REVISTA DEL DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN DE LA UNIVERSIDAD DE CHILE

# Bits DE CIENCIA

EDICIÓN N°22 | PRIMER SEMESTRE 2022

fcfm Ciencias de la Computación
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

## REGULACIÓN DIGITAL Y PLATAFORMAS PARA PROCESOS PARTICIPATIVOS

**La trivialidad de los neuroderechos**
/ Alejandra Zúñiga, Luis Villavicencio, Danielle Zaror y Ricardo Salas

**Interpretabilidad de modelos de inteligencia artificial con garantías formales**
/ Bernardo Subercaseaux

**Alfred V. Aho y Jeffrey D. Ullman, ACM Turing Award 2020**
/ Gonzalo Navarro

# Contenidos


1 **Editorial**
/ Federico Olmedo

2 **Interpretabilidad de modelos de inteligencia artificial con garantías formales**
/ Bernardo Subercaseaux

9 **Alfred V. Aho y Jeffrey D. Ullman, ACM Turing Award 2020**
/ Gonzalo Navarro

17 **Banderas rojas: El proyecto de ley sobre plataformas digitales, sus múltiples errores y las lecciones que podemos aprender de ellos**
/ Vladimir Garay

24 **La trivialidad de los neuroderechos**
/ Alejandra Zúñiga Fajuri, Luis Villavicencio Miranda, Danielle Zaror Miralles y Ricardo Salas Venegas

30 **Votación electrónica remota para la Universidad de Chile**
/ Alejandro Hevia, Camilo Gómez, Catalina Burgos Kreither y Marta Apablaza

36 **Participación digital y plataformas en el contexto del proceso constituyente**
/ Ariel Barraza

43 **Técnicas formales de privacidad de datos: ¿Está el Servel protegiendo nuestra privacidad?**
/ Matías Toro

49 **Estudiantes DCC**



**COMITÉ EDITORIAL**
María Cecilia Bastarrica
Claudio Gutiérrez
Alejandro Hevia
Ana Gabriela Martínez
Bárbara Poblete
Jocelyn Simmonds
Iván Sipirán

**EDITOR GENERAL**
Federico Olmedo

**EDITORA PERIODÍSTICA**
Ana Gabriela Martínez

**PERIODISTA**
Karin Riquelme

**DISEÑO**
Paulette Filla

**FOTOGRAFÍAS E IMÁGENES**
Comunicaciones DCC



Revista BITS de Ciencia del Departamento de Ciencias de la Computación de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile se encuentra bajo Licencia Creative Commons Atribución-NoComercial-CompartirIgual 3.0 Chile. Basada en una obra en www.dcc.uchile.cl



**Revista Bits de Ciencia N°22**
ISSN 0718-8005 (versión impresa)
**www.dcc.uchile.cl/revista**
ISSN 0717-8013 (versión en línea)



**Departamento de Ciencias de la Computación**
*Avda. Beauchef 851, 3º piso, edificio norte. Santiago, Chile. 837-0459 Santiago*

www.dcc.uchile.cl
56 22 9780652
bitsdeciencia@dcc.uchile.cl
/dccuchile



*El contenido de los artículos publicados en esta Revista, son de exclusiva responsabilidad de sus autores y no reflejan necesariamente el pensamiento del Departamento de Ciencias de la Computación de la Universidad de Chile.*

# Editorial

**FEDERICO OLMEDO**
*Editor General*
*Revista Bits de Ciencia*

Las tecnologías digitales nos han abierto la puerta a un sinfín de tareas antes inimaginables, desde permitir una interconexión global instantánea a través de redes sociales hasta, en el contexto del proceso constituyente que se está llevando a cabo en Chile, permitir a cualquier persona presentar propuestas de normas constitucionales a través de la plataforma de Iniciativa Popular de Norma.

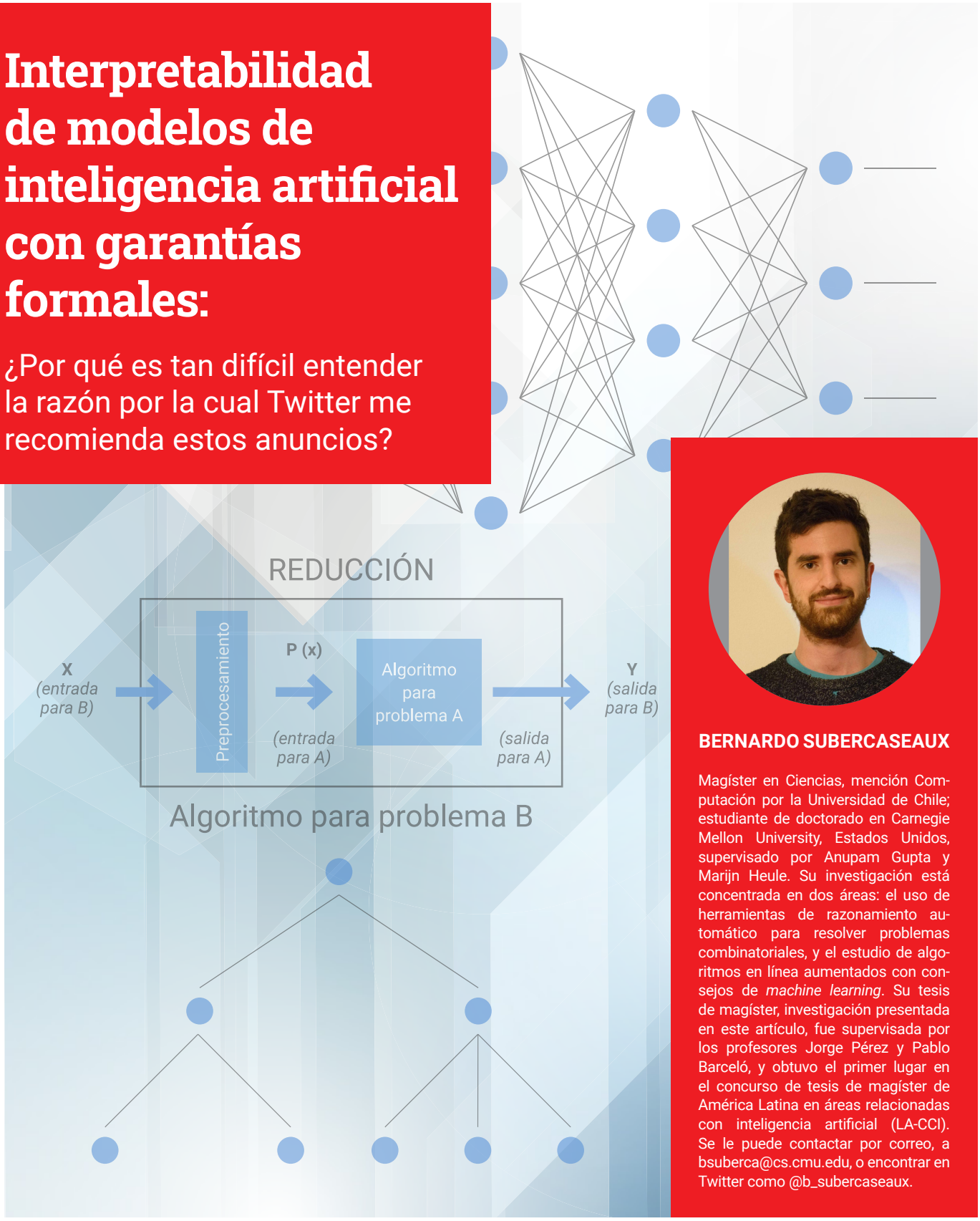
En esta edición de la Revista abordamos, como tema central, el apoyo a distintos procesos participativos que estas tecnologías digitales pueden brindarnos, y los desafíos regulatorios que generan sus nuevos alcances. Por un lado, discutimos dos propuestas legislativas recientes sobre la regulación de plataformas digitales y establecimiento de neuroderechos y, por otro lado, el uso de herramientas digitales para la participación política en dos contextos diferentes: el del proceso constituyente y el de votación electrónica dentro de la Universidad de Chile. Analizamos también el problema de la privacidad de datos a través de un estudio de caso de publicación de datos por parte del Servel.

En la sección de Investigación Destacada estudiamos un problema fundamental de la inteligencia artificial: cómo comparar formalmente la interpretabilidad de diferentes modelos. En la sección de Premio Turing presentamos la premiación que recientemente recibieron Aho y Ullman, y finalmente en la sección de Estudiantes recapitulamos los trabajos de varixs estudiantes recientemente graduadxs del DCC.

Esperemos que disfruten de esta edición de la Revista. Cualquier consulta o sugerencia, pueden hacérnosla llegar a bitsdeciencia@dcc.uchile.cl. ▪

# Interpretabilidad de modelos de inteligencia artificial con garantías formales:

## ¿Por qué es tan difícil entender la razón por la cual Twitter me recomienda estos anuncios?






**BERNARDO SUBERCASEAUX**

Magíster en Ciencias, mención Computación por la Universidad de Chile; estudiante de doctorado en Carnegie Mellon University, Estados Unidos, supervisado por Anupam Gupta y Marijn Heule. Su investigación está concentrada en dos áreas: el uso de herramientas de razonamiento automático para resolver problemas combinatoriales, y el estudio de algoritmos en línea aumentados con consejos de *machine learning*. Su tesis de magíster, investigación presentada en este artículo, fue supervisada por los profesores Jorge Pérez y Pablo Barceló, y obtuvo el primer lugar en el concurso de tesis de magíster de América Latina en áreas relacionadas con inteligencia artificial (LA-CCI). Se le puede contactar por correo, a bsuberca@cs.cmu.edu, o encontrar en Twitter como @b_subercaseaux.

**RESUMEN.** Considerando el alto impacto que tienen las decisiones que toman algoritmos de inteligencia artificial en nuestras vidas, se vuelve fundamental que como ciudadanas y ciudadanos podamos confiar en tales algoritmos, y que podamos interpretar o explicar sus decisiones. Sin embargo, un problema importante del área de interpretabilidad en inteligencia artificial, que apunta a resolver estos problemas, es su falta de definiciones precisas sobre conceptos como *interpretabilidad*, *explicación*, etc. En particular, a pesar de aparecer frecuentemente en la literatura, afirmaciones como *"los árboles de decisión son más interpretables que las redes neuronales"* no tienen un significado formal concreto y, por tanto, es difícil construir una ciencia precisa en base a ellas. Mi trabajo de magíster en el DCC justamente consistió en intentar matizar y formalizar ese tipo de afirmaciones, construyendo una noción matemáticamente precisa de interpretabilidad que permite comparar diferentes tipos de modelos de inteligencia artificial en torno a distintos tipos de explicaciones para sus decisiones. Usando técnicas de algoritmos y complejidad computacional, nuestros resultados consisten en una variedad de teoremas que apoyan y fundamentan la sabiduría popular del área.

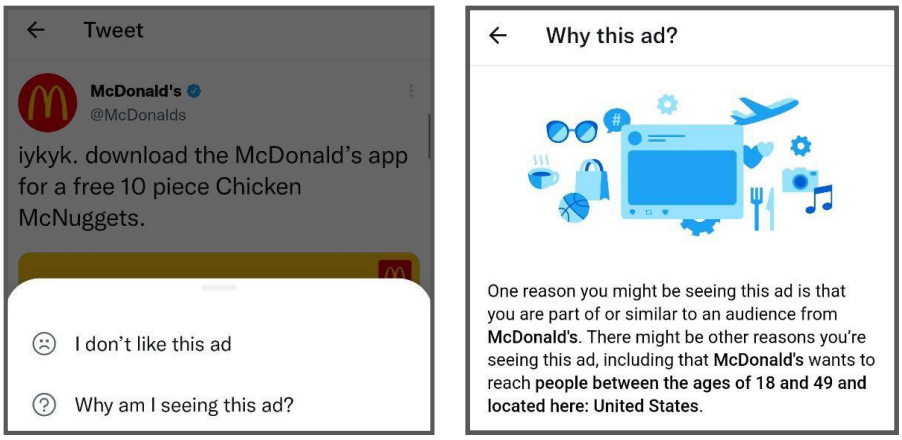



**Figura 1.** Explicación entregada por Twitter sobre los anuncios mostrados a los usuarios.

## ¿Qué es la interpretabilidad y por qué importa?

En el mundo moderno prácticamente todas y todos estamos sometidos a decisiones tomadas por modelos de inteligencia artificial: sistemas de recomendación que deciden cuáles anuncios sugerirnos, qué película recomendarnos, o peor aún, cuál de dos noticias contradictorias mostrarnos primero al buscar un asunto político en Internet. Otras aplicaciones tienen influencias aún más directas en nuestra vida práctica: aseguradoras que usan nuestros datos para predecir nuestro riesgo y así ajustar sus tarifas, bancos que utilizan modelos de inteligencia artificial para decidir si otorgar o no un préstamo, y el escalofriante caso en Estados Unidos del sistema penal utilizando un modelo para predecir reincidencia criminal y con ello influir en sentencias.

Debido a su alto impacto en la ciudadanía, el uso actual de modelos de inteligencia artificial levanta una variedad de preocupaciones éticas e ingenieriles. Un concepto de particular importancia al desplegar tecnologías que hacen uso del aprendizaje de máquinas es el de "confianza"; si la ciudadanía va a ser objeto de decisiones tomadas por algoritmos, es fundamental que pueda confiar en ellos y la calidad de sus decisiones. Más aún, esa confianza no tiene por qué ser a ciegas, y en esa línea hay movimientos que abogan por el "derecho a explicaciones" [1], visión en la cual se entiende como un derecho de las personas el poder inspeccionar y recibir explicaciones sobre las decisiones a las cuales están siendo sometidas.

Por ejemplo, al recibir un anuncio en Twitter, es posible preguntar *por qué uno ha recibido tal anuncio*. La respuesta usualmente consiste en algunas características de nuestro perfil a las cuales los anunciantes apuntan como público objetivo. Como puede apreciarse en la Figura 1, la explicación es bastante incompleta e imprecisa; no a toda persona entre 18 y 49 años en Estados Unidos se le presentan anuncios de McDonald's, lo que implica que otros factores están siendo considerados para decidir si presentarnos tales anuncios.

El área de *interpretabilidad* o *explicabilidad* en inteligencia artificial (XAI)[1] tiene por objetivo el diseño de modelos

1 La diferencia entre los conceptos de *interpretabilidad y explicabilidad* es materia de discusión y, más en general, aún no hay consenso sobre sus definiciones [2,3].

**Figura 2.** Los mapas de saliencia para predicciones contradictorias pueden ser extremadamente similares. Imagen obtenida de [9], quien rinde a su vez cortesía a Chaofan Chen y Checkermallow.

de inteligencia artificial en los cuáles los humanos podamos confiar y utilizar eficazmente [4]. Una variedad de definiciones de interpretabilidad puede reunirse alrededor de la siguiente idea: "*La interpretabilidad tiene que ver con el grado en el cual los humanos podemos entender las causas de una decisión*" [5–7]. Notemos que se trata de una idea bastante vaga; la semántica de las palabras "entender" o "causa" es bastante ambigua. A pesar de su creciente relevancia [8], en particular entendiendo el creciente número de aplicaciones de inteligencia artificial que afectan a la ciudadanía, el área de interpretabilidad o explicabilidad sufre de un fuerte problema definicional: no hay común acuerdo en lo que interpretabilidad o explicabilidad quieren decir exactamente, o cómo deben medirse. Es en este contexto que se enmarcó mi trabajo de magíster, con el objetivo de estudiar a través de formalismos matemáticos la interpretabilidad de diferentes modelos de inteligencia artificial.

## Hacia una noción formal de interpretabilidad

Antes de presentar nuestro trabajo, es relevante entender por qué tiene sentido esforzarse en entender la interpretabilidad desde un punto de vista matemático y formal. El famoso artículo *The Mythos of Model Interpretability* [3] ilustra cómo la falta de definiciones claras para conceptos como interpretabilidad es un problema importante en el área; a falta de definiciones claras y métricas consistentes no es posible evaluar adecuadamente el progreso en el área, o comparar los diferentes métodos. Más aún, el artículo *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead* [9], destaca una serie de problemas prácticos con la forma en que los métodos actuales buscan explicar las decisiones que toman modelos opacos como las redes neuronales profundas. Expondré un ejemplo de Rudin. Una tarea clásica en visión computacional es la clasificación de imágenes; detectar, por ejemplo, si una imagen contiene o no un perro siberiano. Imaginemos por un momento que tenemos una red neuronal profunda que ha aprendido a resolver esta tarea con gran precisión. ¿Cómo podemos estar convencidos de que realmente ha aprendido a reconocer perros siberianos? Una hipótesis alternativa, por ejemplo, es que la mayoría de las fotos de perros siberianos con que ha sido entrenada la red ocurren en un paisaje con nieve, y en realidad la red no ha aprendido nada sobre siberianos, sino simplemente que sobre una cierta cantidad de pixeles blancos es conveniente predecir que hay un siberiano en la imagen. Para descartar tales hipótesis, una idea bastante natural es intentar identificar cuáles *pixeles,* o características más en general, están siendo relevantes en una decisión. El problema empieza al definir lo que "relevantes" quiere decir. Varias técnicas de interpretabilidad sobre visión computacional, bajo el nombre de *mapas de saliencia* [8], se basan en computar (o aproximar) el gradiente de la predicción con respecto a las características (i.e., por cada característica, estudiar el efecto que tiene en la predicción el perturbarla según una pequeña cantidad). Es decir, implícitamente definen las características más *relevantes* como aquellas que más efecto tienen localmente en la función de predicción. La Figura 2 ilustra cómo esta idea puede carecer completamente de poder explicativo. Fundamentalmente, el problema es que decir que un conjunto $S$ de características tiene un gradiente de gran magnitud en una localidad del espacio no tiene por qué explicar nada. En otras palabras, si entendemos los mapas de saliencia como una explicación, pareciera que estos requieren a su vez otra explicación para entender

***Si la ciudadanía va a ser objeto de decisiones tomadas por algoritmos, es fundamental que pueda confiar en ellos y en la calidad de sus decisiones.***

lo que implican. Este problema, de "explicaciones que requieren explicación" se ha comentado repetidas veces en la literatura.

En este contexto de falta de definiciones, y abundancia de técnicas de explicabilidad cuya semántica no es del todo clara, nuestro trabajo busca avanzar la discusión sobre interpretabilidad en una dirección más formal y basada en principios. En particular, es muy común encontrar en la literatura afirmaciones como *"Los árboles de decisión son modelos claramente interpretables, mientras que las redes neuronales profundas son cajas negras no interpretables"*, sin embargo, precisamente por la ausencia de definiciones concretas, no es fácil verificar tales afirmaciones.

Una alternativa posible es llevar a cabo estudios experimentales, en los cuales científicas o científicos de datos evalúan manualmente qué tan interpretables son diferentes modelos en particular, entrenados sobre un conjunto de datos particular. Esta alternativa, si bien útil en la práctica, posee algunos problemas metodológicos. En primer lugar, en ausencia de una definición clara de interpretabilidad, los experimentos suelen definir criterios de interpretabilidad particulares a la situación, los cuales pueden ser cuestionables. Por ejemplo, un experimento puede consistir en presentar a los participantes dos modelos, sus decisiones sobre un conjunto de datos, y luego frente a datos nuevos hacer que los participantes predigan lo que cada modelo decidirá. La idea sería entonces que, si un modelo es más interpretable, será más fácil para el participante humano predecir su comportamiento frente a nuevos datos. De manera implícita, un estudio en esta línea ha definido la interpretabilidad en base a nuestra capacidad de predecir futuras predicciones de un modelo. Por otra parte, una comparación experimental en un momento particular, con un grupo particular de participantes no necesariamente reflejará una verdad subyacente; de la misma forma, no puede decirse de manera definitiva que un algoritmo *A* es más eficiente que un algoritmo *B* solamente en base a un experimento particular, ejecutado en una máquina particular, en base a unas implementaciones particulares.

El aporte principal de nuestro trabajo [10] es presentar una definición formal para establecer cuándo un modelo *A* es más interpretable que un modelo *B*, lo que nos permite establecer afirmaciones como *"para este determinado tipo de explicación, los árboles de decisión son más interpretables que las redes neuronales"* no ya apelando a la intuición humana, sino a través de un teorema matemático. Nuestro punto de partida es flexibilizar la noción de *"A es más interpretable que B"* al parametrizarla según diferentes tipos de explicaciones (que detallaré más adelante); en lugar de decir *"A es más interpretable que B"*, diremos *"A es más interpretable que B con respecto al tipo de explicación E"*. Una ventaja de esta decisión es su flexibilidad frente a los distintos tipos de explicaciones posibles, al no forzar la noción de interpretabilidad a referir a ninguna forma de explicación en particular. Ahora, dado un tipo de explicación *E*, diremos que *"A es más interpretable que B con respecto al tipo de explicación E"* si computar una explicación de tipo *E* sobre el modelo *A* tiene menor complejidad computacional que hacerlo sobre el modelo *B*.

## Una brevísima introducción a la complejidad computacional

La complejidad computacional es un formalismo matemático, cuyo desarrollo empezó en los años sesenta, que permite comparar formalmente la dificultad de problemas computacionales. Usualmente, la complejidad computacional de un problema se define en términos del mínimo número de pasos que un computador (i.e., una máquina de Turing) debe realizar para resolver el problema. Determinar tal mínimo número de pasos, incluso asintóticamente, es extremadamente difícil en la práctica; la comunidad matemática no ha podido determinar tal número para prácticamente ningún problema. Por ejemplo, no se sabe cuál es el mínimo número de pasos que se requieren para un problema tan simple como: *dado un arreglo de números enteros, potencialmente negativos, determinar si existen 3 de ellos que suman 0*. No obstante, el área de complejidad computacional es capaz de determinar que un problema *A* es más difícil que otro problema *B*, sin necesariamente saber la complejidad exacta de *A* o de *B*. ¡Esto suena muy útil para nuestro problema de interpretabilidad, porque podremos comparar la interpretabilidad de distintos modelos sin estar forzados a definirla para un modelo aislado! ¿Pero cómo es posible decir que un problema *A* es más difícil que un problema *B* sin saber qué tan difícil es cada uno? La técnica usual se llama *reducción* y consiste en mostrar que, si uno fuese capaz de resolver el problema *A* eficientemente, entonces eso permitiría resolver el problema *B* eficientemente. Por ejemplo, si consideramos *A = dados dos números x e y, computar x+y*; *B = dado un número x, computar 2x,* podremos decir que *A* es más difícil que *B* ya que, de saber resolver *A* eficientemente, puede computarse *B*(x)

***El área de interpretabilidad sufre de un fuerte problema definicional; no hay común acuerdo en lo que interpretabilidad quiere decir exactamente, o cómo debe medirse.***

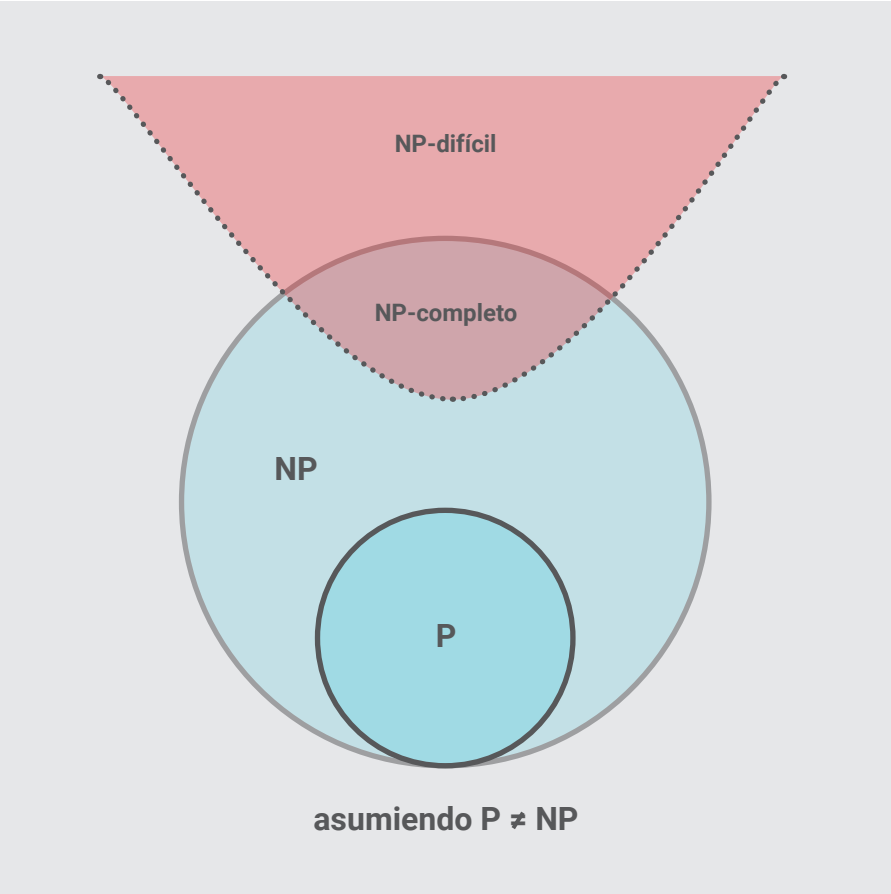



**Figura 3.** Diagrama de las clases de complejidad P y NP.

simplemente como *A(x, x)*. Por el contrario, no es evidente cómo una máquina que resuelve *B*, es decir, que duplica números, puede usarse para resolver el problema *A*. Para organizar estas comparaciones de dificultad entre problemas, el área de complejidad computacional ha definido una serie de *clases*, las cuáles contienen todos los problemas que no son más difíciles que uno dado (ver Figura 3). Por ejemplo, la clase *NP* puede definirse como la clase de problemas que no son más difíciles que el problema de *satisfacibilidad booleana*, es decir, determinar si una fórmula de la lógica proposicional puede hacerse verdadera bajo alguna asignación de sus variables. Por otra parte, la clase *P* puede definirse como los problemas que no son más difíciles que una restricción particular del problema de satisfacibilidad booleana [11]. Cuando un problema *A* no es más difícil que el problema de satisfacibilidad booleana, decimos que *A pertenece a NP*. Por el contrario, si es igual o más difícil decimos que es *NP-difícil* (i.e., difícil con respecto a *NP*). Finalmente si, un problema es difícil con respecto a una clase a la cual pertenece, se dice que es *completo* para esa clase.

