

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.

Topological data analysis for classification of noisy and high-dimensional datasets

Estudiante

Rolando Kindelan

Profesores guías

Nancy Hitschfeld Kahler y
Mauricio Cerda



Nací en Cuba, donde estudié Ingeniería en Ciencias Informáticas y luego realicé un magíster en la misma Universidad de las Ciencias Informáticas. Durante más de siete años trabajé en la industria, en proyectos de desarrollo de software y equipos médicos para el sistema de salud cubano. Esa experiencia me enseñó lo valioso que es cuando la investigación se traduce en soluciones concretas para mejorar la vida de las personas.

En 2018 llegué a Chile para comenzar el Doctorado en Computación en la Universidad de Chile, bajo la guía de Nancy Hitschfeld Kahler y Mauricio Cerda. