a la población, sobre este contexto y necesidad frente a las largas listas de espera de donantes.

Con este estudio se espera que se identifiquen algunos factores de ayuda a las personas con base a determinadas situaciones ciudadanas de las distintas provincias de la República Argentina, mediante los datos disponibilizados del sistema de SINTRA [2] que es el Sistema Nacional de Información de Procuración y Trasplante de la República Argentina.

Cabe mencionar que “ampliar el acceso a la información pública se fortalece la rendición de cuentas y se enriquece el debate público, a la vez que se crean nuevas oportunidades para generar valor agregado” [3]. En cuestiones de salud pública, es importante comprender que es una de las necesidades básicas más relevantes del ser humano, por lo que la búsqueda de procedimientos que mejoren la calidad y tiempo de vida, resultan sumamente necesarios en avances para la medicina moderna, en donde la sustitución de un órgano enfermo por otro sano, constituye probablemente, el tratamiento más espectacular de la medicina [4], en este

contexto la medicina avanzó y avanza permanentemente con mayor tecnología, logrando una sobrevida de los pacientes mediante el estudio de trasplantes de órganos como avance para la sociedad [5].

El estudio de datos públicos en conjunto con las nuevas tecnologías, permitirán proponer mejoras en las formas en que se lleva a cabo los distintos procedimientos de trasplante de órganos en cuanto a la calidad y seguridad de estos. Como problemática, se analizó que la disponibilidad de órganos no va de la mano y parece ser independiente de cualquier mejora tecnológica actual, ya que existe una gran desproporción entre el elevado número de candidatos que requieren un trasplante y la relativamente escasa cantidad de donadores disponibles [6]. Esto provoca largas listas de espera en donde el potencial donante debe ser evaluado previamente, buscando el mayor parecido genético entre el donante y el receptor [7], para esto la comunidad hospitalaria tiene un rol clave en la detección de potenciales donantes [8]. Por lo explicado anteriormente, es preciso conocer que “la cultura de datos abiertos trata de obtener un valor añadido de la información [...], la información no genera sólo su valor por estar reservada a unos pocos, sino que lo hace por su disponibilidad para ser interpretada y traducida por cualquier actor interesado en trabajar con ella” [9], es decir, un dato abierto puede ser accedido a través de un formato especial que permite la interoperabilidad con otros softwares, y así lograr conocer que tipos de aspectos pueden ser mejorados en la sociedad, por ejemplo, si son necesarios determinados cuidados y/o servicios en provincias específicas debido a la falta de recursos. Todo esto se podría lograr con un estudio de los datos para conocer que falencias y que mejoras son las que se podrían trabajar para lograr un valor agregado a la comunidad.

Existen algunos trabajos enfocados a las mediciones en aspectos de calidad de datos públicos [10] [11] [12], por otro lado, otras investigaciones apuntan a medidas otorgadas a la calidad de los datos, o bien trabajos realizados con informes sobre el estado de los

datos elaborados por fundaciones comprometidas con el impacto de estos datos públicos [13].

Como parte de este trabajo, se analizaron las expresiones de voluntad registradas por el INCUCAI, con el fin de identificar los posibles motivos que expliquen la falta de interés por parte de la población, así como la importancia de esta información para mejorar la tasa de detección de potenciales donantes en menor tiempo, ya que no actuar con celeridad puede provocar la pérdida de órganos. Básicamente, en casos estudiados a través de estos datos, se identificó que disponer de la información del donante con previa antelación (tramites de expresión de voluntad), podría optimizar el proceso de donación, aportando una ayuda a los profesionales de la salud. Además, al aumentar la participación de las personas y por ende la cantidad de datos, se podría pensar en modelos predictivos que fortalezcan la actual procuración de órganos. Algunos de los ejemplos estudiados en este proyecto, son parte de o bien se encuentran en desarrollo, por ejemplo, a continuación, se analiza la evolución de los donantes reales versus los donantes de los que efectivamente se utilizó alguno de sus órganos. En la Figura 1, se muestra la tendencia de donantes Efectivos vs Reales en Argentina, período [2000-2023]. Se visualiza la línea que indica la cantidad de donantes Reales por año, es decir, órganos extraídos y aún no trasplantados. Por otro lado, las barras indican la cantidad de donantes efectivos por año, órganos extraídos y trasplantados, siempre será menor o igual al indicador de línea (Reales). En el período de los años 2000 al 2007, se observa un crecimiento en las donaciones realizadas, casi duplicando su valor en 7 años. Finalmente, se diferencian los donantes reales (línea azul claro) de los donantes efectivos (Barras azul oscuro).