Probablemente el problema abierto más importante de la computación teórica es si acaso las clases *P* y *NP* son iguales. Mientras tanto, se cree fuertemente que *NP* contiene problemas más difíciles que *P*, y por tanto al demostrar que un cierto problema *A* es *NP-difícil*, y que en cambio otro problema *B* está en *P*, se suele considerar que esto es evidencia contundente para decir que *A es más difícil que B*. ¡Usaremos estas ideas para demostrar que interpretar modelos como una red neuronal profunda es más difícil que interpretar un árbol de decisión, bajo distintos tipos de explicaciones!

## Tipos de explicaciones

Se presentan tres tipos de explicaciones distintas a través de un ejemplo. Consideremos un banco no particularmente ético que utiliza un modelo de inteligencia artificial para decidir si un postulante debiese o no recibir un determinado préstamo.

Habiendo sido rechazada la postulación presentada en la Figura 4, consideremos los siguientes tipos de explicaciones:

1. El banco no da préstamos a personas con antecedentes penales.
2. Si el postulante hubiese tenido una casa propia, se le habría dado el préstamo.
3. La mayoría de los postulantes casados y sin trabajo estable son rechazados.

A una explicación del primer tipo le llamamos *mínima razón suficiente*, y se formaliza como un subconjunto minimal de las características tal que

***El aporte principal de nuestro trabajo es presentar una definición formal para establecer cuándo un modelo A es más interpretable que un modelo B.***

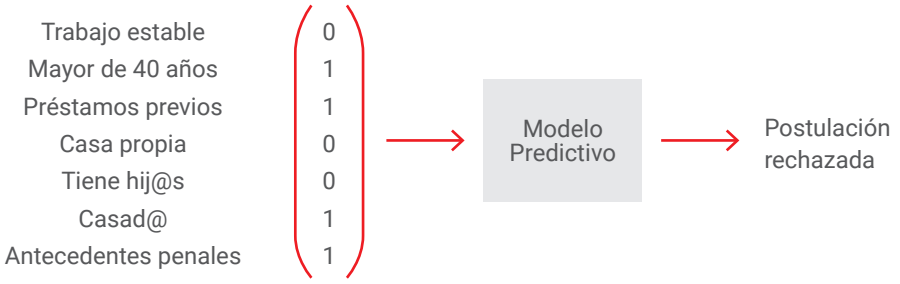



**Figura 4.** Ilustración simplificada de un modelo predictivo siendo usado por un banco para decidir si otorgar un préstamo a un postulante.

| | Á. de Decisión | Perceptrón | MLPs |
|---|---|---|---|
| **Mínimo cambio necesario** | P | P | NP-completo |
| **Mínima razón suficiente** | NP-completo | P | $\Sigma_2^p$ |
| **Conteo de compleciones** | P | #P-completo | #P-completo |

**Tabla 1.** Resumen de resultados de complejidad para distintos tipos de explicaciones sobre distintos tipos de modelos.

cualquier entrada que respete tales características obtendrá el mismo veredicto. En el ejemplo presentado, cualquier instancia posible que tenga un 1 en la componente correspondiente a *antecedentes penales* será rechazada. A una explicación del segundo tipo le llamamos *mínimo cambio necesario*, en cuanto se basa en el mínimo número de características que se deberían cambiar en una instancia para cambiar el veredicto que un modelo en cuestión le da. El tercer tipo, *conteo de compleciones*, es un poco más complejo, y se basa en computar la proporción de instancias que son aceptadas o rechazadas condicionando en algún conjunto de características, como por ejemplo estar casado y no tener trabajo estable. Más aún, en este contexto es posible formalizar la noción de *decisión sesgada*, como aquella en que el veredicto puede cambiar al modificar una característica protegida (edad, género, raza, etc.). Un teorema sencillo de demostrar es que una decisión es sesgada si y solamente si es posible explicarla a través de una mínima razón suficiente que contiene una característica protegida.

Nuestro trabajo considera también otros tipos de explicaciones [12] y está fuertemente inspirado por el trabajo de [13]. Más aún, hemos extendido nuestro análisis a tipos arbitrarios de explicaciones expresables en un lenguaje lógico de primer orden [14].

## Resultados

Nuestro estudio se centra en tres clases de modelos: *perceptrones* (o modelos lineales), *árboles de decisión*, y *perceptrones multicapa* (redes neuronales). La Tabla 1 resume nuestros resultados, determinando la complejidad de los diferentes tipos de explicaciones, sobre los diferentes modelos estudiados. Las clases de complejidad *#P* y $\Sigma_2^p$ tienen definiciones ligeramente más complicadas que aquí omitimos, pero lo relevante es que son clases que se creen incluso más difíciles que *NP*. Es decir, nuestros resultados se pueden resumir diciendo que, para los diferentes tipos de explicaciones considerados, el computarlas sobre perceptrones multicapa es más difícil que sobre árboles de decisión o perceptrones, pero que la comparación entre estos dos últimos depende crucialmente del tipo de explicación considerada.

A continuación presentaré una versión simplificada de las demostraciones más sencillas. Consideremos el problema del *mínimo cambio requerido*. Para demostrar que es *NP*-completo para redes neuronales, usaremos primero un resultado conocido que dice que dada una fórmula booleana $F$, es posible construir eficientemente una red neuronal $M_F$ que la simula. Esta propiedad no la tienen ni los árboles de decisión ni los modelos lineales. Asumamos ahora que queremos decidir si $F$ es satisfacible, y recordemos que la definición de *NP* nos dice que si demostramos que resolver el problema de *mínimo cambio requerido* sobre redes neuronales permite determinar si $F$ es satisfacible, entonces habremos demostrado que nuestro problema es *NP*-díficil. Asumamos además que $F$ no es satisfecha al asignar todas sus variables a 0, pues de lo contrario la satisfacibilidad de $F$ es trivial. Construyamos ahora la red $M_F$, y observemos que dada su equivalencia a

***Usaremos esta idea [de reducción del área de complejidad computacional] para demostrar que interpretar modelos como una red neuronal profunda es más difícil que interpretar un árbol de decisión, bajo distintos tipos de explicaciones.***

*F*, debe cumplir que $M_F(\bar{0}) = 0$, donde $\bar{0}$ representa el vector de 0s. Ahora, sí el *mínimo cambio requerido* a $\bar{0}$ para cambiar su veredicto en $M_F$ es finito, entonces eso quiere decir que existe una instancia $\bar{x}$ tal que $M_F(\bar{x}) = 1$, y por tanto una asignación de variables asociada a $\bar{x}$ que satisface a *F*. Esto muestra que resolver el problema de *mínimo cambio requerido* permite determinar la satisfacibilidad de *F* y, por tanto es un problema *NP*-difícil. Por el contrario, para los árboles de decisión y los modelos lineales, el problema puede resolverse en tiempo polinomial, mediante programación dinámica y un algoritmo *glotón*, respectivamente. Curiosamente, si a las diferentes características se les asocia un costo de cambio, por ejemplo diciendo que cambiar de "no tener casa propia" a "tener casa propia" cuesta más que de "no tener auto" a "tener auto", entonces el problema se vuelve *NP*-completo para modelos lineales, pero sigue pudiendo resolverse eficientemente sobre árboles de decisión mediante programación dinámica. ■

## REFERENCIAS

[1] Bryce Goodman and Seth Flaxman. European Union Regulations on Algorithmic Decision Making and a "Right to Explanation". AI Magazine, 38(3):50–57, 2017.

[2] Leilani H. Gilpin, David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. Explaining explanations: An overview of interpretability of machine learning. In 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2018.

[3] Zachary C Lipton. The mythos of model interpretability. Queue, 16(3):31–57, 2018.

[4] David Gunning and David Aha. DARPA's explainable artificial intelligence (XAI) program. AI Magazine, 40(2):44–58, 2019.

[5] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267:1–38, 2019.

[6] Or Biran and Courtenay V. Cotton. Explanation and Justification in Machine Learning: A Survey. 2017.

[7] Finale Doshi-Velez and Been Kim. A Roadmap for a Rigorous Science of Interpretability. CoRR, abs/1702.08608, 2017.

[8] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García, Sergio Gil-López, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58:82–115, 2020.

[9] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5):206–215, 2019.

[10] P. Barceló, M. Monet, J. Pérez, and B. Subercaseaux. Model interpretability through the lens of computational complexity. In 34th Conference on Neural Information Processing Systems (NeurIPS), 2020.

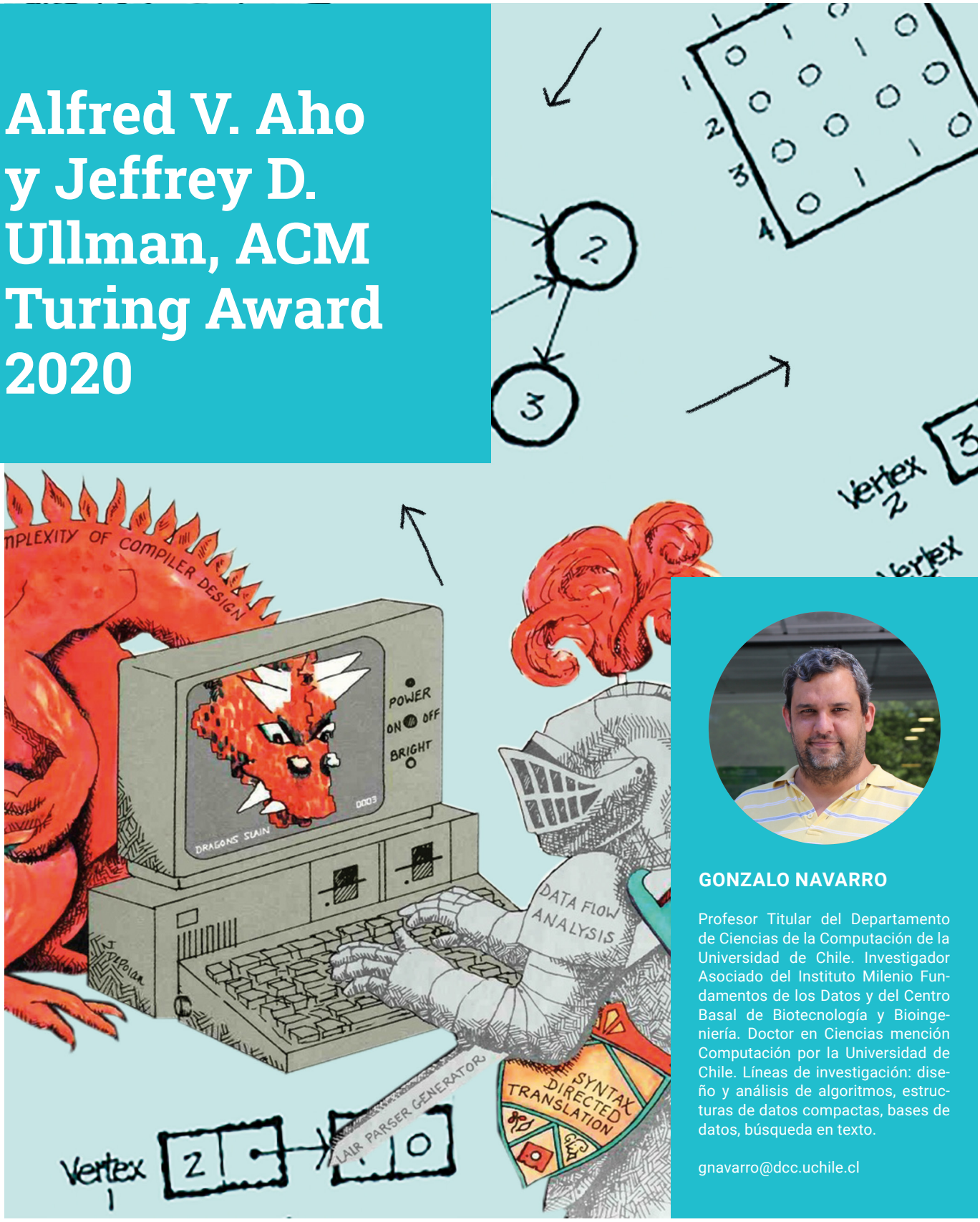
[11] Sanjeev Arora and Boaz Barak. Computational complexity: a modern approach. Cambridge University Press, 2009.

[12] B. Subercaseaux. Model Interpretability through the Lens of Computational Complexity. Tesis de Magíster. Universidad de Chile, 2020.

[13] A. Darwiche and A. Hirth. On the reasons behind decisions. In ECAI, pages 712–720, 2020.

[14] M. Arenas, D. Báez, P. Barceló, J. Pérez and B. Subercaseaux. Foundations of Symbolic Languages for Model Interpretability. In 35th Conference on Neural Information Processing Systems (NeurIPS), 2021.

# Alfred V. Aho y Jeffrey D. Ullman, ACM Turing Award 2020





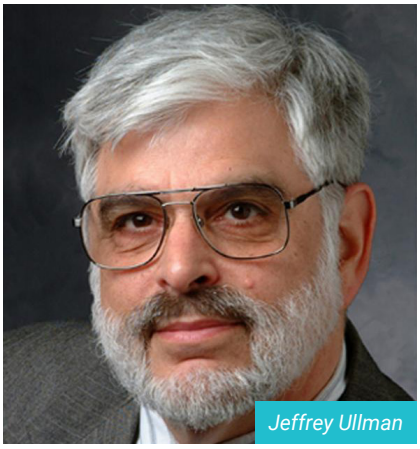

**GONZALO NAVARRO**

Profesor Titular del Departamento de Ciencias de la Computación de la Universidad de Chile. Investigador Asociado del Instituto Milenio Fundamentos de los Datos y del Centro Basal de Biotecnología y Bioingeniería. Doctor en Ciencias mención Computación por la Universidad de Chile. Líneas de investigación: diseño y análisis de algoritmos, estructuras de datos compactas, bases de datos, búsqueda en texto.

gnavarro@dcc.uchile.cl

**RESUMEN.** El que Aho y Ullman hayan sido distinguidos en 2020 con el ACM Turing Award, más que una sorpresa, se siente como un "¡ya era hora!". Sus nombres aparecen asociados a las ramas clave de la computación desde los años sesenta, en que la disciplina dejaba de ser una curiosidad y se empezaba a tomar en serio.

Aho y Ullman realizaron contribuciones fundamentales en la teoría, implementación y divulgación en el área de lenguajes formales, compilación, algoritmos, lenguajes de programación y bases de datos. En un estilo poco común hoy en día, muchas de sus contribuciones se encuentran en forma de libros donde recopilaron todo lo que había que saber, y que han sido claves en la formación de millones de programadores e investigadores por décadas. Parafraseando a Aho, sus técnicas llevaron la compilación de ser un arte a ser una ciencia. ¡Y no son personajes obsoletos! Hoy en día se interesan en temas como integración de datos, paralelismo masivo o computación cuántica.




Fuente: https://awards.acm.org/about/2020-turing

**Ganadores del Premio Turing 2020.**

Una de las experiencias más hermosas de las que tengo memoria al estudiar Computación en los ochenta es aquella perfecta combinación de lenguajes formales, *parsing*, compilación, semántica, optimización de código, lenguajes de programación y algoritmos. Estos temas formaban un todo armónico que siempre percibí como el núcleo científico de nuestra maravillosa disciplina, posiblemente con el agregado de bases de datos. Los nombres de Aho, Hopcroft y Ullman aparecían donde fuera que uno se asomara, como una versión siglo XX de aquellos últimos universalistas del Renacimiento que sabían todo de todo, y que eran los padres fundadores de todo el conocimiento. El que Aho y Ullman hayan sido distinguidos en 2020 con el ACM Turing Award (Hopcroft lo fue en 1986), más que una sorpresa, se siente como un "¡ya era hora!" para estos todavía activos fundadores de la Computación moderna.

No es casualidad que ambos recibieran juntos la distinción más alta de nuestra disciplina. El comienzo de su historia académica se remonta a la época en que la Computación dejaba de ser una curiosidad y se empezaba a tomar en serio. Ambos se habían graduado en 1963 (Aho en la Universidad de Toronto, Ullman en la Universidad de Columbia) y se conocieron el mismo año, en la fila de admisión del Doctorado en Ingeniería Eléctrica y Computación de la Universidad de Princeton. Fue el comienzo de una prolífica amistad y colaboración científica que duraría muchas décadas, y que literalmente le daría forma a nuestra disciplina como la conocemos hoy.

Corrieron suerte distinta en sus doctorados: aparentemente Ullman lo tuvo más fácil. Aho recuerda que "Jeff había encontrado su tema de tesis al comienzo de su segundo año, y se pasó todo el tercer año tipeando su tesis, página por página. Mientras tanto, yo me tiraba de los pelos porque todavía no tenía un tema". Para peor, el profesor guía que había reclutado a Aho se fue a trabajar a la Universidad de Stanford, dejándolo huérfano. Como una confirmación de que no hay mal que por bien no venga, el nuevo profesor guía que encontró Aho fue John Hopcroft.

Ullman se doctoró en 1966, con una tesis titulada "Synchronization Error Correcting Codes", y Aho un año después, con la tesis "Indexed Grammars, an Extension of Context Free Grammars". Ullman publicó sus primeros resultados en el *IBM Journal of Research and Development* (1965) y en las *IEEE Transactions on Information Theory* (1966), mientras que Aho publicó su tesis en *SWAT'67* (hoy *FOCS*), y la versión de revista en el *Journal of the ACM* en 1968.

Apenas doctorados, ambos se fueron a trabajar a los Bell Laboratories e inmediatamente comenzaron a colaborar en el desarrollo de algoritmos eficientes para el *parsing* y compilación de lenguajes de programación. No era una casualidad. Estamos en los "míticos años sesenta", en los que nacieron el sistema operativo Unix y el lenguaje de programación C, precisamente en los Bell Labs. Las publicaciones de sus resultados científicos también comienzan a aparecer desde 1968.

Asimismo, su vocación por resumir y divulgar el conocimiento muestra las primeras señales, con su bello survey "The Theory of Languages" en *Mathematical Systems Theory* (1968).

Ullman retornó pronto a Princeton, en 1969, y se mudó a la Universidad de Stanford en 1979, donde pasó a ser profesor emérito en 2002. Aho se quedó mucho más en Bell Labs, hasta que se mudó a la Universidad de Columbia en 1995, de donde también es profesor emérito. Su colaboración continuó ininterrumpidamente todo este tiempo, y es difícil exagerar el impacto que tuvo.

No es necesario explicar a alguien con formación en Computación la relevancia de los lenguajes de programación, que nos permiten expresar complejos procesos lógicos a un nivel suficientemente alto como para desarrollar el software tremendamente sofisticado que se requiere en un mundo cada vez más dependiente de la Computación. Los computadores no entienden estos lenguajes, sino unos mucho más elementales (el lenguaje de máquina), en los cuales la programación de tareas relativamente sencillas es un desafío intelectual formidable. Seríamos incapaces de hacer funcionar el mundo moderno programando en lenguaje de máquina. Simplemente no tenemos la capacidad intelectual necesaria. El desarrollo de sistemas cada vez más elaborados y complejos ha sido posible solo gracias al desarrollo de los lenguajes de programación llamados de alto nivel, y de otros lenguajes de aún más alto nivel por sobre ellos, como lenguajes para manipular bases de datos (SQL, SPARQL), para realizar procesamiento paralelo (MapReduce), para realizar procesamiento matemático y estadístico (Maple, R), y muchos otros.

Los sofisticados sistemas a los que estamos acostumbrados hoy en día utilizan varios niveles de lenguajes: un lenguaje gráfico para acceder a bases de datos se compila a SQL, el cual a su vez se traduce a C, y este finalmente a lenguaje de máquina. ¡Incluso escribimos artículos en LaTeX, un lenguaje que se compila a Tex, y este a su vez se traduce a PDF! Para que esta arquitectura funcione, es fundamental poder traducir eficientemente un programa escrito en un lenguaje de alto nivel a otro de más bajo nivel (ver Figura 1). Los años sesenta vieron surgir los primeros lenguajes de programación de alto nivel, y el problema de cómo traducirlos eficientemente a lenguaje de máquina era crucial para el desarrollo de la disciplina. Pero este era un problema complejo. Debe existir un conjunto de reglas sintácticas para escribir programas correctos, y esas reglas en sí se escriben en un lenguaje formal, por ejemplo con gramáticas libres del contexto. Debe diseñarse un mecanismo automático, un *parsing*, que utilice las reglas sintácticas y convierta el programa en una representación interna a partir de la cual se pueda generar un programa equivalente en lenguaje de máquina (ver Figura 2). Debe existir una semántica clara que describa qué hace exactamente un programa de alto nivel, de forma de asegurarse de traducirlo correctamente a lenguaje de máquina e incluso optimizarlo para que ejecute más rápido sin cambiar lo que debería hacer. Y todo esto debe hacerse eficientemente para poder traducir programas complejos de miles de líneas.

Aho y Ullman realizaron contribuciones fundamentales en la teoría, implementación y divulgación en el área de lenguajes formales, *parsing*, compilación, y algoritmos y su análisis, estableciendo además importantes conexiones entre ellas. En un estilo que es menos común hoy en día, muchas de sus contribuciones se encuentran en forma de

***En los ochenta los nombres de Aho, Hopcroft y Ullman aparecían donde fuera que uno se asomara, como una versión siglo XX de aquellos últimos universalistas del Renacimiento que sabían todo de todo.***

**Figura 1.** Proceso de compilación.

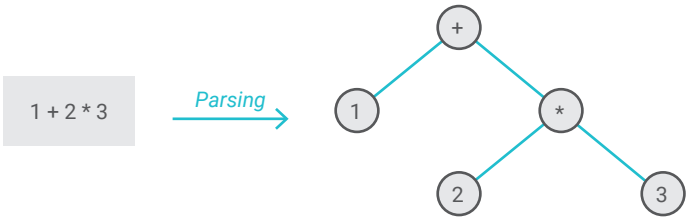



**Figura 2.** Ejemplo de *parsing* de una expresión aritmética.

**Figura 3.** Serie de libros de Aho y Ullman sobre compilación.

libros (ver Figuras 3 y 4) donde recopilaron lo que había que saber acerca de varias de estas áreas fundamentales, incluyendo técnicas desarrolladas por ellos y por otros. Estos libros han sido claves en la formación de millones de ingenieros de software e investigadores en Computación durante décadas, y en muchos casos son la forma más directa de exposición que hemos tenido todos nosotros a los logros de Aho y Ullman (y Hopcroft, en muchos casos). Parafraseando a Aho, estas técnicas llevaron la compilación de ser un arte a ser una ciencia. Aho también cuenta que la idea de escribir estos libros se la dio Richard Hamming, que también estaba en los Bell Labs. "Si quieren que su investigación tenga realmente impacto, deben enseñar a los demás a utilizarla, y la mejor forma de hacer eso es un libro", les dijo.