Si bien aún falta un largo camino, en este trabajo se pudo analizar que la oferta de órganos es escasa principalmente por la rápida caducidad que tienen estos componentes orgánicos, y las condiciones especiales para ser considerados como “Reales”, es decir

potencialmente útiles, los cuales, en muchos casos, se pierden en el camino y no llegan a ser “Efectivos”, es decir, trasplantados en una persona. Esto se debe a múltiples factores, ya que es un proceso complejo con muchas variables de por medio, pero que se podría mejorar considerablemente si se contara con la misma información que se tiene del receptor. Lograr identificar más información en aspectos a futuros, agilizaría los tramites médicos y potenciaría mucho más la procuración actual de órganos, en conjunto con la utilización de las herramientas adecuadas para el procesamiento de datos.



**Figura 1.** Tendencia de donantes Efectivos vs Reales en Argentina, período [2000-2023].

Otro punto para trabajar en este proyecto es la utilización del aprendizaje automatizado (Machine Learning), el cual permitiría utilizar datos para entrenar a diferentes modelos, que luego, éstos a través de algoritmos, colaborarían en la toma de decisiones en el estudio de escenarios de análisis. Para ello, es fundamental comprender que es necesario contar con datos de calidad en los portales de datos abiertos que serán utilizados.

## 2. LÍNEAS DE INVESTIGACIÓN DESARROLLO

Este proyecto pertenece a la línea de investigación de Ingeniería de Software del Centro de Altos Estudios en Tecnología Informática (CAETI). Para esta última instancia del proyecto, se hace enfoque en los siguientes ejes:

- Culminar con el diseño y desarrollo de algoritmos para identificar patrones en datos públicos abiertos orientados en temáticas de salud, políticas de buenas prácticas en calidad de datos, tecnologías, y aspectos internacionales que puedan llevar a cabo un valor agregado a la comunidad.
- Analizar los modelos predictivos orientados a contextos gubernamentales, como así también, analizar posibles predicciones sobre nuevos datos encontrados con técnicas de machine learning.
- Establecer una comparativa entre las herramientas de tecnología y análisis de datos en concepto de implementación de técnicas de Machine Learning con análisis predictivo.
- Implementar nuevas funciones en una herramienta de desarrollo propio de validación de datasets gubernamentales basada en métricas de calidad de datos estándares como geoespaciales;
- Análisis de casos predictivos y aplicación con casos reales y prototipos desarrollados.
- Generar una guía de buenas prácticas para las técnicas de Machine Learning en la utilización de datos abiertos.

## 3. RESULTADOS OBTENIDOS/ESPERADOS

### *Enfoque del proyecto:*

Actualmente, el contexto de Gobierno Abierto permite la implementación de diversas herramientas tecnológicas en las que se logran trabajar con políticas públicas en aspectos de salud, infraestructura básica, seguridad ciudadana, entre otros. Este proyecto se

enfoca en la mejora de falencias detectadas sobre el análisis de los datos públicos que se encuentran disponibilizados en sitios gubernamentales. Esto se logra, a través de la utilización de prototipos de software en los que se orientan a validaciones sobre aspectos de calidad de datos. Un adecuado análisis, permitirá evitar errores en el estudio de éstos, y mitigar posibles interpretaciones erróneas. Mantener la calidad de datos, permitirá trabajar con algoritmos en los que se pueda detectar patrones que fomenten el estudio de predicciones en escenarios posibles, y que esto traerá aparejado, beneficios sociales a las personas.

### *Objetivos principales:*

Este trabajo propone un análisis sobre los alcances, desafíos y mejores prácticas a partir de casos prácticos y datos reales con los que se trabajan en Gobierno Abierto. El objetivo principal es realizar el análisis, diseño y desarrollo de herramientas de software para la gestión y validación de la calidad de los datos públicos en dicho contexto. Detectando el “estado de salud” de las fuentes de datos provenientes de casos de aplicación gubernamentales con los prototipos desarrollados, incorporando algoritmos para identificar patrones que logren predicciones sobre nuevos datos públicos de diversos tipos (geolocalizados, entre otros).

### *Metodología y Técnicas:*

El proceso metodológico de investigación utilizado para esta última parte del proyecto, es a través de un proceso sistemático cualitativo, que implica con la culminación de las siguientes últimas etapas: revisión de datos, clasificación y análisis de datos; estudio de escenarios de identificación de patrones sobre éstos con enfoques predictivos sobre nuevos datos; desarrollo de prototipos (análisis cuantitativo de datos); visualización de la predicción de patrones; validación de la solución propuesta; identificación de las limitaciones del trabajo.