El más antiguo de estos libros es *"The Theory of Parsing, Translation, and Compiling"*: parte 1, Parsing, de 1972 y parte 2, Compiling, de 1973. Su sucesor, "*Principles of Compiler Design*" (1977), conocido como "el libro del dragón verde" por la ilustración de su tapa[1], contenía toda la teoría relevante a la compilación, desde lenguajes formales y parsing hasta diseño de compiladores, estableciendo un esquema claro y modular para llevar a cabo el proceso completo. Contiene también contribuciones algorítmicas de los autores para hacer el proceso más eficiente. Este libro y sus sucesivas ediciones se consideran la biblia del área y son usados en cursos universitarios y en la industria alrededor del mundo. La última edición de este libro, de 2007, se titula *"Compilers: Principles, Techniques and Tools"* e incluye como nuevos coautores a Monica Lam y Ravi Sethi.[2]

El siguiente libro más antiguo es el también famosísimo "*The Design and Analysis of Computer Algorithms*" de 1974, con Hopcroft, que fue por décadas el libro clásico en los cursos de Diseño y Análisis de Algoritmos en todas las universidades. El libro introduce, por ejemplo, el modelo RAM de computación (Random-Access Machine), que permite analizar los algoritmos de un modo realista para los computadores que utilizamos, y a la vez suficiente-

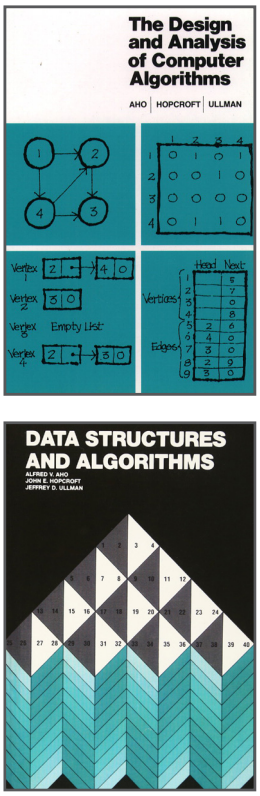



**Figura 4.** Libros de Aho y Ullman sobre algoritmos y estructuras de datos.

1 El dragón, una idea de Ullman, representaba la dificultad del problema, y era enfrentado por un caballero medieval. Cuenta Ullman que las tapas de los libros eran usualmente formales entonces, y que logró colar la idea. El presidente de Addison-Wesley vino luego a decirle que creía que era la peor tapa de la historia. A los tres meses, con las ventas disparadas, volvió para decirle que era la mejor.

2 La tapa de la edición de 1986 apareció fugazmente en la película "Hackers" en 1995. Aho cuenta que solo cuando vieron su libro ahí, sus hijos empezaron a pensar que su padre hacía algo interesante.

***Además de sus nueve libros en conjunto, Aho y Ullman tienen una amplia producción de artículos científicos de gran impacto (Aho sobre 100, con 90 mil citas, y Ullman sobre 300, con 135 mil citas).***

mente general para que el análisis sea factible y universal. El libro también contiene varias contribuciones propias de los autores. A pesar de que en décadas recientes han aparecido libros que son hoy más populares (como el de Cormen, Leiserson, Rivest y Stein), ¡algunas partes del curso CC4102 Diseño y Análisis de Algoritmos de nuestro Departamento todavía se basan en aquel libro! Un buen complemento que apareció después es "*Data Structures and Algorithms*" (1983). En su época, los conocíamos como "el Ajo verde" y "el Ajo negro" por el color de sus tapas (ver Figura 4) .

Además de sus nueve libros en conjunto y de varios otros escritos por separado, Aho y Ullman tienen una amplia producción de artículos científicos de gran impacto (Aho sobre 100, con 90 mil citas, y Ullman sobre 300, con 135 mil citas), así como software de amplísima difusión y utilidad. Es un ejercicio muy interesante pasear por sus publicaciones para descubrir los temas de interés de estos gigantes de la disciplina. No hay una que no toque un tema relevante, no hay compromiso entre calidad y cantidad. Además de sus trabajos conjuntos en teoría de la computación, compiladores y algoritmos, incursionaron en áreas como sistemas operativos (donde Aho profundizó más) y bases de datos (donde Ullman profundizó mucho más).

Un ejemplo de contribución algorítmica es su artículo conjunto (con Hopcroft) para encontrar el ancestro común más bajo de dos nodos en un árbol, un problema que aparece todo el tiempo al desarrollar algoritmos para datos jerárquicos. Ellos propusieron el problema y las primeras soluciones (*STOC*'73 y *SIAM Journal on Computing* 1976), la óptima (por Dov Harel y Robert Tarjan) tuvo que esperar hasta 1984. Otro ejemplo algorítmico es la cota inferior cuadrática para calcular la subsecuencia común más larga entre dos secuencias (con Daniel Hirschberg, *SWAT*'74 y *Journal of the ACM* 1976), un problema crucial en el alineamiento de genomas, detección de plagio y traducción automática. Un tercero es su algoritmo para representar eficientemente la clausura transitiva de un grafo (con Michael Garey, *SIAM Journal on Computing* 1972). En el área de lenguajes para bases de datos, un importante artículo conjunto es su estudio del poder expresivo del álgebra relacional, mostrando que los puntos fijos (por ejemplo, la relevante clausura transitiva) no son expresables, y proponiendo extensiones que lo permitieran (*POPL*'79). Este es un tema sumamente actual en bases de datos de grafos, por ejemplo, donde lenguajes como Datalog y SPARQL permiten clausuras transitivas. Finalmente, Aho destaca en su charla de recepción del Turing Award un artículo en el que muestran cómo describir la sintaxis de lenguajes en forma de gramáticas ambiguas más reglas de desambiguación, de una forma que puede trasladarse directamente a *parsers* eficientes (*POPL*'73 y *Communications of the ACM* 1975, con Stephen Johnson).

Aho trabajó por separado, por ejemplo, en problemas relacionados con búsqueda en texto, tales como el famoso algoritmo de Aho-Corasick (*Communications of the ACM* 1975, con Margaret Corasick) para buscar un gran número de patrones en un texto en tiempo lineal (e implementado por él en el software *fgrep* de Unix). Creó también famosos programas básicos y universales para Unix (y, hoy, Linux), como el software *egrep* de búsqueda de expresiones regulares y el software *awk* de manipulación de texto (la "a" es por Aho) que usan todo el tiempo los usuarios de Linux. Estas herramientas se usaron, por ejemplo, en el conocido software *lex*, que provee la base de análisis sintáctico para desarrollar *parsers* de lenguajes. *Lex* se usa en particular en el igualmente conocido generador de *parsers yacc*, desarrollado sobre las técnicas creadas por Aho, Hopcroft y Ullman. La combinación *lex/yacc*, o sus derivados como *flex/bison*, está en la base de la mayoría de los compiladores modernos.[3] En el área de lenguajes formales, el artículo de su tesis (*Journal of the ACM* 1968), donde propone un tipo de gramática más potente que las libres del contexto para describir lenguajes de programación y que es todavía procesable eficientemente, es uno de los más citados. Aho declara que sus intereses actuales son lenguajes de programación, compiladores, algoritmos y computación cuántica.

Ullman contribuyó en forma decisiva en el área de bases de datos, al punto de que se lo considera uno de los fundadores de la disciplina. Su libro "*Principles of Database Systems*" (1980) es fundacional. Ullman se ha mantenido a la par de la evolución del área, con contribuciones en minería de datos, datos semiestructurados, OLAP, integración de datos, bases de datos deductivas y ciencia de datos. Un ejemplo de sus contribuciones es su técnica de "*magic sets*" para evaluar reglas lógicas en

3 Aho comenta que con estas herramientas sus estudiantes desarrollaban un compilador para algún lenguaje acotado en grupos de cinco durante tan solo un semestre. Me recordó, con nostalgia, que nosotros hicimos lo mismo en esos años, compilando un lenguaje tipo Pascal al Assembler del Intel 8088, optimización incluida, ¡no es una exageración de Aho!

bases de datos, que las reescribe de forma que se puedan evaluar *bottom-up* eficientemente en base a *joins*. Si bien es una heurística, esta técnica es la base para optimizar estos programas (publicado con varios coautores en *PODS'*86). Su artículo más citado (*SIGMOD'*97, con Sergey Brin, Rajeev Motwani y Shalom Tsur) presenta nuevos algoritmos para analizar información de grandes volúmenes de compras pasando pocas veces sobre los datos, de modo de detectar conjuntos de ítems que se compran juntos frecuentemente. Este es un problema muy popular de minería de datos. En términos de tecnología, Ullman fue profesor guía de Sergey Brin (cofundador de Google) y formó parte del directorio técnico de la empresa. Finalmente, su libro "*Introduction to Automata Theory, Languages and Computation*" (1972, con Hopcroft) es uno de los más conocidos y usados en el área. Ullman declara que sus intereses actuales son teoría de bases de datos, integración de datos, minería de datos y educación en línea.

***Es un ejercicio muy interesante pasear por sus publicaciones para descubrir los temas de interés de estos gigantes de la disciplina. No hay una que no toque un tema relevante, no hay compromiso entre calidad y cantidad.***

A los lectores más jóvenes, que tal vez sientan que estamos hablando de caballeros medievales, los invito a ver la charla de recepción del Turing Award, que usa el interesante concepto de abstracción computacional como hilo conductor. Van a descubrir a dos personajes atemporales, que pasan en un instante de hablar de los inicios de los lenguajes de programación y del álgebra relacional a los temas más candentes de integración de datos, paralelismo masivo o computación cuántica. La entrevista también nos deja entrever algo de las personas detrás de los nombres: mientras Aho está a sus anchas en la entrevista y la disfruta, Ullman sufre visiblemente con cada pregunta. Al final, después de repasar las contribuciones de estos dos gigantes, es toda una lección de humildad el escuchar decir a Aho que simplemente fue afortunado por entrar a la disciplina cuando todavía era fácil hacer contribuciones fundamentales, o decir a Ullman (en otra entrevista) que siente un poco el "síndrome del impostor". ¡Qué nos queda a los demás! ■

**Agradecimientos**

*Agradezco las sugerencias varias de Pablo Barceló y Éric Tanter.*

## REFERENCIAS

- Artículo de ACM sobre el Turing Award 2020.
  https://awards.acm.org/about/2020-turing
- Video en YouTube de la charla y entrevista asociadas al otorgamiento del premio.
  https://www.youtube.com/watch?v=ixIlknu7svM&ab_channel=AssociationforComputingMachinery%28ACM%29
- Artículo de la Universidad de Princeton sobre el Turing Award 2020.
  https://www.dailyprincetonian.com/article/2021/04/princeton-alumni-turing-award-aho-ullman-computer-science
- Páginas de DBLP de Aho y Ullman.
  https://dblp.org/pid/a/AVAho.html
  https://dblp.org/pid/u/JeffreyDUllman.html
- Páginas de Google Scholar de Aho y Ullman.
  https://scholar.google.cl/citations?user=gb2r2ssAAAAJ&hl=en&oi=ao
  https://scholar.google.cl/citations?user=wUJ2bXgAAAAJ&hl=en&oi=ao
- Páginas de Wikipedia de Aho y Ullman.
  https://en.wikipedia.org/wiki/Alfred_Aho
  https://en.wikipedia.org/wiki/Jeffrey_Ullman
- Páginas personales de Aho y Ullman.
  http://www.cs.columbia.edu/~aho/
  http://infolab.stanford.edu/~ullman/

# Nota del editor:

## Sobre las acusaciones a Ullman por su sostenido comportamiento discriminatorio

*La siguiente columna refleja pura y exclusivamente la opinión y punto de vista del editor general de Revista Bits de Ciencia, y es independiente del artículo anterior.*

Cada año, el anuncio del Premio Turing —el Premio Nobel de la computación— es esperado con mucha expectativa y entusiasmo por parte de toda la comunidad científica. En ese sentido, el anuncio del último año fue una rara excepción que sembró un manto de dudas sobre la pertinencia del proceso de selección. Nadie puso en duda el valor de las contribuciones científicas de los premiados, Aho y Ullman; sin embargo muchos levantaron una bandera roja por el sostenido comportamiento discriminatorio que Ullman ha mostrado hacia ciertos grupos a lo largo de las últimas décadas.

Los primeros signos de este tipo de comportamiento se remontan a mediados de los 2000 cuando, cansado de recibir y tener que responder correos de estudiantes iraníes, Ullman publicó su infame página web "Answers to All Questions Iranian", que mantuvo en línea hasta finales de 2020.[1] En la página web, entre otros, manifestaba abiertamente sus sentimientos contra el pueblo iraní y justificaba el genocidio sufrido por los pueblos nativos de Estados Unidos argumentando que "*esa es la manera en la que ocurren y siempre han ocurrido las cosas. Civilizaciones tecnológicamente más avanzadas reemplazan a civilizaciones menos avanzadas*".

Por otro lado, consideraba que los estudiantes que habían crecido en Irán luego de la Revolución Islámica de 1979 representaban un peligro para los Estados Unidos y no se les debía permitir estudiar en el país. Esto se ve reflejado en una discusión que mantuvo por Google+, apoyando el veto a los iraníes que la Universidad de Massachusetts estaba por instaurar en 2015, donde sostenía que *"tenemos que distinguir entre los estadounidenses de ascendencia iraní que han elegido hacer su camino junto a los Estados Unidos, y los iraníes que no se marcharon de Irán cuando los fanáticos religiosos se hicieron del control, y que pueden estar alineados con los deseos de Irán de construir armas nucleares y desplegar terroristas por todo el mundo. Si bien estoy convencido de que muchos de los estudiantes que viven en Irán no tienen mayor anhelo que marcharse del país mientras este siga gobernado por fundamentalistas islámicos, ¿podemos correr el riesgo de educarlos y que luego utilicen esa educación contra nosotros?"*.[2]

Un último ejemplo de la discriminación deliberada de Ullman hacia el pueblo iraní, en este caso bajo la presunción de ciertas posturas políticas, puede verse en una comunicación por correo electrónico que en 2010 mantuvo con un estudiante iraní sobre el proceso de admisión a la Universidad de Standford, de la cual Ullman era profesor emérito. A la consulta del estudiante, Ullman respondió: *"Incluso si estuviera en condiciones de ayudar, tampoco ayudaría a ningún estudiante iraní hasta que Irán no reconozca a Israel como tierra del pueblo judío. Sé que puede que no compartas la misma postura demencial de los mulás que gobiernan tu país, pero es una cuestión de principios. Si los iraníes quieren gozar de los beneficios de Stanford u otras instituciones de los Estados Unidos, deben respetar los valores que tenemos en los Estados Unidos, incluída la libertad de religión y el respeto a los derechos humanos"*.[3]

En respuesta a esto, el National Iranian American Council elevó una queja formal a la Universidad de Standford.[4] Sin embargo, alegando que las observaciones hechas por un académico ya retirado no reflejaban la postura de la Universidad ni su política de admisión, las autoridades de la Universidad decidieron no tomar ninguna medida al respecto, incluso permitiendo que Ullman siga manteniendo su página como parte del sitio web institucional.[5]

Al florecer todos estos antecedentes luego de la premiación, diversos miembros de la comunidad científica decidieron retirar el apoyo que inicialmente habían dado a Ullman. Entre ellos se destacan, por ejemplo, Shafi Goldwasser y Yann LeCun, antiguos ganadores del Premio Turing.[6] Este repudio se volvió mucho más masivo luego de la publicación de una carta abierta firmada por más de

1 https://web.archive.org/web/20200129080549/http:/infolab.stanford.edu/~ullman/pub/iranian.html.

2 https://www.ics.uci.edu/~eppstein/gplus/20150212-VDYSkY69tGe.html.

3 https://drive.google.com/file/d/1ZjydBnYeO8LhU3TyxvYxBTziWsQvSf2t/view.

4 https://web.archive.org/web/20110124051512/http:/www.niacouncil.org/site/DocServer/Stanford_Discrimination_Letter.pdf.

5 https://web.archive.org/web/20140809043733/http:/www.paaia.org/CMS/stanford-university-president-responds-directly-to-paaia-over-retired-professors-anti-iranian-remarks.aspx.

6 https://www.facebook.com/SimonsInstitute/posts/4003638433027737.

mil adherentes de la academia e industria, bajo la etiqueta *CSForInclusion*.[7]

Paradójicamente, la institución encargada de otorgar la premiación —la Association for Computing Machinery (ACM)—, se define como una *"organización científica y educativa global dedicada a promocionar el arte, ingeniería, ciencia y aplicaciones de la computación [...] promoviendo los más altos estándares éticos y profesionales"*, además de declarar a la diversidad, igualdad e inclusión como uno de sus cuatro valores fundamentales.[8]

Ante el rechazo público por la controvertida premiación, la ACM primero reaccionó con un escueto tweet diciendo que *"las afirmaciones hechas en el pasado por Jeffrey Ullman [...] no reflejan la visión de la ACM"*.[9] Luego publicó en su sitio web una respuesta oficial a la masiva carta en la que explicaba que la ACM nunca había recibido ninguna denuncia contra Ullman, que los miembros del comité encargado de la premiación se enteraron de sus antecedentes por la polémica que se disparó luego del anuncio y que a partir de la próxima entrega iban a ajustar su proceso de selección, exigiendo a los nominadores que notifiquen si están al tanto de algún comportamiento del candidato propuesto, que sea inconsistente con los valores de la ACM.[10]

Mientras ese es un buen primer paso, ¿es suficiente?, especialmente teniendo en cuenta todo el esfuerzo que la comunidad de las ciencias de la computación, y STEM (Science, Technology, Engineering, and Mathematics) en general, está haciendo para ampliar la participación a sectores que han estado tradicionalmente excluidos o subrepresentados. Posiblemente sea hora de repensar el rol que queremos que los valores éticos efectivamente ocupen dentro de nuestra comunidad. Abrazarlos seguramente requerirá un cambio cultural muy grande de las instituciones educativas y organizaciones científicas que ante este tipo de situaciones, deberán ir más allá de salvar su imagen, al menor costo posible.

Resulta sorprendentemente paradójico que un científico sostenidamente contrario a la diversidad e inclusión se haga acreedor de una premiación que conmemora el legado de Alan Turing, uno de los fundadores de la teoría de la computación quien tanta discriminación sufrió por su orientación sexual, habiendo sido procesado y condenado por ello. Tuvieron que pasar más de cincuenta años para que recibiera el indulto y perdón póstumo que tantos reclamaban. ¿Tendrán las víctimas de la discriminación por parte de Ullman la misma suerte? Solo el tiempo nos dará una respuesta…

7 https://csforinclusion.wordpress.com/.

8 https://www.acm.org/about-acm/mission-vision-values-goals.

9 https://twitter.com/TheOfficialACM/status/1379891090246004744.

10 https://www.acm.org/response-to-letter.

Banderas rojas:

# El proyecto de ley sobre plataformas digitales, sus múltiples errores y las lecciones que podemos aprender de ellos

**VLADIMIR GARAY**

Director de incidencia y comunicaciones en Derechos Digitales, organización civil que trabaja en la intersección entre tecnología y derechos humanos. Es licenciado en comunicación social y periodista por la Universidad de Chile y magíster en arte, pensamiento y cultura latinoamericana por la Universidad de Santiago de Chile.

vladimir@derechosdigitales.org

*El polémico proyecto de ley presentado por la Comisión de Desafíos del Futuro del Senado erra de forma tan ejemplar, que una revisión de sus formas puede ser útil para futuros esfuerzos legislativos.*

En el mundo anglosajón, muy dado a la creación de neologismos, hace un tiempo se viene hablando del "techlash", la creciente animosidad pública hacia las grandes empresas de tecnología de plataformas de Silicon Valley y sus equivalentes chinos [1]. Si a principios de la década pasada la llamada "Web 2.0" era celebrada como una tecnología eminentemente democrática y democratizante —desde la "Primavera Árabe" a Occupy Wall Street y la primera elección de Barack Obama en Estados Unidos— hoy estaríamos viviendo en el exacto doble opuesto, donde las redes sociales son responsables del colapso de las democracias occidentales. Pruebas de ello son las revelaciones de Edward Snowden sobre el espionaje masivo de la Agencia de Seguridad Nacional (NSA por sus siglas en inglés), la elección de Donald Trump en Estados Unidos y el auge del terraplanismo. Evidentemente, ambos polos son exageraciones y dificultan nuestra capacidad de entender el impacto real de estas tecnologías en nuestras vidas.

Una consecuencia directa de la creciente desconfianza contra las grandes empresas de tecnología es la creación de condiciones favorables a los impulsos regulatorios de todo tipo. No es posible comprender cabalmente los esfuerzos normativos surgidos desde la Unión Europea fuera de este contexto, incluyendo el Reglamento General de Protección de Datos ni la Directiva sobre los derechos de autor en el Mercado Único Digital. La primera es reconocida como el más alto estándar existente en la materia, mientras que la segunda es profundamente problemática.[1] En Estados Unidos actualmente existen distintos intentos por reformar la llamada "Section 230",[2] ampliamente considerada como una ley fundamental para el desarrollo de las empresas de Internet. Mientras tanto, en el Reino Unido se está discutiendo un proyecto denominado "Online Safety Bill", altamente cuestionado por asuntos relativos a la recolección de información personal y en materia de libertad de expresión [2].

Chile no está exento de este fenómeno. En septiembre de 2021 ingresó al Congreso Nacional el proyecto de ley que regula las plataformas digitales [3], firmado por los senadores Francisco Chahuán, Juan Antonio Coloma, Alfonso De Urresti, Guido Girardi y Carolina Goic, todos miembros de la Comisión de Desafíos del Futuro, Ciencia, Tecnología e Innovación. El proyecto ha generado un rechazo transversal entre la comunidad de especialistas a nivel nacional e internacional, y ha sido sindicado como un riesgo importante al ejercicio de derechos fundamentales en entornos digitales.

Este no es el primer intento por legislar respecto al funcionamiento de las plataformas digitales ni lo que las personas hacen en ellas. En julio de 2014, el diputado demócratacristiano Jorge Sabag presentó un proyecto de ley para sancionar con multas y cárcel a quienes atentaran contra la "dignidad de las autoridades" en las redes sociales. La iniciativa —conocida como "la ley antimemes"— fue retirada menos de 24 horas después, debido a la indignación transversal con la que fue recibida. El diputado no solamente reconoció que la propuesta era un error, sino que admitió no haberla leído [4]. Otras propuestas normativas que lidian con la regulación de las plataformas digitales o algunos de sus aspectos son el proyecto de ley contra la violencia de género en línea[3] y el proyecto de ley que sanciona a personas que detentan altos cargos de elección popular en caso de promover o financiar noticias falsas, que hace alusión explícita a las redes sociales.[4] La regulación de plataformas digitales fue abordada también en un proyecto presentado por el ex diputado Marcelo Díaz, el que no llegó a discutirse.

***El proyecto ha generado un rechazo transversal entre la comunidad de especialistas a nivel nacional e internacional.***

Algunas de estas propuestas son buenas, otras no. Sin necesariamente ser exhaustivo, este listado demuestra el interés de las y los legisladores por abordar materias relativas a las plataformas en Internet y su uso. En ese sentido, el proyecto presentado por la Comisión de Desafíos del Futuro tiene una serie de particularidades que lo hacen ejemplar y es un buen caso de estudio para futuros intentos legislativos.

1 Para saber más sobre las críticas a la Directiva sobre los derechos de autor en el Mercado Único Digital, recomiendo chequear la documentación publicada por la European Digital Rights (EDRi), disponible en https://edri.org/our-work/copyright-reform-document-pool/.

2 La sección 230 establece el principio de no responsabilidad de los intermediarios de Internet en Estados Unidos, que básicamente establece que los proveedores de servicios de Internet no pueden ser declarados responsables de los actos realizados por sus usuarios.

3 Boletín 13928-07. Proscribe, tipifica y sanciona la violencia digital en sus diversas formas y otorga protección a las víctimas de la misma. Disponible en http://www.senado.cl/appsenado/templates/tramitacion/index.php?boletin_ini=13928-07.

4 Boletín 12.314-07. Moción de los Senadores señor Navarro, señora Provoste y señores Latorre, de Urresti y Quintana con la que inician un proyecto de reforma constitucional que establece la cesación en los cargos de presidente de la república, senador, diputado, consejero regional, alcalde y concejal por la difusión, promoción o financiamiento de noticias falsas. Disponible en https://www.bcn.cl/laborparlamentaria/wsgi/consulta/verParticipacion.py?idParticipacion=1771361.

**Figura 1.** Declaración de Guido Girardi, respecto al proyecto de ley.

## Banderas rojas

La propuesta legislativa presentada por la Comisión de Desafíos del Futuro presenta una serie de deficiencias y problemas que es necesario examinar con detención, en tanto señales de peligro.

En primer lugar, se trata de un proyecto de ley misceláneo, que pretende sentar norma sobre una larga lista de fenómenos asociados al uso de Internet, profundamente distintos entre sí. En diez páginas y quince artículos, el proyecto se pasea por temáticas altamente complejas, como el tratamiento de datos personales, la libertad de expresión, el controversial "derecho al olvido", la toma de decisiones algorítmicas, la inteligencia artificial y las expresiones injuriosas, por mencionar algunas. Durante las distintas instancias de debate que se han realizado en torno a la propuesta se agrega al listado cuestiones como la "manipulación algorítmica" y la "adicción a Internet", fenómenos sobre los cuales no existe consenso entre especialistas, y en más de una ocasión se han mencionado tópicos derechamente esotéricos, como la "amenaza transhumanista" o "el fin de la sociedad occidental",[5] cuya pertinencia en esta discusión es misteriosa. De lo anterior se desprende no solo una preocupación por la liviandad con la que cada una de estas cuestiones ha sido abordada en el proyecto, sino una duda legítima respecto al grado de conciencia y conocimiento que tienen los legisladores sobre la temática que han decidido regular y las motivaciones que los han llevado a tomar esa determinación.

Otra cuestión que llama la atención sobre el proyecto de ley es que introduce una serie de conceptos absolutamente inéditos en el campo de la regulación de Internet, de significados y alcances imprecisos y profundamente problemáticos. Por ejemplo, se usa la noción de "libertad de expresión digital", que solamente puede ser entendida como una forma particular de libertad de expresión, contraviniendo las numerosas recomendaciones que los organismos internacionales de derechos humanos han hecho a los estados respecto a la manera correcta de abordar y proteger el ejercicio de derechos en los entornos digitales, es decir, del mismo modo en que se protegen en el mundo material.[6] Otro concepto problemático es el de "consumidor digital", que ubica la normativa en el contexto del intercambio de bienes y servicios, y no como una materia propia de derechos fundamentales. Incluso la definición de "plataforma digital" propuesta es cuestionable, pues no establece ninguna distinción entre Facebook, el sitio web del Servicio de Impuestos Internos o un blog personal autoalojado; del modo en que está expresado en el proyecto, serían exactamente lo mismo y sujetos a las mismas consideraciones y obligaciones.

Otros aspectos donde el proyecto busca innovar de forma torpe es mediante el establecimiento de un régimen de "responsabilidad objetiva" por todos los daños ocasionados por una plataforma. Se trata de una medida sin par en ninguna regulación sobre Internet en

5 Estos tópicos fueron abordados de forma recurrente por el (ex) Senador Guido Girardi. Por ejemplo, en la sesión de la Comisión de Desafíos del Futuro realizada el 29 de noviembre de 2021 y disponible aquí: https://tv.senado.cl/tvsenado/comisiones/permanentes/desafios-del-futuro-ciencia-tecnologia-e-innovacion/comision-de-desafios-del-futuro-ciencia-tecnologia-e/2021-11-29/073606.html.

6 Sobre este punto, vale la pena revisar la resolución sobre resolución sobre Promoción, protección y disfrute de los derechos humanos en Internet del Consejo de Derechos Humanos de Naciones Unidas. Disponible en https://documents-dds-ny.un.org/doc/UNDOC/LTD/G18/203/76/PDF/G1820376.pdf?OpenElement y la resolución de la Asamblea General de Naciones Unidas, aprobada el 17 de diciembre de 2018, disponible en: https://undocs.org/pdf?symbol=es/A/RES/73/173.

el mundo, contraria a la recomendación realizada por la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos [5] y que contradice lo establecido nueve artículos antes en el mismo proyecto de ley. Además, al facultar al tribunal a duplicar la indemnización por daños, el proyecto crea la figura de los daños punitivos, que no tiene reconocimiento legal ni consistencia con el sistema jurídico chileno. Otros ámbitos donde el proyecto no considera la legislación vigente son la protección de datos personales, los derechos de los consumidores y la protección de los derechos de niños, niñas y adolescentes.

Esta vocación pionera fue enfatizada a punta de declaraciones grandilocuentes sobre el carácter inédito del proyecto. "Con el PL [sic] de Plataformas Digitales, estamos regulando el ciberespacio, algo que nunca se ha regulado" publicaba el (ex) Senador Girardi en su cuenta de Twitter (ver Figura 1) [6]. "Chile tendrá la primera ley del mundo que regula plataformas digitales" titulaba La Tercera a fines de noviembre (ver Figura 2) [7]. No solamente llama la atención la falta de rigurosidad del medio de comunicación,[7] sino además la amplia cobertura que recibió el proyecto de ley, probablemente asociada al carácter fundacional que se le intentó impregnar a la iniciativa. De hecho, gran parte de la discusión se trasladó a la prensa en vez de generar mejores espacios de trabajo sobre el proyecto.

La liviandad con la que se realizan estas afirmaciones contrasta con la discusión global entre especialistas, todavía incapaces de producir una solución plenamente satisfactoria al problema. Por eso no sorprende que la carta coordinada por Derechos Digitales y dirigida a la Comisión de Desafíos del Futuro (ver Figura 3) [8] haya logrado la adhesión de connotadas personalidades a nivel mundial, sumamente preocupadas por el contenido del proyecto: Daphne Keller, directora del Programa de regulación de plataformas en Stanford; Sonia Livingstone, profesora en psicología social en la London School of Economics y eminencia en derechos de la infancia en la era digital; Michael Karanicolas, director ejecutivo del Institute for Technology, Law & Policy de UCLA; Jessica Fjeld, Directora asistente en la clínica sobre Cyberlaw del Berkman Klein Center for Internet & Society y docente de derecho en la Harvard Law School; el destacado activista y escritor Cory Doctorow; Evgeny Morozov, uno de los pensadores más citados a nivel mundial en materias relativas al impacto social de las tecnologías. Organizaciones como la World Wide Web Foundation, la Wikimedia Foundation, la Electronic Frontier Foundation, la Association for Progressive Communications, Article 19, Access Now y las principales organizaciones latinoamericanas de derechos digitales hicieron un llamado a la Comisión a no perseverar e iniciar un nuevo proceso transparente, abierto y multisectorial, donde participen todas las partes interesadas y con un compromiso de derechos humanos.

## ¿Qué podemos aprender de todo esto?

Tras un par de meses bastante agitados en la Comisión de Desafíos del Futuro, y sin que llegara a discutirse en el pleno del Senado, la tramitación del proyecto entró en receso, directamente relacionado con el recambio de autoridades en el Legislativo. Al momento de escribir estas palabras, la Comisión no ha elegido todavía su nueva presidencia y el futuro del proyecto es incierto. La discusión podría reactivarse mañana o podría suspenderse de forma indefinida. En Chile los proyectos de ley no caducan y se mantienen vigentes siempre y cuando no hayan sido rechazados en el Congreso, por lo que solo queda esperar.

Mientras tanto, vale la pena revisar algunas señales de alerta presentes en la tramitación del proyecto de ley de plataformas digitales y que pueden ser indicadores de problemas en futuras iniciativas legales.

***Incluso la definición de "plataforma digital" propuesta es cuestionable, pues no establece ninguna distinción entre Facebook, el sitio web del Servicio de Impuestos Internos o un blog personal autoalojado.***

**Figura 2.** Repercusión en los medios sobre el proyecto de ley.

7 Evidentemente, lo que señala el titular no es cierto. A nivel local existen numerosas materias relativas al funcionamiento del "ciberespacio" que están reguladas. Por ejemplo, cuestiones como neutralidad de la red, delitos informáticos, interceptación de comunicaciones, disputas de nombres de dominio e infracciones al derecho de autor, por mencionar algunas. A nivel internacional, la lista es extensa y antigua.

# Sociedad civil internacional alerta sobre los peligros para el ejercicio de derechos de proyecto de ley de regulación de plataformas digitales presentado en Chile

POR: DERECHOS DIGITALES

Las organizaciones e individuos firmantes expresamos nuestra preocupación por el avance del Proyecto de ley de regulación de plataformas digitales, Boletín Nº 14.561-19, que se discute en el Congreso chileno. El proyecto, motivado por la necesidad de equilibrio entre el poder de algunas empresas de tecnología y la ciudadanía, establece reglas que terminan siendo peligrosas para el ejercicio de los derechos fundamentales en internet.

En especial, expresamos nuestra profunda preocupación por las razones siguientes:

1. La iniciativa apunta a regular a las "plataformas digitales", que define como "toda infraestructura digital cuyo propósito es crear, organizar y controlar, por medio de algoritmos y personas, un espacio de interacción donde personas naturales o jurídicas puedan intercambiar información, bienes o servicios". El concepto es demasiado amplio y vago, e incluiría a toda clase de servicios, sin importar su tamaño o base de usuarios, sus funcionalidades o propósitos, o si son sitios web o apps móviles, proponiendo una única regulación para muy distintas realidades.

2. La iniciativa intenta innovar en materias de protección de datos personales, derechos de los consumidores y protección de derechos de niños, niñas y adolescentes. Con esto, deja de lado las leyes vigentes y las discusiones legislativas en esas otras materias sin buscar coordinación con ellas, estableciendo reglas separadas que no mejoran el estado general de la protección de esos derechos.

3. La propuesta presenta una definición de "consumidor digital" (artículo 3, letra d), centrando la protección ofrecida en aspectos vinculados al consumo, ignorando impactos sociales, políticos y culturales más allá de las relaciones de consumo que se vinculan al uso de plataformas digitales.

4. La propuesta busca trasladar las reglas que rigen el mundo offline a las plataformas en líneas. Pero lo hace a través de la creación de obligaciones y reglas (artículo 6) que solamente entorpecen el funcionamiento en línea e imponen nuevos deberes de responsabilidad a los intermediarios generando incentivos a la remoción de contenido para limitar esas responsabilidades. De este modo, se discrimina negativamente a las plataformas digitales y se desincentiva a su creación, reduciendo la competencia y afectando a la formación de nuevos espacios de interacción en línea.

5. El proyecto de ley busca combatir la información "manifiestamente falsa", abordando el fenómeno de la desinformación de manera desconectada de la experiencia internacional en la materia. Con esto, lesiona las garantías de derechos humanos de libertad de expresión y libertad de información sin censura previa, contraviniendo las reglas vigentes según el sistema interamericano de derechos humanos.

6. La introducción de mecanismos de verificación de edad apropiados que se propone (artículo 8), aunque es loable en su objetivo de protección de la infancia, hace caso omiso que precisamente la propuesta supone recoger más datos de las usuarias para su identificación.

7. La referencia a no discriminación introducida (artículo 9) resulta vaga y prescinde de cualquier coordinación con la normativa vigente en la materia, imponiendo al proveedor de servicios la implementación de mecanismos de control de sesgos respecto de los cuales no se ofrecen parámetros y se dejan librados a su arbitrariedad en definitiva.

8. El proyecto de ley atribuye la "responsabilidad objetiva" por todos los daños ocasionados por una plataforma (artículo 15), en contradicción con sus propias reglas de exención de responsabilidad (artículo 6), y facultando a los tribunales a duplicar la indemnización por esos daños, creando en Chile la figura de los daños punitivos que no tiene reconocimiento legal ni consistencia con el sistema jurídico chileno. A la vez, ese establecimiento de responsabilidad objetiva es contrario a la recomendación de la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos, al expresar que "un esquema de responsabilidad objetiva en el ámbito de la comunicación electrónica o digital es incompatible con estándares mínimos en materia de libertad de expresión".

9. La propuesta incluye la posibilidad de suspensión o bloqueo temporal de una plataforma digital, lo que contraviene de forma directa las recomendaciones de los organismos de protección de derechos humanos para asegurar y promover una internet libre y abierta, y garantizar la libertad de expresión.

10. Por último, la regulación de plataformas digitales debe fomentar desconcentración de poder, transparencia y rendición de cuentas, no promover censura y entregar más poder a quienes ya lo tienen como controladores de la tecnología.

Por todo lo anterior, pedimos al Congreso chileno:

- No perseverar la tramitación del Boletín Nº 14.561-19.
- Iniciar un nuevo proceso de discusión de las reglas para las plataformas de internet, como un debate transparente, abierto y multisectorial, donde participen todas las partes interesadas, con los compromisos de Chile frente al Derecho Internacional de los Derechos Humanos como la base para cualquier propuesta.

**Figura 3.** Carta dirigida a la Comisión de Desafíos del Futuro.

*Las declaraciones grandilocuentes sobre el carácter inédito de un fenómeno suele ser un indicador de desconocimiento respecto del campo que se quiere regular.*

1. **La vocación pionera:** el hecho de que una tecnología sea nueva no quiere decir que sus impactos sean particularmente distintos a todo lo conocido, ni suficientemente extendidos para merecer una regulación especial. Lo contrario significaría tener que generar leyes para cada nueva tecnología, lo que es absurdo e imposible.

   Muy rara vez es necesario partir de cero. Usualmente existen ejemplos de legislación que lidian con fenómenos análogos en la legislación local y en la experiencia comparada a nivel internacional. Las declaraciones grandilocuentes sobre el carácter inédito de un fenómeno suele ser un indicador de desconocimiento respecto del campo que se quiere regular.

   El mejor ejemplo de proyectos de ley imbuidos de esta vocación pionera son aquellos sobre "neuroderechos", presentados por la misma Comisión de Desafíos del Futuro, los que equivocadamente asumen que, puesto que las normas se refieren a tecnologías que todavía no existen, requieren de la creación de un nuevo derecho, en vez de fortalecer derechos existentes, mediante una legislación robusta en materias como la protección de datos personales o leyes antidiscriminación.

2. **No consultar con especialistas:** una cuestión particularmente alarmante del proyecto de ley de plataformas digitales es que prácticamente todas las personas con experiencia en materias regulatorias relativas a Internet en el país coinciden en que el proyecto tiene deficiencias importantes y su eventual aprobación crea más problemas de los que resuelve.

   Pero en vez de escuchar las críticas, los promotores del proyecto se han dedicado a descalificar a quienes las han emitido.[8] A algunas organizaciones, como Wikimedia Chile y Optia, no se les permitió presentar frente a la comisión, al mismo tiempo que se invitó a psicólogos, filósofos e ingenieros. Esto último no es necesariamente malo: hay una riqueza en la discusión interdisciplinar. Sin embargo, no tiene sentido marginar de la discusión a aquellas personas que efectivamente tienen conocimiento sobre el campo y experiencia regulatoria, especialmente si están levantado alertas respecto de cuestiones que no están bien resueltas en el proyecto.

   Es deber de los legisladores trabajar en favor de la mejor regulación posible y la aproximación de múltiples partes interesadas, con todos sus problemas, sigue siendo la mejor aproximación posible y la que garantiza un proceso más transparente y democrático.

3. **Regular a partir del sentido común o de percepciones sesgadas:** durante los últimos cinco años, la sombra de Cambridge Analytica ha eclipsado cualquier discusión sobre el impacto social de las redes sociales y usualmente se cita para ejemplificar el colapso de las democracias occidentales producto de nuestro uso de la tecnología. Sin embargo, no hay pruebas sólidas de que Cambridge Analytica haya jugado un rol relevante en la elección presidencial estadounidense de 2016 [9].

   De hecho, hay muchísimo que no sabemos respecto del impacto que estas tecnologías tienen a distintos niveles. El año pasado se armó gran revuelo cuando Frances Haugen, una ex empleada de Facebook, filtró a la prensa documentos de la compañía, incluyendo un estudio donde tres de cada diez niñas afirmaban que

8 "Atacan los aspectos más gravosos para las plataformas digitales del referido proyecto de ley, sosteniendo que serían exactamente lo contrario: cargas para los usuarios. Con ello, dan pie a la sospecha de que defienden los intereses de dichas plataformas como lobbistas disfrazados de activistas". Amunátegui, C., Cornejo, I., Donoso, L., Ramos, P., Walker, N. and Girardi, G., 2021. Respuesta a las críticas del proyecto de ley sobre plataformas digitales. [online] El Mercurio Legal. Disponible en: https://www.elmercurio.com/legal/Registro/Login.aspx?urlBack=/Legal/Noticias/Opinion/2021/10/20/respuesta-criticas-proyecto-plataformas-digitales.aspx.

Instagram las hacía sentirse peor consigo mismas. Pero como afirma Laurence Steinberg en una columna publicada en el New York Times [10], no solamente la investigación tenía serios problemas metodológicos, sino que en ningún caso puede considerarse como el fin de un proceso investigativo serio. Evidentemente el dato es interesante y sin embargo, a partir de él no es posible establecer causalidad entre el sentimiento de malestar y el uso de la plataforma. ¿Las adolescentes están deprimidas porque usan Instagram o el hecho de que estén deprimidas resulta en una experiencia poco sana a la hora de usar Instagram? Es difícil poder establecerlo sin mayor investigación y las empresas que administran las plataformas de Internet históricamente han sido extremadamente reticentes en dar acceso a la información requerida a los esfuerzos independientes de investigación.

A falta de certezas, nos quedamos con titulares que omiten la complejidad de los fenómenos y que pasan a formar parte del sentido común, inspirando luego políticas públicas. No hay que minimizar tampoco el rol de lo que Lee Vinsel ha denominado "crity-hype" [11], la tendencia a enfocarse en problemáticas rimbombantes y poco realistas asociadas a la tecnología, dejando de lado problemas reales, concretos y aburridos.

Es importante generar una distancia crítica que permita entender en qué medida las propuestas legislativas son producto de una preocupación sincera y cuánto hay de marketing político en ellas.

Finalmente, es necesario destacar que la discusión no pasa por regular o no regular. Nadie se opone a generar condiciones que garanticen un mayor desarrollo de los derechos fundamentales en los entornos digitales y muchas veces para ello es necesario crear políticas públicas. Pero hay que legislar bien: las consecuencias que puedan tener articulados bienintencionados, pero técnicamente deficientes pueden ser peores que los problemas que están intentando solucionar. ▪

## REFERENCIAS


[1] R. Foroohar, 2018. Year in a Word: Techlash. [online] Ft.com. Disponible en: https://www.ft.com/content/76578fba-fca1-11e8-ac00-57a2a826423e.

[2] Higson-Bliss, L., 2022. Online safety bill: ambiguous definitions of harm could threaten freedom of speech – instead of protecting it. [online] The Conversation. Disponible en: https://theconversation.com/online-safety-bill-ambiguous-definitions-of-harm-could-threaten-freedom-of-speech-instead-of-protecting-it-179514.

[3] Boletín 14561-19. Regula las plataformas digitales. Disponible en https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=15047&prmBOLETIN=14561-19.

[4] D. Charpentier, 2022. Sabag retirará proyecto sobre memes de Internet: "No me fijé lo que habían redactado mis asesores". [online] BioBio Chile. Disponible en: https://www.biobiochile.cl/noticias/2014/07/11/sabag-retirara-proyecto-sobre-memes-de-Internet-no-me-fije-lo-que-habian-redactado-mis-asesores.shtml.

[5] Informe anual de la Comisión Interamericana de Derechos Humanos 2013, volumen II. Informe de la Relatoría Especial para la Libertad de Expresión. p. 531. Disponible en: https://www.oas.org/es/cidh/docs/anual/2013/informes/LE2013-esp.pdf.

[6] G. Girardi. Cuenta personal de Twitter (29 de noviembre de 2021), Disponible en: https://twitter.com/guidogirardi/status/1465409626451685394.

[7] C. Yáñez, , 2021. Chile tendrá la primera ley del mundo que regula plataformas digitales como Facebook, Google, Instagram y Twitter. [online] La Tercera. Disponible en: https://www.latercera.com/que-pasa/noticia/chile-tendra-la-primera-ley-del-mundo-que-regula-plataformas-digitales-como-facebook-google-instagram-y-twitter/.

[8] Derechos Digitales, 2021. Sociedad civil internacional alerta sobre los peligros para el ejercicio de derechos de proyecto de ley de regulación de plataformas digitales presentado en Chile. [online] Disponible en: https://www.derechosdigitales.org/16879/sociedad-civil-internacional-alerta-sobre-los-peligros-para-el-ejercicio-de-derechos-de-proyecto-de-ley-de-regulacion-de-plataformas-digitales-presentado-en-chile/.

[9] E. Gibney, 2018. The scant science behind Cambridge Analytica's controversial marketing techniques. [online] Nature.com. Disponible en: https://www.nature.com/articles/d41586-018-03880-4.

[10] L. Steinberg., 2021. Opinion | Does Instagram Harm Girls? No One Actually Knows. [online] Nytimes.com. Disponible en: https://www.nytimes.com/2021/10/10/opinion/instagram-facebook-mental-health-study.html.

[11] L. Vinsel, 2021. You're Doing It Wrong: Notes on Criticism and Technology Hype. [online] Medium. Disponible en: https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5.

# La trivialidad de los neuroderechos[1]

1 Este trabajo es una versión traducida, actualizada y resumida del trabajo "Neurorights in Chile: Between neuroscience and legal science", en Hevia, Martín (ed.) Developments in Neuroethics and Bioethics: Regulating Neuroscience: Transnational Legal Challenges, Volumen 4, Elsevier-Academic Press, 165 -179.

**ALEJANDRA ZÚÑIGA FAJURI**

Doctora en Derecho por la Universidad Autónoma de Madrid. Profesora de la Facultad de Derecho de la Universidad de Valparaíso e investigadora del Centro de Investigaciones de Filosofía del Derecho y Derecho Penal de la misma institución.

alejandra.zuniga@uv.cl

**LUIS VILLAVICENCIO MIRANDA**

Doctor en Derecho por la Universidad Autónoma de Madrid. Profesor de la Facultad de Derecho de la Universidad de Valparaíso y director del Centro de Investigaciones de Filosofía del Derecho y Derecho Penal de la misma institución.

luis.villavicencio@uv.cl

**DANIELLE ZAROR MIRALLES**

Doctora en Derecho por la Universidad de Chile. Investigadora del Centro de Estudios en Derecho Informático de la Facultad de Derecho de la Universidad de Chile.

dzaror@derecho.uchile.cl

**RICARDO SALAS VENEGAS**

Doctor en Derecho por la Universidad de Chile. Profesor de la Facultad de Derecho de la Universidad de Valparaíso e investigador del Centro de Investigaciones de Filosofía del Derecho y Derecho Penal de la misma institución.

ricardo.salas@uv.cl

**RESUMEN.** Este trabajo analiza críticamente la noción y la protección de "neuroderechos" contenidos en dos proyectos de ley presentados recientemente en el Congreso Nacional. El objetivo es mostrar que, tras ciertas simplificaciones filosóficas, las iniciativas legales buscan resguardar, mediante mecanismos inadecuados e impertinentes desde el punto de vista de la racionalidad legislativa, persistentes y viejas amenazas a los derechos de las personas.

## Introducción

Diariamente entregamos a los gigantes tecnológicos datos para permitirles comprender y predecir nuestro comportamiento, pero ¿pueden *leer nuestra mente*? Dos proyectos de ley insinúan que tal lectura sería posible y que, por lo mismo, deberíamos proteger nuestra integridad mental creando los neuroderechos. El primer proyecto de ley busca la "protección de los neuroderechos y la integridad mental" (boletín 13.828-19), mientras que el segundo se cristalizó en la Ley 21.833 que agregó un nuevo inciso al artículo 19 Nº 1 de la Constitución para resguardar la "actividad cerebral" (boletín 13.827-19). La idea es promover por medio de este reconocimiento no solo un "nuevo derecho humano", sino impulsar su reconocimiento internacional.

Para efectos de este trabajo, vamos a entender por neurotecnologías un campo de la ciencia y la ingeniería en el que se exploran y desarrollan métodos que permiten interconectar el sistema nervioso con dispositivos [1]. Durante 2018, se hizo famoso el caso de David Mzee, una persona tetrapléjica que pudo volver a caminar distancias cortas con la ayuda de un andador ortopédico y gracias al implante de electrodos que revivieron

***Los proyectos de ley pretenden cautelar de nuevas amenazas a viejos derechos, generando un típico caso de redundancia normativa.***

su médula espinal. Identificar, en cambio, una definición estándar de neuroderechos es una tarea difícil. Lo máximo que podríamos decir es que se trataría de pretensiones destinadas a proteger a las personas de los eventuales peligros de las neurotecnologías.

En lo que sigue, analizaremos la noción de "neuroderechos". Nuestro propósito es mostrar que, camuflado tras algunas simplificaciones filosóficas, las iniciativas apuntadas buscan proteger a las personas de persistentes amenazas a sus derechos humanos por medios equivocados. Por falta de prolijidad, de la constatación de nuevas perspectivas en el campo de la ciencia y de la tecnología, se termina por inferir una conclusión jurídica equivocada.

## Avances tecnológicos en espera

Ambos proyectos de ley citan como fundamento investigaciones que, arguyen, estarían *ad-portas* de producir los efectos que las iniciativas pretenden neutralizar. Entre dichos avances se menciona el trabajo de Gallant quien indicó, el año 2011, que estaba cerca de reproducir las imágenes que una persona estaba observando [2]. Sin embargo, transcurridos diez años, el experimento no ha generado los resultados anunciados y el propio Gallant ha reconocido que mapear el cerebro está, por ahora, fuera de nuestro alcance. En la misma línea, los proyectos mencionan a la empresa Neuralink que anunció hace poco la posibilidad de cargar y descargar pensamientos. Este objetivo se lograría a través de un "encaje o cordón neural" que permitiría comunicarse con un computador sin una interfaz física. Aunque Neuralink previó tener su tecnología operativa para el año 2020, ello no ocurrió.