### Resultados obtenidos:

En lo que fue el inicio del proyecto, a grandes rasgos, se analizaron las falencias en aspectos de calidad de datos públicos abiertos, se lograron identificar técnicas de validación de calidad de datos abiertos en formatos estándares y, además, se realizó un estudio detallado sobre los datos públicos geolocalizados. Si bien, este es un proyecto que abarca varias temáticas de datos (tecnología, salud, financiamiento, entre otros), en esta oportunidad, se está trabajando sobre datos en un entorno de salud pública, orientados a el estudio de disponibilidad de Órganos y Tejidos en Argentina, analizando el sitio web de INCUCAI. Por otra parte, se están estudiando posibles modelos predictivos (mediante herramientas de software), para efectuar predicciones y mejorar los resultados de análisis en resultados finales con técnicas de Machine Learning.

### 4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto se compone por una docente investigadora, Doctora en Ciencias Informáticas de la Universidad Nacional de La Plata (UNLP), directora del Proyecto de Investigación, 3 (tres) docentes tesistas, 1 (un) docente auxiliar de la carrera Ingeniería en Sistemas Informáticos y 2 (dos) estudiantes tesistas de la carrera de Maestría en Tecnología Informática en la Universidad Abierta Interamericana (UAI). También cuenta con un docente tesista de la carrera del Doctorado en Ciencias Informáticas de UAI, 2 (dos) estudiantes colaboradores de la carrera de Ingeniería en Sistemas Informáticos de UAI, y 2 (dos) tesistas de la carrera Licenciatura en Gestión de Tecnología Informática en UAI. En relación directa con la línea de I/D presentada para el proyecto, los miembros del equipo se encuentran en realización de: 1 tesis doctoral, 4 tesis de maestría y 2 tesinas de grado de la UAI.

### 5. BIBLIOGRAFÍA

- [1] Argentina.gob.ar - INCUCAI (2024). “INCUCAI”. Disponible en: <https://www.argentina.gob.ar/salud/incuca>
- [2] INCUCAI - SINTRA (2024). “SINTRA - El Sistema Nacional de Información de Procuración y Trasplante de la República Argentina”. Disponible en: <https://sintra.incucan.gov.ar/>
- [3] Buenos Aires provincia (2017). “Kit de Apertura Municipal”. Disponible en: <http://escueladefiscales.com/Kit%20de%20Apertura%20Municipal%202017%20-%20provincia%20de%20buenos%20aires.pdf>
- [4] Guerra-Sáenz, E. K., Narváez-Navarro, Á. G., Hernández-López, A. C., Saucedo, J. B., Ruiz-Cantú, G., Cordero-Pérez, P., ... & Escobedo-Villarreal, M. M. (2016). Actitud hacia la donación y el trasplante de órganos y tejidos de tres instituciones de salud. Gaceta médica de México, 152(5), 688-696.
- [5] Cyunel, M., Garea, M., & Neira, P. (2019). Procuración y trasplante de órganos y tejidos. Rev Hosp Niños (B Aires), 61(273), 88-94.
- [6] Arias, T. T. M., Weber, F. L. R., de Pablo, P. A. S., Patiño, M. A. G., Balbuena, M. M. M., & Reyes, P. A. S. (2010). Encuesta de opinión sobre donación de órganos 2009. Medicina Interna de México, 26(4), 313-318.
- [7] Defelitto, J. (2011). Trasplante de órganos: generalidades. PARTE III, 540. Disponible en: <https://acortar.link/wfPyQB>
- [8] Cyunel, M., Garea, M., & Neira, P. (2019). Procuración y trasplante de órganos y tejidos. Rev Hosp Niños (B Aires), 61(273), 88-94.
- [9] Magallón Rosa, R. (2017). Datos abiertos y acceso a la información pública en la reconstrucción de la historia digital.
- [10] ISO 25012 (2008). Ingeniería de software - Requisitos de calidad y evaluación de productos de software (SQuARE) - Modelo de calidad de datos.
- [11] Martínez, R. et al. (2021). Metrics proposal to measure the quality of governmental datasets. IEEE Latin America Transactions, Vol. 100. ISSN 1548-0992.
- [12] de España, G. (2017). Manual práctico para mejorar la calidad de los datos abiertos. Madrid.
- [13] ODI Open Data Institute. “The 2019 Data Skills Framework”. Disponible en: <https://theodi.org/article/open-data-skills-framework/>