Como puede verse, los "hitos" tras la justificación de los proyectos de ley son de dudosa trascendencia y generan una falsa sensación de urgencia. Por otro lado, sus patrocinantes han señalado que buscan situar a Chile a la vanguardia internacional. Pero, ¿son este tipo de argumentos suficientes para legislar? La teoría legisprudencial nos entrega herramientas para evaluar el razonamiento del legislador de modo que la ciudadanía pueda indagar sobre los verdaderos motivos que determinan la creación de reglas y sus fines.

## Neuroderechos y racionalidad legislativa: razones para no legislar

El proyecto de reforma constitucional sobre neuroderechos se sustenta en la idea de que es necesario construir un nuevo derecho a fin de proteger la integridad física y psíquica de las personas de cualquier mecanismo tecnológico que pretenda "aumentar, disminuir o perturbar dicha integridad individual sin el debido consentimiento", es decir, proteger la libertad de conciencia y la privacidad mental de ciertas neurotecnologías. ¿Existen motivos suficientes para legislar? ¿Cumplen estos proyectos con los niveles mínimos exigibles desde la teoría de la legislación? Según Atienza una ley debe tener *racionalidad lingüística*, es decir, el texto debe ser capaz de transmitir un mensaje claro a sus *destinatarios* [3]. Para satisfacer este criterio, el uso correcto del lenguaje es crucial. El proyecto de neuroderechos usa términos tales como "continuidad psicológica y psíquica", "mente", "pensamientos", "sustratos mentales", entre otros. El problema de estos conceptos es que son tan oscuros como para hacer imposible que se cumpla este requisito. Desde ya, aun cuando el proyecto define lo que se va a entender por neurotecnologías, no explica en qué medida podríamos distinguir estos métodos de otro tipo de avances que, de manera habitual, observan, alteran y determinan nuestra psique. Desde la televisión a Internet, desde las terapias psicológicas y psiquiátricas hasta las tomografías, desde el implante coclear hasta las intervenciones quirúrgicas cerebrales, todos son instrumentos capaces de acechar o alterar aquello que el proyecto llama "identidad personal". El proyecto trabaja con conceptos metafísicos, discutidos desde hace siglos por la teología, la filosofía y, más recientemente, por las ciencias de la psicología y la psiquiatría. Ello evidencia que el aparato conceptual de la neurociencia es demasiado limitado todavía para ser regulado.

En segundo lugar, Atienza plantea que un texto legal debe tener *racionalidad jurídico-formal*, o sea, la nueva ley debe insertarse armónicamente en el sistema jurídico y, además, debe ser comprendida por los sujetos imperados, los sujetos destinatarios y por el sistema de justicia. Aquí lo primero que salta a la vista es que los proyectos de ley pretenden cautelar de nuevas amenazas a viejos derechos, generando un típico caso de redundancia normativa. El derecho a la privacidad, por ejemplo, cuyo origen puede remontarse a la Edad Media, permanece como tal sea que se lo amenazara en esos siglos cuando el párroco perforaba la pared para espiar a un monje, sea que se lo amenace en el siglo XXI cuando el Estado instala microcámaras o hace volar drones. Por ello la mayoría de los juristas y científicos que han reflexionado sobre estos problemas sostienen que los procedimientos que

pretenden la lectura de la mente están ya limitados por los clásicos derechos a la libertad de expresión, la libertad de pensamiento y el derecho a la privacidad, reconocidos en la legislación y en tratados internacionales [4].

Los proyectos no clarifican, además, quiénes son los sujetos imperados ni menos quiénes serían los órganos encargados de hacer cumplir sus reglas. No se cumple, por tanto, lo que Wintgens llama *densidad* normativa, es decir, hay falta de certeza sobre infracciones y sanciones [5]. Por último, los proyectos de ley se refieren a un "derecho a la protección de sesgos", lo que no sería más que una especificación del derecho a la no discriminación arbitraria, la que ya se encuentra mejor recogida en el proyecto de nueva ley de protección de datos personales a través de lo que se denomina el "derecho a la explicación".

En tercer lugar, Atienza se refiere a la *racionalidad pragmática*. En este nivel se demanda que la ley determine de qué forma toda la sociedad logra adecuarse a lo prescrito por la ley. Frente al tipo de realidad descrita en los proyectos de neuroderechos resulta difícil imaginar que se logre el objetivo apuntado por cuanto ese tipo de "invasiones a la privacidad mental" no existen. No hay cómo adecuar un comportamiento a una situación imaginaria. En la misma línea, tampoco se satisface lo que Waldron llama el *principio de legislación explícita* que exige que la ley pueda influir o alterar realmente la vida de las personas [6]. En términos de Wintgens, se incumple la exigencia de la *racionalidad contextual* ya que el uso de tecnologías cognitivamente invasivas, en el sentido descrito en los proyectos, no cuenta con evidencia que la respalde [7].

El término "lectura de la mente" es esencialmente metafórico. Aquello que llamamos "mente" comprende estados mentales como la imaginación, emociones, intenciones, percepción, toma de decisiones, etc. Con las tecnologías de

***El proyecto usa términos como "continuidad psicológica y psíquica" y "sustratos mentales" que son tan oscuros como para hacer imposible que se cumpla el requisito [de racionalidad lingüistica].***

interfaz cerebral, la neurociencia puede resaltar algunas correlaciones entre los estados mentales y la actividad cerebral. Pero, como afirman Rainey y otros, el acceso a alguna base material de los estados mentales es fragmentaria [8]. Los correlatos neuronales son incipientes huellas físicas de eso que llamamos "mente", datos tan fraccionados que no están todavía ni cerca de descifrar un pensamiento. Este hecho indesmentible ha sido reconocido, incluso, por Rafael Yuste, principal impulsor de los neuroderechos [9].

El doctor Edward Chang, quien ha llevado adelante la investigación más avanzada en la actualidad sobre interfaces para comprender cómo el cerebro controla nuestra capacidad para hablar, ha declarado que la técnica solo puede leer "señales" del habla [10,11]. Hasta ahora, no es posible saber lo que estamos pensando. Aun cuando pudiéramos distinguir las palabras que alguien intenta decir de las señales que emite su cerebro, ni siquiera estaríamos cerca de la lectura de la mente. La tecnología solo nos permite mirar las áreas que son relevantes para los aspectos motores de la producción del habla, no aquello que constituye un pensamiento. Lo cierto es que no sabemos ni siquiera conceptualmente *qué es y cómo se produce un pensamiento*.

Enfoquémonos ahora en los últimos dos requisitos para legislar que destaca Atienza: la *racionalidad teleológica* y la *racionalidad ética*. La primera se refiere a

***Hasta ahora, no es posible saber lo que estamos pensando. Aun cuando pudiéramos distinguir las palabras que alguien intenta decir de las señales que emite su cerebro, ni siquiera estaríamos cerca de la lectura de la mente.***

los fines sociales de la ley. La segunda, por su parte, juzga su justificación axiológica. A estas alturas debemos preguntarnos ¿será ético ocupar un tiempo escaso de la agenda legislativa sin justificación? Imagine que asignáramos un puntaje a estos proyectos de ley y los hiciéramos competir con otros que busquen mejorar, por ejemplo, la educación, la conectividad, el cuidado sanitario, el acceso a alimentación básica o la vivienda. Conforme al *principio de cuidado* [6] o el *principio de alternatividad* [5] que debe guiar la tarea del legislador, tenemos que demandarle que nos pruebe que las restricciones signifiquen mejoras. Cuando una ley es sencillamente superflua, erosiona la idea de justicia. Las leyes injustificadas no tienen un efecto inocuo. Al contrario, el sobregiro democrático merma la legitimidad de las instituciones representativas.

Tampoco se satisface el principio de representación. Como indica Waldron, el parlamento es un foro que debe dar voz a intereses públicamente relevantes [6]. ¿Existe en la sociedad una preocupación sobre los neuroderechos? ¿Qué problemas resuelve? Además, el proyecto de ley de reforma constitucional se aprobó en general de forma unánime en la Sala del Senado con 37 votos a favor y sin abstenciones. Algo similar ocurrió con el proyecto de ley: fue aprobado en general con 39 votos a favor, sin votos en contra ni abstenciones. Esta falta de debate no es una buena noticia. Se encuentra en entredicho lo que Waldron llama el *principio de deliberación* y el *principio de respecto del desacuerdo* [6], o sea, el deber que tienen los parlamentarios de debatir, reaccionando a las propuestas divergentes. Ahogar las diferencias bajo un consenso artificioso puede ilustrar la falta de una deliberación suficiente.

El análisis jurídico de los proyectos sobre neuroderechos enuncia la antigua novedad de que los derechos fundamentales están permanentemente expuestos a nuevas amenazas. Por cierto que aun si el Estado o una trasnacional dedicada a la neurotecnología lograra "leer" el pensamiento (técnicamente, los "datos neuronales") de una persona, tampoco estaría afectando un "nuevo derecho humano", sino el persistente *derecho a la privacidad*. Dicho de otro modo, el que surjan nuevas formas de matar no altera el contenido del derecho a la vida ni es fundamento para la creación de nuevos derechos. Por lo mismo, una consagración de estos neuroderechos no está justificada. La redundancia de los neuroderechos se observa al reconducirlos a derechos fundamentales que ya están asegurados en la actual Constitución, tratados internacionales sobre derechos humanos y demás legislación nacional (V. gr. Ley 19.628 sobre protección de la vida privada y de los datos sensibles y Ley 20.584 que regula los derechos y deberes del paciente).

## Un poco de filosofía

Si se analiza el proyecto desde una perspectiva filosófica se observa que quienes defienden los "neuroderechos" adscriben a una teoría *reduccionista* de la neurociencia cognitiva [12]. El reduccionismo nace de la antigua confusión cartesiana expresada en el dualismo mente/cuerpo, que hoy se ha reemplazado por otro dualismo erróneo: cerebro/cuerpo. Ambos reduccionismos comparten los mismos problemas conceptuales. El principal de ellos —denominado "falacia mereológica"— explica que la mente no es ni una sustancia idéntica ni distinta del cerebro y adscribir atributos psicológicos al cerebro es incoherente. El pensamiento y la sensación son *atributos del ser humano, no de su cerebro*, del todo, no solo de una parte. El ser humano es un animal autoconsciente que puede percibir, razonar, emocionarse y usar un lenguaje. No es un cerebro dentro del cráneo de un cuerpo. De modo que resulta una pretensión de corte "reduccionista cartesiana" argüir que es necesario construir nuevos derechos a fin de proteger *una parte específica del cuerpo humano,* el cerebro, pues en ella se encontraría la identidad.

En la línea del clásico experimento mental de Rorty [13], se puede afirmar que para que el cerebroscopio pueda interpretar un patrón particular de actividad neuronal necesita más que registrar la actividad de las neuronas en el presente. Tiene que haber estado acoplado a ese cerebro y cuerpo desde la fecundación para poder registrar toda su historia neuronal y hormonal. Entonces, y solo entonces, podría decodificar la información neuronal [8]. En el mismo sentido, supongamos que usted pone su firma en un documento. Aunque el acto de poner la firma va acompañado de disparos neuronales en su cerebro, esos disparos neuronales no *explican* lo que ha hecho. Al firmar con su nombre podría estar firmando un cheque o dando un autógrafo. En cada caso, el disparo neuronal es el mismo y, sin embargo, el *significado* de lo que ha hecho es diferente en cada caso y esas diferencias son "dependientes de las circunstancias", no solo el producto de disparos neuronales. Los disparos neuronales acompañan al acto de firmar, pero solo las circunstancias de la firma, incluyendo la intención de hacerlo, son los factores significativos para explicar lo hecho [14]. La neurociencia cognitiva solo es capaz de identificar de dónde provienen los "disparos" neuronales, pero no tiene más idea de su significado que el que tiene Google o Facebook cuando hacemos "Like".

## Conclusión

Regular la tecnología plantea desafíos en distintos niveles, especialmente cuando se encuentra en fases incipientes. La prudencia impone el deber de hacer las cosas con cuidado para evitar consecuencias indeseables como inhibir la investigación o entorpecer su financiamiento. Por otro lado, ni la Constitución ni las leyes parecen ser las fuentes jurídicas más adecuadas para este tipo de regulaciones. Dado el intenso dinamismo de la tecnología y su carácter global, parece más adecuado concentrarse en directrices técnicas. Una posibilidad óptima podría ser la dictación de normas sanitarias o tecnológicas, conocidas en algunos contextos como los *sandbox* regulatorios. Se trata de mecanismos para situaciones en progreso, acotados y flexibles, en términos que se pueden ir ajustando y revisando conforme al contexto. En Chile ya existen reglas constitucionales y legales que entregan respuestas adecuadas para los peligros que el uso de la neurotecnología pudiese ocasionar en el futuro. En definitiva, los problemas técnicos del proyecto de ley y la constitucionalización de los "neuroderechos" se funden con grandes preguntas todavía abiertas que reclaman, en conjunto, cierta modestia parlamentaria.

Finalmente, vale la pena citar a Anil Seth cuando arguye que "lo que significa ser 'yo' no se puede reducir ni subir a un programa. Somos animales biológicos de carne y hueso, cuyas experiencias conscientes se forman en todos los niveles mediante mecanismos biológicos que nos mantienen vivos. Simplemente hacer las computadoras más inteligentes no va a hacerlas sensibles" [15]. Ni tampoco, agregamos, les va a permitir leer nuestra mente. Por ello, al menos por ahora, se hace necesario renunciar a la pretensión reduccionista de legislar *para una parte de nuestro cuerpo*, porque no es cierto que solo en el cerebro se encuentra nuestra mente, identidad y consciencia. ▪

### REFERENCIAS


[1] Stieglitz, T. (2019). Why Neurotechnologies? About the Purposes, Opportunities and Limitations of Neurotechnologies in Clinical Applications. *Neuroethics*. https://doi.org/10.1007/s12152-019-09406-7.

[2] Gallant, J. "Brain Mapping, Brain Decoding, and Future Neurotechnology". https://www.youtube.com/watch?v=qekfk-lBgb8.

[3] Atienza, M. (1997). *Contribución a una teoría de la legislación*. Madrid: Civitas.

[4] Ligthart, S. (2020). Freedom of thought in Europe: Do advances in 'brain-reading' technology call for revision? *Journal of Law and the Biosciences*, 1-27. https://doi.org/10.1093/jlb/lsaa048.

[5] Wintgens, L. (2003). Legisprudencia como una nueva teoría de la legislación. *Doxa*, 26, 261-287.

[6] Waldron, J. (2006). "Principles of legislation". In R. Bauman & T. Kahana (Eds.), *The Least Examined Branch. The Role of Legislatures in the Constitutional State* (pp. 15-32). New York: Cambridge University Press.

[7] Wintgens, L. (2012). *Legisprudence. Practical Reason in Legislation*. Farnham: Ashgate.

[8] Rainey, S., Martin, S., Christen, A., Mégevand, P. & Fourneret, E. (2020). Brain Recording, Mind-Reading, and Neurotechnology: Ethical Issues from Consumer Devices to Brain-Based Speech Decoding. Sci Eng Ethics, 26, 2295-2311.

[9] Yuste, R. Proyecto BRAIN: "Mapear el cerebro es el mayor desafío de la ciencia". https://youtu.be/iVyTEu4FDvw.

[10] Anumanchipalli, G.K., Chartier, J. & Chang, E.F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature*, 568, 493-498.

[11] Chang, E. "Why computers won't be reading your mind any time soon", Wired. https://www.wired.co.uk/article/brain-computer-interfaces.

[12] Bennett, M.R. & Hacker, P.M.S. (2003). *Philosophical Foundations of Neuroscience*. Blackwell Publishing.

[13] Rorty, R. (1980). *Philosophy and the Mirror of Nature* (2 ed.), Princeton University Press.

[14] Patterson, D. (2003). Reviewed Bennett, M.R. & Hacker, P.M.S., *Philosophical Foundations of Neuroscience,* Blackwell Publishing. Disponible en: https://ndpr.nd.edu/news/philosophical-foundations-of-neuroscience/.

[15] Seth, A. "Cómo el cerebro alucina con la realidad", Futuro 360. https://www.futuro360.com/data/anil-seth-neurocientifico-como-el-cerebro-alucina-con-la-realidad_20190108/.

# Votación electrónica remota para la Universidad de Chile

**ALEJANDRO HEVIA**

Académico del Departamento Ciencias de la Computación de la Universidad de Chile. Coordinador Académico en Participa UChile, y director del Laboratorio de Criptografía Aplicada y Ciberseguridad, CLCERT, Universidad de Chile. Doctor en Ciencias de la Computación por la University of California, San Diego, Estados Unidos.

ahevia@dcc.uchile.cl

**CAMILO GÓMEZ**

Magíster en Ciencias de la Computación por la Universidad de Chile. Coordinador Operativo en Participa UChile.

cjgomez@uchile.cl

**CATALINA BURGOS KREITHER**

Diseñadora Gráfica por la Universidad del Pácífico. Diseñadora Gráfica en Participa UChile. En Twitter la encuentras como @catalina_bk.

**MARTA APABLAZA**

Licenciada en Historia por la Universidad de Chile y Periodista por la Universidad Católica. Encargada de comunicaciones en Participa UChile.

marta.apablaza@gmail.com

*Investigadores de la Facultad de Ciencias Físicas y Matemáticas desarrollaron un sistema de votación electrónica remota para ser utilizado por toda la comunidad universitaria, que puede tener un impacto de largo alcance en cómo se ejercita la democracia en pequeñas comunidades.*

La interrupción de nuestra vida cotidiana debido a la pandemia fue un evento que dejó grandes interrogantes respecto a cómo ejercemos ciertos hitos políticos de nuestra vida cotidiana como lo son elecciones y derecho a voto en espacios comunitarios. ¿Cómo crear y ejercer un sistema de votación electrónica de forma remota y segura? La respuesta a esta interrogante la está construyendo Participa UChile, en la forma de un nuevo sistema de votación electrónica remota universitaria, impulsado por la Prorrectoría de la Universidad de Chile, cuyo desarrollo está a cargo de un grupo de investigadores de la Facultad de Ciencias Físicas y Matemáticas (FCFM), y que cuenta con el apoyo técnico de la Vicerrectoría de Tecnologías de la Información (VTI).

Este nuevo sistema fue diseñado para ser utilizado en la mayoría de los procesos de votación al interior de la Universidad de Chile, tales como elecciones de decanos, directores de departamentos, centros de estudiantes y representantes de personal de colaboración, entre otras. Al ser un sistema de votación electrónica remota, los integrantes de la comunidad universitaria pueden sufragar desde un computador, un teléfono o una *tablet*, sin necesidad de recurrir físicamente a un recinto de votación. Cabe destacar que para ingresar al sistema de votación, las personas solo requieren tener activa su cuenta institucional (Pasaporte UChile).

El desarrollo de este proyecto comenzó en marzo de 2021 buscando expandir las opciones de participación dentro de la comunidad universitaria sin dejar de cumplir los niveles de seguridad y confiabilidad asociados a nuestra Universidad.

## ¿En qué consiste Participa UChile?

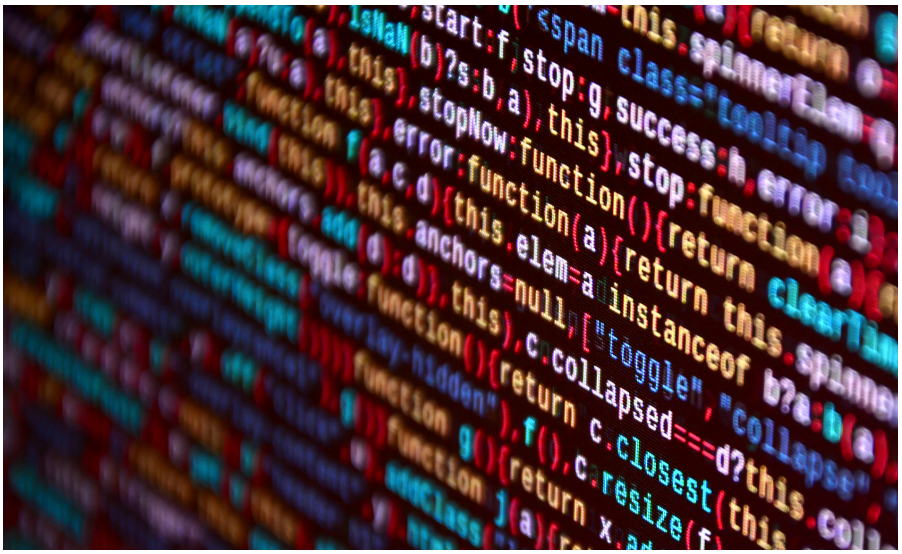
Participa UChile (https://participauchile.cl/) está basado en Helios, un sistema de votación electrónica remota y verificable para votaciones de bajo riesgo (https://vote.heliosvoting.org/). Fue creado en 2008 por Ben Adida, por entonces estudiante de doctorado en criptografía en MIT, quien lo desarrolló como un sistema de código abierto. Este acceso público no solo permite a quien quiera conocer cómo fue construido sino además tomar ese código, utilizarlo y modificarlo.

El profesor Alejandro Hevia, académico del Departamento de Ciencias de la Computación y director del Laboratorio de Criptografía Aplicada y Ciberseguridad (CLCERT) de la FCFM, lidera el equipo de desarrollo de Participa UChile. "Helios es un sistema diseñado y mejorado por expertos, que ha sido utilizado por importantes entidades como la Asociación Internacional de Investigación Criptográfica y por universidades como Princeton en Estados Unidos y la Universidad Católica de Lovaina en Bélgica", explica. Y agrega: "Su uso previo en un contexto académico, y el ser de código abierto, nos da la confianza que ha sido revisado, testeado y mejorado desde su creación. Es entonces un sistema probado que adaptamos para nuestro contexto universitario". Destaca que Helios utiliza criptografía, lo que permite garantizar el secreto del voto. En términos simples, esto significa que se emplean técnicas matemáticas que permiten "sellar" el voto, o en términos técnicos "cifrarlo", lo que impide que una persona distinta al votante pueda ver o modificar su contenido, incluso si se trata del administrador del sistema.

En el caso particular de Participa UChile, se agregaron también funcionalidades específicas requeridas por reglamento, como la posibilidad de emitir voto ponderado, o poder votar nulo o blanco.

*[Participa UChile] fue diseñado para ser utilizado en la mayoría de los procesos de votación al interior de la Universidad de Chile, tales como elecciones de decanos, directores de departamentos, centros de estudiantes y representantes de personal de colaboración.*

El profesor Alejandro Hevia indica que Participa UChile está diseñado —inicialmente— para llevar a cabo elecciones dentro de la comunidad universitaria. Esto, porque al tratarse de un ambiente con bajo riesgo de ataque informático se pueden practicar diferentes formas de elección e ir mejorando en seguridad y procesos. "En elecciones comunitarias los incentivos económicos o de otro tipo para perjudicar la elección son acotados y manejables". Sin embargo, agrega el académico, los desafíos son significativos incluso en este escenario. "Las elecciones comunitarias aun cuando sean pequeñas o medianas también requieren dar garantías de privacidad, integridad, transparencia y usabilidad", indica el profesor Hevia, al tiempo que clarifica su objetivo: "Participa UChile busca resolver el problema de las elecciones y consultas seguras en forma científica, paso a paso, construyendo sistemas que combinan la teoría y la práctica, desde situaciones de bajo riesgo, en forma sistemática, robusta, y participativa".



*Al basarse en el software Helios, una plataforma usada y testeada extensivamente, y ser de código abierto, Participa UChile entrega la posibilidad de auditar errores y realizar las mejoras necesarias en forma transparente.*

## Ventajas de una elección remota con fundamentos criptográficos

La utilización de técnicas criptográficas en Participa UChile permite entregar tres importantes garantías a sus usuarios: secreto del voto, mitigación de la coerción y realización de auditorías online.

Respecto al secreto del voto, en general, en otros sistemas de votación electrónica este depende en gran medida de la probidad de los administradores del sistema, quienes eventualmente podrían acceder al servidor y conocer quiénes votaron y por qué candidato. Participa UChile innova en este aspecto, ya que el secreto del voto depende de los llamados *custodios de clave*.

Este último concepto es un símil a los tradicionales vocales de mesa: son un número pequeño de personas, encargadas de realizar el conteo final de votos. Para esto, los custodios comparten una clave privada que crean especialmente para la votación y que tiene la particularidad que ningún custodio por sí solo puede abrir esta urna virtual.

El sistema está diseñado para "dividir" la clave en trozos, uno para cada custodio y, por lo tanto, requiere que los custodios ingresen su clave en forma conjunta. En otras palabras: "Esto es similar a una puerta con tres candados de seguridad donde la llave de cada candado la tienen tres personas distintas. Una sola persona puede abrir su candado, pero si el resto no lo hace la puerta seguirá cerrada. En nuestro ejemplo, necesitaríamos de las tres llaves para abrirla. Si bien esto es bueno, permitiría a una persona boicotear el proceso si se niega a abrirla. Por eso, nuestro sistema va más allá y permite 'abrir la puerta', esto es realizar el conteo, siempre que al menos dos de los tres los custodios participen. Eso agrega solidez al sistema", explica el profesor Hevia.

En la práctica, los custodios de clave deben ser designados (o posiblemente elegidos) en forma previa a cada elección.

Camilo Gómez, coordinador operativo de Participa UChile, profundiza: "Nuestra propuesta es que sean tres a cinco 'custodios de clave', según la elección. Si bien desde el punto de vista técnico no hay problema en designar una cantidad mayor, un número más amplio podría hacer más complejo coordinar su labor y eventualmente afectar las garantías del sistema. Con todo, esta es solo una sugerencia de operación, y la junta electoral es quien debiera tomar la decisión final".

Además de mantener el secreto del voto, Participa UChile mitiga la coerción, es decir, que alguien obligue al votante a marcar una determinada preferencia. La solución es permitir votar más de una vez, pero preservando la unicidad del voto. Si un votante decide cambiar su preferencia una vez emitido su voto puede ingresar a la plataforma y votar nuevamente, sobrescribiendo ese nuevo voto al anterior. Es decir, se puede reemitir el voto mientras el proceso esté abierto, pero solo se cuenta el último voto. "Esta estrategia es considerada por la comunidad científica como una mitigación razonable para el problema de la coerción", explica Alejandro Hevia.

La tercera garantía que otorga Participa UChile tiene que ver con la realización de auditorías online. El sistema permite que, mientras se están emitiendo los votos e, incluso, mientras el sistema está contándolos, se pueda verificar externamente que cada uno

de los pasos se ha ejecutado correctamente sin comprometer el secreto de los votos ni las claves de los custodios. El truco está en la existencia de algoritmos matemáticos que operan sobre valores públicos emitidos por el sistema, permitiendo "testear" que todos los votos válidos fueron considerados, y que los custodios operaron correctamente. Esta auditoría indirecta permite tener certeza sobre la integridad del cómputo final y, eventualmente, detectar si alguien intentó modificar el resultado.

## Roles en una elección

Al igual que en una elección presencial, donde existen vocales de mesa y encargados de recintos de votación, Participa UChile considera roles y actores definidos para cada elección: Junta Electoral, administrador del sistema, custodios de claves y votantes.

La Junta Electoral es el equivalente al Tricel del Servel para una elección específica. El reglamento universitario establece que debe existir una para toda elección; si bien su composición depende del tipo de elección, debe supervisar la correcta realización del proceso de votación según lo establecido en el reglamento, e informar sobre la fecha de votación, candidatos y universo de votantes, entre otras tareas.

El administrador de sistema es el encargado de la configuración de la elección de acuerdo con lo especificado por la Junta Electoral. Este rol lo ejercen los encargados de Participa UChile, actualmente el profesor Alejandro Hevia y Camilo Gómez.

"Los custodios, como se mencionó anteriormente, son los encargados de las claves y se designan o eligen en forma previa a una elección y solo para ella. Es importante destacar que los custodios de clave no tienen por qué ser los administradores. Un custodio es un equivalente al vocal de mesa, con la diferencia que en vez de estar presentes todo el proceso de elección, solo deben hacerlo en dos oportunidades: antes de abrir la votación para crear las claves, situación equivalente a constituir mesa. Y al cierre del proceso, para abrir la urna y realizar el conteo de votos", explica el profesor Hevia.

***La utilización de técnicas criptográficas en Participa UChile permite entregar tres importantes garantías a sus usuarios: secreto del voto, mitigación de la coerción y realización de auditorías online.***

## Riesgos y mitigación

Los investigadores a cargo del desarrollo de Participa UChile explican que todos los sistemas de votación electrónica remota presentan riesgos inherentes. Sin embargo, aclaran, también los mecanismos utilizados para mitigarlos en el caso particular del sistema de la Universidad de Chile.

Un riesgo ya mencionado tiene que ver con la coerción. Si bien en Participa UChile el riesgo se mitiga, no se elimina. Tal como explica Camilo Gómez, "el sistema no puede evitar que alguien te presione externamente a votar por una opción determinada. Tampoco previene la venta del voto. A nivel mundial, no existe sistema de votación remoto o presencial que lo prevenga completamente; es un problema que las democracias modernas están abordando". La estrategia de reemitir el voto, al menos por ahora, parece ser la opción más efectiva.

Un segundo riesgo tiene que ver con fallas provocadas por ataques informáticos, o lo que comúnmente se conoce como "*hackeos*". Por ejemplo, alguien

*Todos los sistemas de votación electrónica remota presentan riesgos inherentes.*

podría crear un virus que ataque el computador o tablet del votante, de modo que, si emite preferencia por el candidato A, el sistema por atrás "marque" el candidato B sin que el votante sepa. También está la posibilidad de ataques masivos que buscan interrumpir la conectividad de los votantes, tales como el corte de los cables de fibra óptica o ataques de denegación de servicios con los cuales se impide el acceso al servidor de votación. Todo esto podría impedir el acceso del votante a la plataforma de votación.

Otro ataque tiene que ver con *hackear* el servidor donde se almacenan los votos. Tal acción, si bien disruptiva, en Participa UChile tendría un efecto limitado. Gracias a que los votos están encriptados y las claves de los custodios no se almacenan en el servidor, el atacante no podría conocer por quién votó cada persona. "Entonces en términos del secreto del voto, este no se ve afectado, aunque sí podría obligar a realizar nuevamente todo el proceso de votación", indica Gomez.

Alejandro Hevia afirma que, para el ámbito de acción de Participa UChile, este tipo de ataques masivos no debieran ocurrir: "Crear un virus efectivo, por ejemplo, requiere de conocimiento y financiamiento. Nuestra apuesta es que, en un entorno comunitario como el universitario, la existencia de personas o grupos dispuestos a este esfuerzo o gasto es menos probable, versus en una elección de alto riesgo como la de presidente o parlamentarios".

Si bien estos son riesgos inherentes a todos los sistemas de votación electrónica, reitera que "en nuestro caso de uso, los incentivos no están alineados para que los participantes u observadores quieran llevar a cabo estos ataques masivos y toman preponderancia otros riesgos menores tales como coerción ocasional simple, inyección o manipulación de votos, o un custodio/administrador curioso, todos los cuales son cubiertos razonablemente en Participa UChile". En cualquier caso, expresa el profesor Alejandro Hevia, "la comunidad debe saber que los riesgos más severos, si bien no son esperables, efectivamente existen".

## Beneficios

Uno de los aspectos fundamentales de Participa UChile es entregar confianza frente a posibles riesgos. Para combatir las posibles fallas del sistema, Alejandro Hevia destaca el hecho que, al basarse en el software Helios, una plataforma usada y testeada extensivamente, y ser de código abierto, Participa UChile entrega la posibilidad de auditar errores y realizar las mejoras necesarias en forma transparente. "Esto es una medida de protección. El código está disponible para que cualquier persona pueda mirarlo y si tiene dudas pueda preguntarle a su amigo experto en programación u otra persona que sepa de criptografía, porque la idea es que el proyecto sea lo más transparente posible. Así, cualquier error o falla que alguien pueda encontrar, esperamos que nos la reporten para nosotros arreglarla. Esperamos con eso seguir el ejemplo de Helios, sobre el cual si bien se detectaron fallas, estas se corrigieron".

Otro beneficio de la votación electrónica entregada por Participa UChile, además de no requerir presencia física en una votación, es el potencial de especialización de los votos, por ejemplo: utilizar idiomas distintos o incluir en el voto preguntas de una mayor complejidad de lo que permite el voto en papel. "Se puede solicitar ordenar opciones según una preferencia. Si incorporamos tales mecanismos, que permitan capturar preferencias más complejas de la comunidad, tenemos el potencial de mejorar la representatividad y legitimidad de las elecciones o decisiones tomadas usando el sistema", indica.

Asimismo, los beneficios de desarrollar una votación electrónica y remota de manera segura se pueden extender más allá del tiempo y de las complejidades de los últimos años donde la cotidianidad se ha visto interrumpida por la pandemia: "Podemos comenzar a ejercitar una forma de votación y ejercer derechos totalmente nueva y podemos empezar a crear conocimiento y confianza en este tipo de sistemas", señala Hevia.

"Construir un sistema de votación electrónica con altos estándares de seguridad en el ámbito técnico y que cuente con la confianza de los electores es un proceso de largo aliento. En Participa UChile estamos invitando a la comunidad a participar para seguir creciendo en seguridad y transparencia en este sistema", finaliza Hevia. ▪

# Participación digital y plataformas en el contexto del proceso constituyente

**ARIEL BARRAZA**

Ingeniero Civil en Computación y estudiante de Magíster en Ciencias mención Computación de la Universidad de Chile.

ariel.barraza.r@gmail.com

## Introducción

El proceso constituyente que se está desarrollando actualmente en Chile, es la culminación de un largo y ascendente proceso de movilizaciones sociales que se han llevado a cabo desde el retorno a la democracia. La actual Constitución Política, que rige desde el año 1981, en plena dictadura militar, impuso un modelo de Estado cuyo único mecanismo para la realización de cambios estructurales, es la obtención de un *quorum* de dos tercios de la votación en ambas cámaras en el Congreso Nacional. Este marco normativo, impide la existencia de mecanismos a través de los cuales se pueda, por la vía institucional, proponer y/o canalizar los cambios exigidos por la ciudadanía, impidiendo al Estado poder otorgar respuesta a las problemáticas sociales, y detonando el estallido social que dio origen al proceso constituyente.

En este escenario de cambios, el uso de herramientas digitales para participación política (no obstante sus limitaciones), puede tener algunas ventajas en los procesos participativos. Por ejemplo, ayudar a la comunicación entre representantes y representados, entre personas alejadas geográficamente, o cuando hay un gran número de participantes; y es posible que su uso se vuelva más recurrente, ya sea por partidos políticos, organizaciones de la sociedad civil, o por instituciones, como ya lo está realizando la Convención Constitucional con su Plataforma Digital de Participación Popular (PDP).

En este artículo se abordan, en el contexto del proceso constituyente, algunos elementos relacionados con la participación digital a través de plataformas computacionales. El contenido fue elaborado a partir del estudio de algunas plataformas digitales de participación política, y a través de elementos que estuvieron presentes en la discusión de las primeras reuniones de concepción de la PDP de la Convención Constitucional, y se espera que otorguen al lector una idea del estado del arte en esta materia.

***Una buena metodología [de un proceso participativo] debería asegurar a las personas participar en igualdad de condiciones, y que haya confianza y aceptación del resultado final.***

## Metodologías de participación en procesos políticos

Una manera simple de definir participación política, es la implementación de mecanismos a través de los cuales las personas, de manera individual o colectiva, puedan manifestar su opinión ante las personas y/o instituciones responsables de llevar a cabo procesos políticos.[1] Todo proceso participativo requiere del uso de una metodología que, implícita o explícitamente, será consecuencia del modelo de democracia elegido por las personas encargadas de su diseño [1]. La metodología tendrá, por lo tanto, decisivas consecuencias en la manera en que se lleve a cabo el proceso, y determinará el diseño de los sistemas digitales utilizados para el desarrollo del mismo [2].

En los modelos de democracia predominantes hoy en día, es de amplio consenso la idea de que en el ejercicio de esta, se asegure que todas las personas tengan la oportunidad de participar en igualdad de condiciones. Además, los procedimientos adoptados deben conducir a que las decisiones del proceso deliberativo sean reflejo de la opinión mayoritaria de las y los participantes, sin que se favorezca a ninguna persona o grupo por sobre otro. Algunos criterios para ello son los definidos por Robert Dahl y explicados en [3]: *participación efectiva*, *igualdad del voto*, *comprensión ilustrada*, *control de la agenda e inclusividad*. Estos criterios son un buen piso para asegurar mínimos democráticos, aunque pudieran no ser suficientes para abarcar todos los aspectos que ha adquirido este asunto con el pasar de los años.

En el caso de la Convención Constitucional, se ha realizado un esfuerzo por reconocer la diversidad de visiones existentes en la sociedad. Esto ha quedado plasmado en los reglamentos de participación popular y de consulta indígena, destacando dos aspectos: la consulta indígena y la inclusión de grupos históricamente excluidos (personas mayores, en situación de discapacidad, niños, niñas y adolescentes, etc.). Estos reglamentos consideran mecanismos específicos de participación para cada uno de estos grupos.

Además, existen sectores organizados en torno a problemáticas específicas: mujeres, disidencias sexuales, grupos medioambientalistas, etc., que han participado colectivamente en distintas

1 Notar que estamos hablando de participación de manera elemental, en la cual las personas o instituciones que ostentan la representatividad popular tienen la última palabra en la toma de decisiones políticas, consistentemente con lo que ocurre en la Convención Constitucional. Por ello, en estos casos no se tienen garantías de incidencia en el proceso deliberativo, ni mucho menos el resultado es vinculante.

instancias del proceso constituyente. Lo anterior es signo de que las personas no solo participan individualmente, sino que también lo hacen de manera organizada, respondiendo a intereses específicos. Y aunque la tendencia parece ir en esa dirección, aún es difuso cómo esto incidirá en el resultado final del proceso.

El diseño metodológico que se adopte en un proceso participativo es, por lo tanto, un asunto delicado, pues tiene impacto en todos los aspectos del proceso, pudiendo incluso incidir o determinar su resultado. Una buena metodología debería asegurar a las personas participar en igualdad de condiciones, y que haya confianza y aceptación del resultado final.

## Los sistemas digitales deben garantizar confianza en el proceso

Obtener la validación ciudadana en los procesos participativos, es una condición necesaria para su éxito. Desde el punto de vista del uso de sistemas computacionales, existen cuestiones importantes a considerar en los procesos de diseño e implementación, pues cualquier falla u omisión puede comprometer la confianza de las y los participantes y, en el peor de los casos, provocar el fracaso del proceso. Algunos de estos aspectos son:

- **Máxima transparencia:** los procesos participativos podrían tener incidencia en la toma de decisiones de mucha relevancia para la vida de las personas que habitan un territorio. Por ello, es vital asegurar la transparencia en los mecanismos, procedimientos y herramientas (digitales y no digitales) utilizadas para llegar a un determinado resultado, ya que estarán bajo permanente escrutinio público.

- **Seguridad:** los sistemas deben estar protegidos ante posibles ataques de terceras personas, puesto que una eventual vulneración comprometería la credibilidad de los resultados. Es importante tener en cuenta que este tipo de sistemas podría ser potencialmente blanco de ataques, debido a la naturaleza de los procesos para los cuales son construidos.

- **Trazabilidad:** contar con mecanismos que permitan el control de las interacciones de personas con los sistemas, asegurando que cualquier intervención humana sea de acuerdo con las normas y permisos establecidos, siempre resguardando que los resultados no sean alterados.

- **Privacidad:** asegurar que los datos de las personas estén debidamente resguardados, protegiendolas de posibles exposiciones o consecuencias negativas asociadas a su participación, tal como ocurre con el secreto del voto.

- **Resiliencia:** asegurar que la infraestructura computacional funcione de manera aceptable ante eventuales fallas o momentos de alta demanda, ya que problemas de disponibilidad podrían tener impacto negativo en la participación efectiva de todas y todos.

La utilización de sistemas digitales, por tanto, va más allá de la implementación de mecanismos de participación, pues requiere del cumplimiento de requisitos no funcionales, críticos para un uso exitoso. Estos aspectos suelen ser de los más débiles en las plataformas digitales

***La utilización de sistemas digitales va más allá de la implementación de mecanismos de participación, pues requiere del cumplimiento de requisitos no funcionales, críticos para un uso exitoso.***

de participación existentes, impidiendo que sean aptas para llevar a cabo procesos de mayor envergadura.

## Algunas dimensiones de los procesos participativos

Llevar a cabo procesos democráticos es un asunto que contempla múltiples dimensiones, cada una de las cuales requiere, por sí sola, la búsqueda de soluciones que permitan implementarlos de manera satisfactoria. A continuación mencionamos algunas de ellas, más o menos en el orden en que se realizarían en una situación con lógica institucional.

1. **Información:** poner a disposición de las personas los elementos necesarios para que puedan participar (idealmente) con el mismo nivel de conocimiento sobre los asuntos que se tratarán en el diálogo democrático. La información debiese ser, indistintamente del perfil de las personas que participen (nivel socioeconómico, ubicación geográfica, educación, edad, situación de discapacidad, etc.), accesible, comprensible, oportuna y ordenada.
2. **Consulta:** realizar distintos tipos de sondeos que permitan conocer la opinión de las y los participantes a través de la respuesta de una o más preguntas. El formato para responder las preguntas suele ser simple, pues esto ayuda a una rápida obtención de los resultados, siendo la votación la forma más sencilla de realizarlos.
3. **Participación:** las personas pueden trabajar, de manera individual o colectiva, en algunas de las actividades del proceso democrático. Algunos ejemplos: presentación de propuestas, organización de reuniones, corredacción de textos, etc.
4. **Deliberación:** mecanismos que ayuden a la toma de decisiones a partir de los datos emanados de los procesos participativos. Esto es, que el resultado sea reflejo de todas las visiones presentes en el debate, "razonablemente" ponderadas según el peso de las partes involucradas. Este proceso es especialmente delicado, porque requiere que todas las personas (especialmente quienes pierden) estén convencidas de que el resultado es reflejo del diálogo democrático, sin que necesariamente una mayoría se imponga a una minoría, como suele ocurrir con las votaciones.
5. **Sistematización:** estructurar la información emanada de los procesos participativos masivos para ser comprendida, analizada y utilizada por el ser humano en otras etapas de estos.
6. **Seguimiento:** llevar la trazabilidad de los resultados de los procesos participativos. Si bien las instancias de consulta ciudadana se han tornado más frecuentes, es difícil obtener información posterior sobre qué ocurrió con el resultado de estas instancias. Idealmente, una persona o grupo debería poder averiguar dónde y cómo fueron considerados los resultados de su instancia de participación, y el estado final del proceso.
7. **Visualizaciones:** poder explorar la información directamente, desde múltiples dimensiones, sin mediar otro tipo de documentos como aquellos que se elaboran en el proceso de sistematización.

Desde el punto de vista computacional, resolver cada una de estas dimensiones presenta un desafío en sí. Por ejemplo, si bien llevar a cabo una votación electrónica es una cuestión aparentemente simple, realizarlo de manera seria, requiere cumplir con criterios complejos: integridad, secreto, autenticación, derecho a voto y disponibilidad [4]. Los procesos de deliberación y sistematización también se complejizan cuando existe un gran número de participantes, pues requiere manejar grandes volúmenes de información no necesariamente estructurada [5].

Estas dimensiones, lejos de estar estandarizadas, dependen fuertemente de la metodología empleada y de la naturaleza del proceso en cuestión. Por ello, en general, los sistemas existentes tienden a ser hechos a la medida de los procesos para los cuales fueron concebidos. Hasta el momento no está claro si es posible estandarizar la participación digital, o alguna parte de ella.

## ¿Y qué pueden hacer las plataformas digitales de participación existentes?

En diversas partes del mundo se están utilizando plataformas digitales para llevar a cabo procesos de participación política.[2] Estos sistemas han sido utilizados por instituciones, partidos políticos y

2 Se puede acceder a un catálogo con algunas de ellas en el contexto de legislación colaborativa en https://catalog.crowd.law/.

organizaciones de la sociedad civil, principalmente con propósitos de consulta y participación en asuntos particulares del lugar donde han sido construidos.

El foco principal de estos sistemas es la recopilación de la opinión de un grupo de participantes, sobre un determinado asunto (*input*). Para ello, se implementan mecanismos consultivos y/o participativos tales como encuestas, votaciones, foros de debate, propuestas, corredacción de documentos, entre otros. Si se trata de un mecanismo directo (votación, encuestas), se suele publicar el resultado. En otros casos (debates, ideas, corredacción de documentos), no siempre está claro cuál es el resultado final, principalmente por dos razones: el proceso abandona la vía digital y sigue su curso de manera "tradicional"; y el procesamiento de la información, al ser más complejo, requiere de la intervención del ser humano. Por ello, si bien estos sistemas permiten obtener la opinión de las personas, no suelen incorporar el resto de las etapas y/o dimensiones de los procesos participativos. La Tabla 1 contiene algunos ejemplos de estos sistemas.

Entre las plataformas mencionadas en la Tabla 1, destacan Consul y Decidim. Ambas, al haber sido desarrolladas para llevar a cabo gobiernos locales participativos, poseen una mayor cantidad de mecanismos, funcionalidades y configuraciones para su uso. Entre estas dos, la segunda destaca por tener mayor flexibilidad en las formas de uso que se le puede otorgar.

Decidim es una plataforma de código abierto, creada por el Ayuntamiento de Barcelona, que ha sido implementada en varias situaciones en distintos países. Su principal característica es poseer un diseño modular, que permite llevar a cabo distintos tipos de procesos partici-

| Plataforma | Creador | País | Descripción | Sitio web |
|---|---|---|---|---|
| Your Priorities | Fundación Citizens | Islandia | Fue desarrollada en el contexto del proceso de reforma constitucional islandesa (2019). Quienes participan pueden proponer ideas para que otras personas argumenten a favor y/o en contra. | https://citizens.is/ |
| Consultas Digitales | ONG Democracia en Red | Argentina | Aplicación utilizada por el Gobierno de Argentina para realizar consultas ciudadanas. Una consulta se compone de una o más preguntas de tipo comentario, votación, jerarquía, por rango, etc. | https://democraciaos.org/es/consultadigital/ |
| Portal Leyes Abiertas | ONG Democracia en Red | Argentina | Aplicación utilizada por la Cámara de Diputados de Argentina para someter a consulta proyectos de ley. Se publican los borradores de algunos textos, para que la ciudadanía realice comentarios sobre estos. | https://democraciaos.org/es/coconstruccionleyes/ |
| Loomio | Iniciativa ciudadana | Nueva Zelanda | Plataforma donde grupos de personas pueden realizar discusiones en hilos deliberativos. La idea es que, dentro de un hilo, se puede argumentar con apoyo de elementos multimedia, hasta llegar a un acuerdo. | https://www.loomio.com/ |
| Consul | Ayuntamiento de Madrid | España | Plataforma desarrollada para llevar a cabo procesos participativos digitales en la ciudad de Madrid. El principal flujo participativo consta de tres etapas: debate previo, propuesta de ideas, comentarios de textos. | https://consulproject.org/ |
| Decidim | Ayuntamiento de Barcelona | España | Plataforma desarrollada para llevar a cabo procesos participativos digitales en la ciudad de Barcelona. Los procesos participativos se organizan en fases, a las que se les pueden asociar mecanismos de participación específicos. Esto último le da un grado de versatilidad que no tienen las otras plataformas. | https://decidim.org/ |

**Tabla 1.** Ejemplos de plataformas desarrolladas para realizar participación digital.

pativos. Estos procesos se organizan en fases, las cuales a su vez tienen asociados mecanismos de participación: debates, propuestas, encuentros, seguimiento de metas, entre otros. Esta lógica de diseño le otorga una gran versatilidad, que incluso le permitió, ser evaluada como alternativa para la implementación de la plataforma de participación de la Convención Constitucional.

Por contraparte, la plataforma aún tiene muchos aspectos que mejorar. Mencionaremos dos de ellos:

1. **Usabilidad:** en el sitio administrativo, la aplicación no es intuitiva y es difícil de configurar.

2. **Robustez:** existen muchas configuraciones que quiebran el funcionamiento de la plataforma. Tampoco se retroalimenta al usuario cuando esto ocurre.

Dedicim, en general, es un buen punto de partida para implementar participación digital, pero una implementación seria, requeriría de una cuidadosa revisión y la realización de algunos ajustes para que realice lo esperado. Pero, a pesar de los contras, la plataforma tiene un nivel de desarrollo superior a las demás, por lo que es muy recomendable su revisión por quienes estén estudiando o deseen implementar participación digital.

**Figura 1.** Comparación de interfaces para firmar una IPN: arriba la PDP, abajo Decidim.

## Participación digital en la Convención Constitucional

La Convención Constitucional ha puesto a disposición de la ciudadanía la posibilidad de participar en el proceso constituyente, a través de su Plataforma Digital de Participación Popular (PDP). Los dos principales mecanismos que han sido implementados son[3]:

1. **Iniciativa Popular de Norma (IPN):** una persona puede presentar una propuesta de norma constitucional, que puede ser apoyada a través de la plataforma. Las propuesta que reúne 15.000 apoyos pasan a votación en la comisión pertinente.

2. **Cabildos:** realización de encuentros de discusión sobre temas constituyentes, con el fin de proveer insumos para la Convención.

Ambos tipos de participación, no obstante los detalles, están presentes en las plataformas mencionadas en la Tabla 1, particularmente en Decidim. En la lógica de participación a escala local,

3 Adicionalmente, en la plataforma están presentes otros tres mecanismos: iniciativa popular de norma de pueblos originarios, consulta indígena, y un buscador de norma. Los dos primeros, son de participación exclusiva de personas pertenecientes a pueblos originarios, mientras que el tercero, consiste en la publicación de las normas aprobadas por el pleno de la Convención Constitucional para ser parte del nuevo texto constitucional.

la Iniciativa Popular de Norma, es una variante de la presentación de propuestas ciudadanas, y los Cabildos, de los encuentros. La Figura 1 ilustra esta similitud para el caso de las IPN.

Una observación importante, es que mientras que en la IPN participaron 982.332 personas, hasta la fecha solo se han registrado 324 cabildos. Y, aunque las razones precisas serán materia de estudio posterior, existen al menos dos que podrían explicar esta situación. La primera, es que la información sobre este mecanismo participativo no llegó de la misma manera a toda la población y, la segunda, es que a diferencia de las IPN, no se tiene claridad en cómo incidirán los cabildos, cuando buena parte de las normas ya han sido aprobadas.

## Para finalizar

La participación digital ha sido la principal manera en que la ciudadanía ha podido incidir en el proceso constituyente en curso. Este proceso, inédito en la historia de Chile, se ha implementado con casi nula experiencia sobre cómo llevar a la práctica procesos de participación masiva y en un contexto de urgencia temporal que agregó dificultad a la tarea. No obstante a ello, la masividad de la participación en la Iniciativa Popular de Norma, desde donde emergieron algunas normas que estarán presente en la propuesta final, dan cuenta de un proceso que, a pesar de las limitaciones, ha cumplido con su cometido.

***Los sistemas existentes tienden a ser hechos a la medida de los procesos para los cuales fueron concebidos. Hasta el momento no está claro si es posible estandarizar la participación digital.***

Implementar participación digital es un asunto complejo, que requiere un cuidadoso diseño metodológico para garantizar la neutralidad del proceso, un sólido diseño que asegure que los sistemas no sean vulnerados, y resolver los problemas de cada una de las dimensiones del problema de participación. Esto último, presenta desafíos en casi todas las áreas de la computación.

A futuro, sin duda, la participación digital se hará más frecuente, posiblemente a través de gobiernos locales participativos, consultas ciudadanas, iniciativa popular de ley, etc. Cuando ello ocurra, un buen entendimiento del problema y la experiencia obtenida de este proceso, serán determinantes para un resultado exitoso. ▪

**Agradecimientos:**
*Quisiera agradecer a Claudio Gutiérrez por sus comentarios y sugerencias varias respecto a la redacción del artículo.*

### REFERENCIAS

[1] Lindner Ralf y Georg Aichholzer: *E-Democracy: Conceptual Foundations and Recent Trends*. Enero 2020, ISBN 978-3-030-27183-1.

[2] Pianini Danilo y Andrea Omicini: *Democratic Process and Digital Platforms: An Engineering Perspective: An Interdisciplinary Approach*. Enero 2019, ISBN 978-3-030-05332-1.

[3] Gutiérrez Claudio: *La tentación tecnológica.* Revista Bits de Ciencia N°19, páginas 3-7. 2020. https://www.dcc.uchile.cl/Bitsdeciencia19.pdf.

[4] Hevia Alejandro: *Votación electrónica y democracia.* Revista Bits de Ciencia N°19, páginas 23-33. 2020. https://www.dcc.uchile.cl/Bitsdeciencia19.pdf.

[5] Hilbert, Martin: *The Maturing Concept of E-Democracy: From E-Voting and Online Consultations to Democratic Value Out of Jumbled Online Chatter*. Journal of Information Technology & Politics, 6(2):87–110, 2009. https://doi.org/10.1080/19331680802715 242.

Técnicas formales de privacidad de datos:

# ¿Está el Servel protegiendo nuestra privacidad?

**MATÍAS TORO**

Doctor en Ciencias Mención Computación por la Universidad de Chile. Posdoctorante del Departamento de Ciencias de la Computación de la misma Universidad. Líneas de investigación: lenguajes de programación, *type-and-effect systems, gradual typing, security typing* y privacidad diferencial.

mtoro@dcc.uchile.cl

Nuestros datos personales están cada día repartidos en más y más aplicaciones o sistemas. Estos datos pueden (o no) contener información sensible o confidencial, como por ejemplo, información médica, financiera, de citas, o de preferencias artísticas. Esto genera un gran problema: el uso de los datos podría filtrar parte de esta información sensible. Informalmente, evitar que estos datos sensibles se hagan públicos o conocidos por terceros se llama *privacidad de publicación de datos*, que por simplicidad, llamaremos *privacidad*. En particular, en un análisis sobre una base de datos, se dice que la privacidad de un individuo es violada, si se aprende algo nuevo al agregar el individuo a la base de datos.

> Decimos que un análisis de datos *preserva la privacidad* si se aprende algo "útil" del análisis y al mismo tiempo no se viola la privacidad de ningún individuo.

Existen varias técnicas para preservar la privacidad, y dentro de las más importantes se encuentran la *anonimización* y la *privacidad diferencial*. En este artículo veremos un ejemplo práctico de violación de privacidad y propondremos cómo intentar revertirla usando técnicas de anonimización.

## El caso del plebiscito

Recientemente, el Servicio Electoral (Servel) dispuso de un listado público con información acerca de la votación del plebiscito constitucional [1]. Cada elemento del listado corresponde a una persona, donde se indican algunos datos de ella como por ejemplo, su país de nacimiento, sexo, comuna, edad, y si votó o no votó en el plebiscito (llamado "votaron"). El *dataset* tiene 14.855.719 filas y cuenta con la información de todo el universo votante (ver Figura 1).

| | Cédula | Circunscripción | Comuna | DV | Edad | Nacionalidad | País Domicilio | País Nacimiento | Partido | Provincia | Rango Edad | Región | Sexo | Sufragio | Voto Exterior | N° de registros | Votaron |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | Coyhaique | Coyhaique | 0 | 22 | chilena | Chile | Chile | [130] Federación Regionalista Verde Social | Coyhaique | 20-24 | De Aysén del General Carlos Ibáñez del Campo | F | sufragó | Nacional | 1 | 1.0 |
| 1 | 0 | El Puerto | Valparaíso | 0 | 89 | chilena | Chile | Chile | Sin Partido | Valparaíso | 80 o + | De Valparaíso | F | no sufragó | Nacional | 1 | NaN |
| 2 | 0 | Iquique | Iquique | 0 | 99 | chilena | Chile | Chile | Sin Partido | Iquique | 80 o + | De Tarapacá | M | no sufragó | Nacional | 1 | NaN |
| 3 | 0 | Río Tranquilo | Rio Ibáñez | 0 | 22 | chilena | Chile | Chile | Sin Partido | General Carrera | 20-24 | De Aysén del General Carlos Ibáñez del Campo | M | sufragó | Nacional | 1 | 1.0 |
| 4 | 0 | El Puerto | Valparaíso | 0 | 95 | chilena | Chile | Chile | Sin Partido | Valparaíso | 80 o + | De Valparaíso | M | no sufragó | Nacional | 1 | NaN |

**Figura 1.** Datos de participación electoral del plebiscito 2020.

df[(df['Pais Nacimiento'] == 'Irlanda') & (df['Sexo'] == 'masculino') & (df['Comuna'] == 'Santiago') & (df['Edad'] >= 30) & (df['Edad'] <= 39)]

| | Cédula | Circunscripción | Comuna | DV | Edad | Nacionalidad | País Domicilio | País Nacimiento | Partido | Provincia | Rango Edad | Región | Sexo | Sufragio | Voto Exterior | N° de registros | Votaron |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1009708 | 0 | Parque Almagro | Santiago | 0 | 35 | extranjero | Chile | Irlanda | Sin Partido | Santiago | 35-39 | Metropolitana de Santiago | M | sufragó | Nacional | 1 | 1.0 |
| 1766909 | 0 | El Centro | Santiago | 0 | 36 | extranjero | Chile | Irlanda | Sin Partido | Santiago | 35-39 | Metropolitana de Santiago | M | no sufragó | Nacional | 1 | NaN |
| 14046592 | 0 | El Centro | Santiago | 0 | 34 | extranjero | Chile | Irlanda | Sin Partido | Santiago | 30-34 | Metropolitana de Santiago | M | no sufragó | Nacional | 1 | NaN |

**Figura 2.** Solo tres personas calzan con los criterios del filtro.

df.groupby(['Edad', 'Pais Nacimiento', 'Sexo' , 'Comuna']).filter(lambda x: x['Número de registros'].count() == 1)

| | Cédula | Circunscripción | Comuna | DV | Edad | Nacionalidad | País Domicilio | País Nacimiento | Partido | Provincia | Rango Edad | Región | Sexo | Sufragio | Voto Exterior | N° de registros | Votaron |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 48 | 0 | Chanco | Chanco | 0 | 95 | chilena | Chile | Chile | Sin Partido | Cauquenes | 80 o + | Del Maule | M | no sufragó | Nacional | 1 | NaN |
| 133 | 0 | Santa Bárbara | Santa Bárbara | 0 | 115 | chilena | Chile | Chile | Sin Partido | Biobío | 80 o + | Del Biobío | F | no sufragó | Nacional | 1 | NaN |
| 194 | 0 | Arica | Arica | 0 | 133 | chilena | Chile | Chile | Sin Partido | Arica | 80 o + | De Arica y Parinacota | M | no sufragó | Nacional | 1 | NaN |
| 209 | 0 | Aysén | Aysén | 0 | 61 | extranjero | Chile | Japón | Sin Partido | Aysén | 60-64 | De Aysén del General Carlos Ibáñez del Campo | M | no sufragó | Nacional | 1 | NaN |
| 212 | 0 | San Joaquín | San Joaquín | 0 | 67 | extranjero | Chile | Argentina | Sin Partido | Santiago | 65-69 | Metropolitana de Santiago | M | no sufragó | Nacional | 1 | NaN |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 14840749 | 0 | 0 | 0 | 0 | 79 | extranjero | Chile | Uruguay | Sin Partido | Santiago | 75-79 | Metropolitana de Santiago | M | sufragó | Nacional | 1 | 1.0 |
| 14840751 | 0 | Isla de Maipo | Isla de Maipo | 0 | 81 | extranjero | Chile | Estados Unidos | Sin Partido | Talagante | 80 o + | Metropolitana de Santiago | M | sufragó | Nacional | 1 | 1.0 |
| 14842069 | 0 | Renaico | Renaico | 0 | 97 | chilena | Chile | Chile | Sin Partido | Malleco | 80 o + | De la Araucanía | M | no sufragó | Nacional | 1 | NaN |
| 14842084 | 0 | Quirihue | Quirihue | 0 | 99 | chilena | Chile | Chile | Sin Partido | Itata | 80 o + | De Ñuble | M | no sufragó | Nacional | 1 | NaN |
| 14851455 | 0 | O'Higgins | O'Higgins | 0 | 21 | chilena | Chile | Chile | Sin Partido | Capitán Prat | 20-24 | De Aysén del General Carlos Ibáñez del Campo | M | sufragó | Nacional | 1 | 1.0 |

*65585 rows × 17 columns*

**Figura 3.** Hay 65.585 grupos de edad, país de nacimiento, sexo y comuna con un solo elemento.

Con la intención de proteger la privacidad de cada persona, las columnas **rut** y **nombre** fueron eliminadas/ofuscadas del listado público. Estas columnas son llamadas *identificadoras* ya que identifican únicamente a cualquier individuo. Las columnas país de nacimiento, sexo, comuna y edad, las llamaremos *cuasi-identificadores* (porque "casi" identifican a un individuo), y la columna "votaron" la llamaremos *atributo sensible* ya que es un valor que se desea proteger.

> Una columna se llama *identificadora* si permite identificar unívocamente a cualquier individuo. Se llama *cuasi-identificadora* si "casi" permite identificar a un individuo.

Lamentablemente, eliminar las columnas rut y nombre no es suficiente para proteger la privacidad de **todos** los individuos; podemos *re-identificar* a ciertos individuos usando *datos auxiliares*. Con datos auxiliares, nos referimos a un listado de datos que contienen las columnas identificadoras y algunas cuasi-identificadoras. Esta información es usada comúnmente por atacantes que buscan asociar columnas identificadoras a datos sensibles. A continuación mostraremos lo que se llama *ataque de asociación de registros* para averiguar si un amigo nuestro votó o no votó en el plebiscito.

Sabemos que nuestro amigo es irlandés, su sexo es masculino, vive en Santiago y en el 2019 tenía entre 30 y 39 años. Si filtramos el listado usando estos datos, podemos ver que solo tres personas cumplen con esas condiciones (ver Figura 2).

Adicionalmente, podemos aprender que hay un 33,33% de probabilidad de que nuestro amigo haya votado en el plebiscito. Peor aún, si conseguimos su edad real en el 2019 (35), podemos saber con 100% de probabilidad que nuestro amigo sí votó en el plebiscito.

De hecho, si agrupamos los datos por los cuasi-identificadores, es decir, contamos cuánta gente hay por cada conjunto de valores de atributos distintos, podemos ver que hay 65.585 individuos que potencialmente podemos re-identificar unívocamente (ver Figura 3).

df.groupby(['Edad', 'Pais Nacimiento', 'Sexo' , 'Circunscripcion', 'Comuna', 'Nacionalidad', 'Pais Domicilio', 'Partido', 'Provincia', 'Region']).filter(lambda x: x['Número de registros'].count() == 1)

| | Cédula | Circunscripción | Comuna | DV | Edad | Nacionalidad | País Domicilio | País Nacimiento | Partido | Provincia | Rango Edad | Región | Sexo | Sufragio | Voto Exterior | N° de registros | Votaron |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 48 | 0 | Chanco | Chanco | 0 | 95 | chilena | Chile | Chile | Sin Partido | Cauquenes | 80 o + | Del Maule | M | no sufragó | Nacional | 1 | NaN |
| 133 | 0 | Iquique | Iquique | 0 | 97 | chilena | Chile | Chile | [3] Unión Democrata Ind. | Iquique | 80 o + | De Tarapacá | F | no sufragó | Nacional | 1 | NaN |
| 194 | 0 | El Golf | Las Condes | 0 | 98 | chilena | Chile | Chile | [2] Partido Demócrata Cristiano | Santiago | 80 o + | Metropolitana de Santiago | F | no sufragó | Nacional | 1 | NaN |
| 209 | 0 | Santa Bárbara | Santa Bárbara | 0 | 115 | chilena | Chile | Chile | Sin Partido | Biobío | 80 o + | Del Biobío | F | no sufragó | Nacional | 1 | NaN |
| 212 | 0 | Temuco centro | Temuco | 0 | 99 | chilena | Chile | Chile | [2] Partido Demócrata Cristiano | Cautín | 80 o + | De la Araucanía | M | no sufragó | Nacional | 1 | NaN |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 14840749 | 0 | Los Lagos | Los Lagos | 0 | 21 | chilena | Chile | Chile | [2] Partido Demócrata Cristiano | Valdivia | 20-24 | De los Ríos | M | sufragó | Nacional | 1 | 1.0 |
| 14840751 | 0 | Natales | Natales | 0 | 21 | chilena | Chile | Chile | [37] Evolución Política | Última Esperanza | 20-24 | De Magallanes y de la Antártica Chilena | M | sufragó | Nacional | 1 | 1.0 |
| 14842069 | 0 | Torres del Paine (C. Castillo) | Torres del Paine | 0 | 21 | chilena | Chile | Chile | [8] Humanista | Última Esperanza | 20-24 | De Magallanes y de la Antártica Chilena | M | sufragó | Nacional | 1 | 1.0 |
| 14842084 | 0 | Natales | Natales | 0 | 22 | chilena | Chile | Chile | [8] Humanista | Última Esperanza | 20-24 | De Magallanes y de la Antártica Chilena | M | sufragó | Nacional | 1 | 1.0 |
| 14851455 | 0 | Huara | Huara | 0 | 21 | chilena | Chile | Chile | [1] Renovación Nacional | Del Tamarugal | 20-24 | De Tarapacá | M | sufragó | Nacional | 1 | 1.0 |

*257349 rows × 17 columns*

**Figura 4.** Si agrupamos los datos por más columnas, podemos obtener más grupos de un elemento.

Por ejemplo, si sabemos que una persona está inscrita en Aysén, tenía 61 años el 2019, es japonés y tiene sexo masculino, sabemos con certeza que no votó: no hay otra persona con esas características en el conjunto de datos.

Peor aún, si consideramos más columnas, la cantidad de personas que potencialmente podríamos re-identificar aumentará aún más, llegando hasta un máximo de 257.349 personas (ver Figura 4).

## ¿Qué podemos hacer entonces?

No todo está perdido, por suerte existen garantías formales que nos ayudarán a obtener un mejor nivel de privacidad, donde la más básica se conoce como *k-anonimato* [2].

> Se dice que un conjunto de datos satisface *k-anonimato* si y solo si cada grupo de cuasi-identificadores aparece al menos *k* veces en el conjunto de datos.

Intuitivamente, esto quiere decir que si un atacante cuenta con información auxiliar, el individuo que busca re-identificar estará "mezclado" con otros *k*-1 individuos con las mismas características.

El conjunto de datos del Servel claramente cumple 1-anonimato, y sería bueno encontrar una serie de transfor-

| | Edad | Nacionalidad | Sexo | Comuna | count |
|---|---|---|---|---|---|
| 93171 | 138 | chilena | masculino | Freire | 1 |
| 80551 | 104 | chilena | masculino | Rauco | 1 |
| 80554 | 104 | chilena | masculino | Renca | 1 |
| 80557 | 104 | chilena | masculino | Rinconada | 1 |
| 80560 | 104 | chilena | masculino | Romeral | 1 |
| ... | ... | ... | ... | ... | ... |
| 10697 | 29 | chilena | femenino | Puente Alto | 5123 |
| 9041 | 27 | chilena | masculino | Puente Alto | 5143 |
| 12059 | 30 | chilena | masculino | Puente Alto | 5209 |
| 11042 | 29 | chilena | masculino | Puente Alto | 5214 |
| 10037 | 28 | chilena | masculino | Puente Alto | 5246 |

*93172 rows × 5 columns*

**Figura 5.** En gris se indican algunos de los grupos de personas con un solo elemento.

| | Rango Edad | pnac | Sexo | Region | count |
|---|---|---|---|---|---|
| 39 | 18-19 | Extranjero | femenino | De los Ríos | 1 |
| 739 | 70-74 | Extranjero | femenino | De Aysén del General Carlos Ibáñez del Campo | 3 |
| 803 | 75-79 | Extranjero | femenino | De Aysén del General Carlos Ibáñez del Campo | 6 |
| 99 | 20-24 | Extranjero | femenino | De Aysén del General Carlos Ibáñez del Campo | 6 |
| 43 | 18-19 | Extranjero | femenino | De Ñuble | 6 |
| ... | ... | ... | ... | ... | ... |
| 287 | 35-39 | Chile | masculino | Metropolitana de Santiago | 257260 |
| 143 | 25-29 | Chile | femenino | Metropolitana de Santiago | 277837 |
| 207 | 30-34 | Chile | femenino | Metropolitana de Santiago | 282399 |
| 159 | 25-29 | Chile | masculino | Metropolitana de Santiago | 283588 |
| 223 | 30-34 | Chile | masculino | Metropolitana de Santiago | 288647 |

*896 rows × 5 columns*

**Figura 6.** Datos agrupados por columnas generalizadas.

maciones a los datos que hagan que los datos publicados cumplan *k*-anonimato para al menos *k*>1 (entre más grande el *k*, mejor).

Como ejercicio, consideraremos solo las columnas edad, nacionalidad, sexo y comuna. Si agrupamos por estas columnas y contamos los distintos valores, podemos ver que hay muchos grupos con un solo miembro (ver Figura 5).

Como primer intento de anonimización podemos *generalizar* las columnas edad, nacionalidad y comuna. La generalización de una columna consiste en reemplazar sus valores por otros valores que corresponden a una categoría más amplia. Para la columna edad, generalizamos en rangos de edad de 5 en 5, para nacionalidad generalizamos las nacionalidades extranjeras como 'extranjera', y para comuna generalizamos publicando solo la región de ella. Al hacer esto obtenemos los grupos de la Figura 6.

Mejoraron los grupos, pero los datos siguen siendo 1-anónimos ya que queda 1 grupo con 1 integrante. Podemos solucionar esto haciendo una unión de los grupos de 18 a 19 con los de 20-24, obteniendo así datos que son 3-anónimos.

Claramente esta no es la solución final, ya que hay más quasi-identificadores en los datos publicados y no sabemos con certeza qué atributos conoce el atacante. Adicionalmente, aumentar la cantidad de atributos dificulta más la tarea de mejorar el *k*-anonimato, ya que los grupos serán más específicos, y por ende se requerirá generalizar aún más, disminuyendo la utilidad de los datos.

Otra solución que es independiente de los datos auxiliares que podría tener un atacante, es usar una técnica más reciente llamada *privacidad diferencial* [3].

> *La privacidad diferencial* consiste en perturbar los resultados agregando ruido aleatorio.

Esta técnica se especializa en publicar datos agregados y no microdatos (un subset de los datos originales), pero genera garantías de privacidad formales más fuertes. De hecho, esta técnica iba a ser usada para publicar los datos del censo del 2020 de Estados Unidos que se postergó por la pandemia por COVID-19 [4].

## Conclusión

Existen muchos otros casos de estudio emblemáticos acerca de violación de privacidad. Por ejemplo, en el 2007 Netflix realizó una competencia abierta para mejorar su sistema de recomendación. Para ello se proveyó de un dataset anonimizado, donde los identificadores fueron reemplazados por otros aleatorios. El problema fue que a estos datos se les hizo un ataque de asociación de registro con data pública de IMDB (Internet Movie DataBase), de-anonimizando así a muchos individuos, revelando sus preferencias de películas y series. Otro caso emblemático ocurrió en 1997 en Estados Unidos, donde un gobernador aprobó la liberación de registros médicos anonimizados de funcionarios públicos. Dos días después el gobernador recibió un correo con todos sus registros médicos, obtenidos mediante un ataque de asociación con el padrón electoral, cruzando código postal, fecha de nacimiento y género.

Vimos que existen técnicas formales que permiten evitar estos escándalos, y publicar datos asegurando la privacidad de los individuos. Pero aún existen desafíos, como por ejemplo (1) que la privacidad de datos aún no está interiorizada en el colectivo, y (2) no se sabe (ni se puede predecir) la información auxiliar con la que va a contar un atacante. Afortunadamente, estas técnicas están ganando terreno en grandes compañías, como el caso de la privacidad diferencial, que está siendo usada por Google y Microsoft para telemetría, LinkedIn para análisis de marketing, y Apple para la recolección de datos personales de usuarios. ■

### REFERENCIAS

[1] https://www.servel.cl/wp-content/uploads/2021/06/VW_VOTARON_2020PLEB_Datos_completos.zip.

[2] P. Samarati and L. Sweeney. 1998. “Protecting Privacy When Disclosing Information: k-Anonymity and Its Enforcement through Generalization and Suppression”. Technical Report SRI-CSL-98-04.

[3] C. Dwork and A. Roth. 2014. The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science 9, 3–4 (2014), 211–407.

[4] S. Ruggles, C. Fitch, D. Magnuson, and J. Schroeder. 2019. “Differential Privacy and Census Data: Implications for Social and Economic Research.” AEA Papers and Proceedings 109:403–408.

# Estudiantes DCC

En esta sección de la Revista estudiantes recientemente graduadxs del Departamento de Ciencias de la Computación (Universidad de Chile) nos cuentan, junto a sus profesores guías, sobre sus trabajos de memoria y/o tesis.

Tesis de doctorado

## Similarity-based Web Queries (Consultas por Similitud en la Web)

***Estudiante:*** Sebastián Ferrada
***Profesores guías:*** Benjamín Bustos y Aidan Hogan

Hacer un doctorado fue una continuación natural tras finalizar el magíster. En el magíster creé IMGpedia (imgpedia.dcc.uchile.cl), una base de conocimiento que contiene información visual y semántica sobre las imágenes de Wikipedia. En IMGpedia se pueden realizar consultas, como por ejemplo: "Obtener pinturas del Louvre que sean similares a un autorretrato de Van Gogh".

La información guardada en IMGpedia es estática, es decir, se calcularon descriptores visuales y relaciones de similitud una sola vez y se guardaron. El siguiente paso parecía lógico: utilizar esa información (o cualquier otra) de forma dinámica, calculando relaciones de similitud de forma eficiente y a pedido de un usuario. Este fue el problema central de mi doctorado.

En mi tesis titulada "Similarity-based Web Queries" o "Consultas por Similitud en la Web" hago tres contribuciones. Primero, propongo un nuevo algoritmo para resolver *similarity joins* de forma aproximada. El *similarity join* es una operación que dados dos conjuntos de datos, obtiene pares de objetos, uno de cada conjunto, tales que son similares entre sí bajo algún criterio. Los criterios de similitud son tan variados que podríamos discutir sobre ellos en un tratado de filosofía intentando descifrar la respuesta a ¿cuándo dos cosas son parecidas? En la práctica, "las cosas" suelen ser vectores de números y "el parecido" se mide a través de funciones de distancia, siendo dos vectores más similares mientras más pequeña sea la distancia entre ellos. Nuestro algoritmo, *root-join*, agrupa objetos cercanos entre sí y calcula distancias solo entre elementos de cada grupo, ofreciendo así garantías de bajo tiempo de ejecución, al costo de no obtener una respuesta 100% correcta.

Como segunda contribución, definí e implementé un operador de bases de datos para extender el lenguaje de Web de consulta SPARQL, lo que permite efectivamente realizar *similarity joins* sobre la Web de Datos, dándole la libertad a los usuarios para seleccionar las dimensiones relevantes para ellos y seleccionar una función de distancia acorde a sus necesidades. La tercera contribución son dos aplicaciones. Por un lado, cargamos los datos de IMGpedia en un servidor que soporta nuestro operador de *similarity join*, para permitir que las consultas sobre las imágenes sean ahora dinámicas. Por otro lado, utilizando el conocimiento adquirido al definir un operador nuevo en un lenguaje de consulta, actualizamos una propuesta antigua para poder realizar *clustering* sobre los resultados de una consulta SPARQL, definiendo una sintaxis que permite al usuario seleccionar un algoritmo de *clustering* y proveer los parámetros necesarios.

Como me fue bien durante el magíster, cometí el error de pensar que el doctorado iba a ser fácil y exitoso. Sin embargo, tuve que lidiar con problemas de salud mental justo a la mitad de mi periodo de estudios, lo que por un lado mermó mi motivación para hacer cosas y, por otro, puso en peligro el proyecto de tesis. Soy afortunado de contar con una gran red de apoyo, lo que me ayudó a salir adelante, realizar tratamiento y finalmente concluir el trabajo. En el último año del doctorado, comenzó también la pandemia, lo que me forzó a terminar mi tesis desde casa y a realizar mi defensa a través de videoconferencia. El doctorado fue una experiencia muy difícil, pero ya habiendo terminado, creo que ha valido la pena, pues estoy haciendo el trabajo que siempre quise hacer: investigar y enseñar.

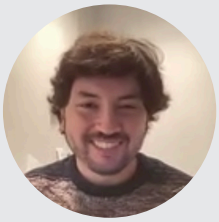
El trabajo con mis profesores guías, Benjamín y Aidan fue muy fructífero. Ellos me daban la independencia necesaria para probar cosas por mi cuenta y luego me ayudaban a discutir conceptos y a resolver dudas que iban surgiendo en el camino. Siempre me sentí muy apoyado por ellos y me enseñaron un montón, no solo sobre temas técnicos de mi tesis, sino también sobre cómo ser un investigador en general. Siempre voy a estar agradecido de haberlos tenido en mi camino.

Actualmente estoy haciendo un postdoc en el Departamento de Ciencias de la Computación de la Universidad de Linköping, en Suecia.

Tesis de doctorado

# Data Structures and Algorithms for Analyzing DNA Sequences in Compressed Space

***Estudiante:*** Diego Díaz
***Profesor guía:*** Gonzalo Navarro



Mi interés por hacer un doctorado surgió a partir de los desafíos técnicos a los que me enfrentaba en mi trabajo. En ese momento era ingeniero de proyectos en el Centro de Modelamiento Matemático de la Universidad de Chile y estaba analizando datos de ADN. Mi problema era que los *datasets* que manejaba eran tan masivos, que los programas que necesitaba para procesarlos solo se podían ejecutar en el clúster de Beauchef (Leftraru).

Buscando investigadores que trabajaran en temas afines a los míos, di con el profesor Gonzalo Navarro. Tuvimos una primera reunión y me explicó cómo sería la dinámica si era su alumno. Sentí que era una buena oportunidad, así que postulé al doctorado del DCC, y afortunadamente quedé.

Mi primer tiempo en el programa fue un poco difícil, ya que mi formación en computación durante pregrado no fue tan rigurosa como la que se da en el DCC. Tuve que nivelarme rápidamente para no fallar en los cursos y eso me generó estrés. A pesar de esto, lo recuerdo como un tiempo muy enriquecedor donde aprendí varias cosas nuevas.

Mi investigación de doctorado tuvo como objetivo desarrollar algoritmos y estructuras de datos eficientes para procesar datos genómicos. Este tipo de colecciones son *strings* que codifican las secuencias de ADN de los genomas. El problema es que estos datos son tan masivos, que operar sobre ellos resulta muy caro computacionalmente. Otra característica de los datos de ADN es que son muy redundantes. Es decir, las mismas secuencias aparecen varias veces en el texto, con pequeñas variaciones. Uno puede sacar ventaja de este hecho para reducir los costos del análisis. La idea general consiste en determinar un pequeño diccionario de frases representativas del texto, realizar cálculos sobre estas frases, y luego extrapolar los resultados a las copias de las frases. Como los datos son repetitivos, el diccionario debería ser mucho más pequeño que el texto original, reduciendo de esta forma el número de operaciones redundantes.

La idea de las frases representativas suena razonable. Sin embargo, tiene sus salvedades. Por ejemplo, si el algoritmo que utilizas para construir el diccionario es más caro que el algoritmo que originalmente querías hacer más eficiente, entonces no tiene mucho sentido. En el doctorado, exploramos técnicas que tuviesen un buen *trade-off* entre nivel de compresión y el uso de recursos computacionales.

Cuando realizas un doctorado, no solo adquieres conocimiento en el área en la que trabajas. También hay otro tipo de aprendizajes que son claves para investigar, como ser capaz de explicar claramente tu trabajo a tus pares, ya sea de forma oral o escrita. Lo que me gustó del DCC es que también se preocupa de que las personas desarrollen este tipo de habilidades. Por ejemplo, el idioma oficial del programa es el inglés. Esto puede parecer una dificultad extra al principio, pero luego que te acostumbras, comunicar tus resultados se va a haciendo cada vez más fácil.

Actualmente estoy realizando un postdoctorado en la Universidad de Helsinki, en el grupo de algoritmos bioinformáticos. Me gustaría continuar trabajando en la academia en el futuro, donde pueda investigar los temas que considero interesantes.

Tesis de doctorado

# Towards a Fine - Grained Linking Approach

***Estudiante:*** Henry Rosales
***Profesores guías:*** Aidan Hogan y Bárbara Poblete

Mi paso por el programa doctoral del DCC partió en el semestre de primavera de 2016. Mi primer desafío fue seleccionar el tema de investigación y el/la supervisor/a que se alineara a mis intereses. Si bien ya me había documentado al respecto mirando principalmente las publicaciones de los profesores en DBLP, me ayudó mucho asistir a las charlas que cada *profe* impartió en el ramo de Investigación. Mi línea de investigación hasta ese momento se orientaba hacia el reconocimiento de patrones y minería de datos, por lo que decidí acercarme a la profesora Bárbara Poblete, quien aborda temas afines. Tuve la suerte de que la profesora Bárbara me propusiera una cotutoría con el profesor Aidan Hogan y así poder unir áreas sumamente interesantes: minería de datos y web semántica.

El segundo desafío que enfrenté fue la selección del tema de investigación. Esto llevó varias semanas de debate junto con los tutores. Finalmente seleccionamos la tarea de *entity linking*. Esta tarea tiene como objetivo identificar las entidades en un texto en lenguaje natural y enlazarlas con sus entidades correspondientes en una *knowledge base/graph*. Luego de revisar por algunos meses el estado del arte de *entity linking* decidí enfocarme en la creación de un sistema que tuviera un enfoque multilingüe, ya que había poco trabajo relacionado. Los primeros trabajos fueron propuestos en el 2007 junto a la popularidad de Wikipedia y la creación de *knowledge bases* como Freebase y DBpedia.

Al ser *entity linking* una tarea relativamente reciente, había muchas ideas por desarrollar. Sin embargo, varias preguntas importantes no quedaban claras en la comunidad, por ejemplo: ¿qué entidades se deberían identificar y enlazar?, ¿son claros los mecanismos de evaluación?, ¿las métricas de evaluación actuales son suficientes?, entre otras. Eran pocas las preguntas que tenían respuestas, por tanto tuvimos que responderlas.

Optando por la definición más amplia de entidad, creamos una herramienta de anotación para crear nuestros propios *datasets* multilingües. Esta herramienta sirvió de base para anotar manualmente y semiautomáticamente un *dataset* en cinco idiomas. Con este *dataset* exploramos la factibilidad de adoptar traducción automática obteniendo resultados favorables. Por otra parte, propusimos una métrica basada en *fuzzy sets* para poder evaluar a un mismo sistema en diferentes dominios de aplicación. En resumen, cuando ya estaban las bases para crear el sistema de *entity linking* multilingüe que soñé ya tenía todos los requisitos para defender la tesis.

El doctorado me dejó muchas experiencias no solo relacionadas con la ciencia de la computación. Pude expandir mis horizontes más allá de América y darme cuenta de que: al pedir un Uber en Hong Kong te pueden recoger en un Tesla; en Bolonia (Italia) puede resultar difícil imprimir un póster; en Colombia no se baila la salsa de la misma forma que en Cuba; la espera de un vuelo de conexión en Estados Unidos puede retrasarse un día completo. Como dice una frase que escuché en la Escuela de Verano ISWS'18 en Italia y con la que no puedo estar más de acuerdo, "la investigación debe ser divertida, sino lo es, es porque algo estás haciendo mal".

Actualmente me encuentro realizando un postdoctorado en el Instituto Milenio Fundamentos de los Datos (IMFD). Estoy participando en la construcción de un *knowledge graph* junto a los profesores Aidan Hogan (Unversidad de Chile) y Renzo Angles (Universidad de Talca) sobre varias fuentes de datos. Por otro lado, tuve la oportunidad de dictar el ramo de Programación Avanzada en la Universidad de O'Higgins como profesor de jornada parcial.

Tesis de doctorado

## Piecewise Adjacent Contours for Multicellular Structures in Fluorescence Microscopy Images

***Estudiante:*** Jorge Jara
***Profesores guías:*** Nancy Hitschfeld y Steffen Härtel

Mi motivación para hacer el doctorado surgió del interés en geometría e imágenes, aplicadas a células y sus componentes en distintos escenarios experimentales, con especial foco en imágenes tridimensionales de fluorescencia en vivo y de alta resolución. Mi tema de doctorado abordó algoritmos de geometría computacional y procesamiento de imágenes para representar membranas celulares dinámicas capturadas en estas condiciones. El título de mi tesis de doctorado es "Piecewise Adjacent Contours for Multicellular Structures in Fluorescence Microscopy Images", que puede traducirse a "Contornos Adyacentes por Tramos para Estructuras Multicelulares en Microscopía de Fluorescencia".

La microscopía de fluorescencia es una técnica óptica que permite observar células y organismos con imágenes 3D y, mediante el uso de múltiples marcadores fluorescentes, distinguir estructuras en distintos colores. Más aún, es posible realizar capturas en el tiempo y producir secuencias de video, ya que las muestras en observación pueden vivir durante varios días. La desventaja de esta técnica es que su resolución es limitada debido a fenómenos de dispersión de la luz, lo que impide distinguir objetos demasiado pequeños (en comparación con la microscopía electrónica, la mejor existente, pero casi imposible de usar con muestras vivas). Además, el comportamiento natural de los especímenes vivientes involucra movimientos y deformaciones muy rápidos que pueden hacer que las imágenes adquiridas pierdan aún más definición, viéndose borrosas. Aun así, la microscopía de fluorescencia ha sido clave para una colorida revolución de imágenes biomédicas en los últimas dos décadas, ya que ha hecho posible observar fenómenos 3D en tiempo real, como el desarrollo del cerebro de mosca, ratón, o pez cebra fluorescente. Este último fue utilizado como modelo para mi trabajo de tesis.

Desarrollé ALPACA, que es una abreviatura del nombre en inglés "ALgorithm for Piecewise Adjacent Contour Adjustment", traducido como "algoritmo para ajuste de contornos adyacentes por tramos", publicado el 2020 en la revista Journal of Microscopy (https://doi.org/10.1111/jmi.12887). ALPACA detecta y optimiza contornos de estructuras celulares adyacentes capturadas en imágenes de microscopía de fluorescencia. ALPACA se enfoca en secciones de contorno adyacentes que son cuantificadas para análisis de forma y organización, para aplicaciones a nivel subcelular, celular y supracelular. Combinamos ALPACA con algoritmos de la familia de contornos activos para interpolar y optimizar la forma de cada contorno y de sus secciones, y lo evaluamos con imágenes sintéticas y reales capturadas por microscopía confocal de disco rotatorio 3D *in vivo*, con el apoyo de usuarios con distintos niveles de experiencia en el sistema biológico modelo y microscopía de fluorescencia.

Tesis de magíster

## Un método interpretable para clasificación general usando teoría de Dempster-Shafer

***Estudiante:*** Sergio Peñafiel
***Profesor guía:*** Nelson Baloian

En esta tesis abordamos el problema conocido en el campo de inteligencia artificial como "clasificación", esto es, dada una muestra caracterizada por un conjunto de parámetros, determinar a qué clase pertenece, de un conjunto predefinido. Las aplicaciones de este problema son tan variadas, que pueden ir desde identificar el tipo de subespecie de una flor dadas sus características, hasta determinar si un individuo tiene riesgo de contraer una enfermedad dado su historial médico. Para medir la eficacia de un método o modelo de clasificación normalmente se usa la precisión que tiene en clasificar una muestra en la clase que le corresponde. Hoy en día ha tomado fuerza la idea de que un modelo de clasificación debe ser también interpretable, es decir, el modelo debe inherentemente explicar por qué tomó tal o cual decisión.

En general los modelos más precisos son poco interpretables y viceversa, por lo que el desarrollo de uno que sacrifique algo de precisión pero que sea interpretable es una tarea interesante de resolver. Para la tesis, entonces, desarrollamos un modelo totalmente nuevo, basado en la teoría de la plausibilidad de Dempster-Shafer, que en precisión es comparable a los mejores que se conocen, pero es ampliamente interpretable. La motivación para el desarrollo de esta tesis fue una invitación de una empresa japonesa dedicada a la informática médica para desarrollar un modelo que permita predecir el riesgo de un individuo de sufrir un ataque al corazón, por lo que tuvimos la oportunidad de viajar a Japón para probarlo con datos recopilados por el Ministerio de Salud japonés, para una cierta área (prefectura) obteniendo muy buenos resultados. Los resultados se publicaron en dos revistas de alto impacto (IEEE Access y Artificial Intelligence with Applications).

Tesis de magíster

## Implementando un reporte seguro de acoso sexual

***Estudiante:*** Ilana Mergudich
***Profesor guía:*** Alejandro Hevia

El trabajo de tesis de Magíster en Ciencias mención Computación de Ilana fue motivado por la necesidad de obtener un sistema eficiente para reportar situaciones de acoso y abuso sexual en forma responsable pero anónima, esto es, donde las acusaciones anónimas tuvieran la garantía de ser reveladas bajo condiciones que colectivamente se consideren adecuadas. Una instancia concreta de esta idea fundamental es un sistema informático, implementado como sistema distribuido, donde las identidades de todas las personas participantes (tanto acusadores/as como acusados/as) se mantienen anónimas mientras una condición específica no se alcance, por ejemplo, no exista un número mínimo (digamos dos) de acusaciones contra la misma persona. De ocurrir, se revelan las identidades y se puede proceder a una investigación formal.

Si bien el sistema propuesto por Ilana no es el primero en resolver este problema con un algoritmo distribuido criptográfico (dicho sistema existente se denomina WhoToo), su contribución principal fue mejorar WhoToo sustancialmente en aspectos internos claves. Para ello, rediseñó varios de los algoritmos criptográficos utilizados por otros igualmente seguros pero significativamente más eficientes. Para comprobarlo más allá de la teoría, Ilana incluso implementó un prototipo factible de ser usado en un contexto real, por ejemplo en una universidad. La fortaleza del trabajo de Ilana fue combinar nuevos algoritmos criptográficos con un diseño orientado a obtener resultados prácticos y eficientes.

El trabajo realizado en esta tesis fue publicado en el artículo "Implementing Secure Reporting of Sexual Misconduct – Revisiting WhoToo" (con Alejandro Hevia como coautor) publicado en octubre de 2021 en una importante conferencia internacional de criptografía (Latincrypt 2021), la cual se realiza en cooperación con la Asociación Internacional de Investigación Criptográfica.

Memoria de pregrado

## Estructuras compactas dinámicas más eficientes para bases de datos de grafos con atributos

***Estudiante:*** Dania de la Puente
***Profesores guías:*** Gonzalo Navarro y Diego Arroyuelo

La memoria de Dania de la Puente, "Estructuras compactas dinámicas más eficientes para bases de datos de grafos con atributos", coguiada por Gonzalo Navarro con Diego Arroyuelo (Universidad Técnica Federico Santa María), aborda el problema de representar bases de datos de grafos en forma más eficiente en términos de espacio. Este tema es muy relevante en la actualidad porque las bases de datos de grafos de distintos tipos están apareciendo en todo tipo de aplicaciones como una forma más flexible que las relacionales para almacenar y buscar información, y están creciendo en tamaño. Una representación más compacta permite manejar mayores grafos en memoria principal (donde el procesamiento es mucho más rápido) o en dispositivos de memoria limitada como celulares.

Si bien existen ya algunas representaciones compactas para bases de datos de grafos, la memoria de Dania se centra en el caso menos estudiado de representaciones dinámicas, es decir, que permiten modificar las aristas del grafo al mismo tiempo que este se consulta. También se centra en un tipo de representación de grafos basada en *tries*, que ven la matriz de adyacencia del grafo como una grilla bidimensional y almacenan las aristas como las coordenadas de los puntos correspondientes, intercalando convenientemente sus dígitos para formar las secuencias que el *trie* almacena, del mismo modo que lo haría un *quadtree*.

La memoria tiene tres contribuciones importantes. La primera es completar una estructura que habían desarrollado los profesores guías, basada en lo que se llama un $k^2$-*tree* dinámico, para agregarle la operación de borrar aristas. El resultado muestra que se puede borrar tan eficientemente como se insertaban aristas en esta estructura. La segunda contribución es comparar la estructura resultante con varias representaciones alternativas de *tries* de coordenadas, algunas ya usadas antes con este propósito y otras aplicadas por primera vez a este problema. El resultado muestra que la nueva estructura adaptada por Dania resulta ser la más conveniente cuando se consideran todos los aspectos de la funcionalidad requerida. La tercera contribución fue integrar la nueva estructura en un sistema de bases de datos existente, que usaba una representación dinámica de $k^2$-*tree* anterior. Los resultados muestran que la nueva estructura es varias veces más rápida que la anterior, al precio de usar una módica cantidad de espacio extra.

Memoria de pregrado

## Análisis del COVID-19 y sus correlaciones a nivel internacional

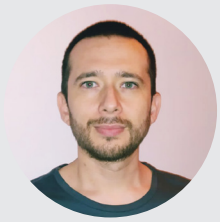
***Estudiante:*** Tamara Novoa - Rodríguez
***Profesor guía:*** Aidan Hogan

Al día de la fecha, aún no se entiende muy bien por qué hay tanta variación entre los diferentes países con respecto a los números reportados, per cápita, de casos y muertes asociadas al COVID-19. Se han planteado varias hipótesis, relacionadas, por ejemplo, con el clima, con la prevalencia de comorbilidades (como la obesidad, la diabetes, etc.), con la edad media de la población, con las políticas públicas del país, etc. Aunque algunas iniciativas ya han analizado esta situación, han estado enfocadas siempre en una hipótesis particular.

En el trabajo de título de Tamara Novoa (alumna del Departamento de Ingeniería Electrica), guiado por Aidan Hogan (profesor del DCC), generamos un "cubo" de datos multidimensional para intentar correlacionar diversas variables sobre países. Habiendo recolectado datos de varias fuentes, separamos las variables en dos categorías: aquellas referentes al COVID-19 (por ejemplo, número de casos, número de muertes, etc.), y aquellas no referentes al COVID-19 (por ejemplo, el PIB per cápita, emisión de dióxido de carbono, saldos promedios, uso de Internet, etc.). Luego buscamos variables correlacionadas entre ambas categorías. Los resultados aportaron nuevas perspectivas de la pandemia y fueron inesperados en algunos casos. Por ejemplo, los países que tienen mayor porcentaje de la población con acceso a instalaciones para lavarse las manos tendían a tener *más* casos de COVID-19. De manera más general observamos un posible factor latente de que los países más desarrollados tendían a tener más contagios y defunciones reportados.

Publicamos un artículo corto en la International Semantic Web Conference (ISWC) y el trabajo fue uno de los tres nominados para el premio "Best Demo".

Memoria de pregrado



# Herramienta hidroinformática V.GeoLinkage: automatización de modelos hidrológicos integrados

***Estudiante:*** Felipe Troncoso
***Profesor guía:*** Pedro Sanzana y Nancy Hitschfeld

En el marco de un proyecto multidisciplinario cuyo objetivo fue llevar a cabo el diagnóstico para la gestión de explotación del acuífero Valle de Azapa, se desarrolló la herramienta V.GeoLinkage que permite trabajar con un modelo que integra en una sola visión la perspectiva superficial y subterránea del flujo de agua. Es común en este tipo de proyectos de recursos hídricos que exista una parte del equipo especializada en cada perspectiva (superficial y subterránea), contando con dos modelos para describir la misma red de drenaje. Este tipo de modelación integrada requiere previamente contar con la relación entre los dominios y topologías de ambas representaciones, las cuales generalmente tienen una discretización espacial diferente. El dominio superficial es típicamente segmentado en unidades hídricas: subcuencas, bandas de elevación o unidades de respuesta hidrológica, las cuales son representadas por triángulos o polígonos irregulares simples, mediante arcos y nodos. Por otro lado, el dominio subterráneo está organizado en unidades hidrogeológicas, que son representadas por grillas, triangulaciones, *quadtrees* o diagramas de Voronoi.

La herramienta desarrollada, V.GeoLinkage, encuentra cómo se relacionan estas representaciones para un modelo superficial de la plataforma WEAP (weap21.org) y un modelo subterráneo en la plataforma MODFLOW (www.usgs.gov). Desarrollamos una interfaz en Python que se integra dentro del sistema de información geográfica GRASS-GIS y extrae como mapas vectoriales ambas representaciones. Estos mapas son procesados y relacionados usando la integración con las herramientas que posee GRASS-GIS. Finalmente, se genera un archivo que almacena esta relación entre las discretizaciones espaciales, el cual es usado por WEAP para ejecutar el modelo acoplado.

La automatización lograda en el preprocesamiento permitió pasar de días a minutos, haciendo factible trabajar con escenarios que enriquecieron el diagnóstico con un mayor número de hipótesis, por ejemplo, un cambio en uso de suelo o una expansión espacial de zonas agrícolas, durante el periodo de estudio. En el portal de la Dirección General de Aguas del Ministerio de Obras Públicas se puede acceder al informe completo (snia.mop.gob.cl/sad/SUB5917v1.pdf). Adicionalmente el trabajo lo presentamos en la sesión de hidroinformática de la conferencia internacional European Geosciences Union (EGU), al cual se puede acceder en https://doi.org/10.5194/egusphere-egu21-13909.

dcc
CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE CHILE

PEC
DCC

# MAGÍSTER EN TECNOLOGÍAS DE LA INFORMACIÓN

Departamento de Ciencias de la Computación
*Universidad de Chile*

Acreditación
7 años acreditada
Carrera Acreditada hasta diciembre 2022
Comisión Nacional de Acreditación CNA-Chile

## ÁREAS DE ESPECIALIZACIÓN:

- Ciencia e Ingeniería de Datos
- Inteligencia Artificial
- Ingeniería de Software
- Gestión de Calidad de Software
- Gestión de Proyectos Informáticos
- Seguridad Computacional

## INFORMACIÓN Y POSTULACIONES EN:

https://www.dcc.uchile.cl/postgrado/mti/

## CONTACTO:

Departamento de Ciencias de la Computación
FCFM, Universidad de Chile

Beauchef #851, Edificio Norte, tercer piso, oficinas 326-327. Santiago.

**Correo:**
capacita@dcc.uchile.cl

**Teléfonos:**
+56 2 2978 4965
+56 2 2978 4976
+56 9 3871 9957 *WhatsApp*

REVISTA DEL DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN DE LA UNIVERSIDAD DE CHILE

# Bits DE CIENCIA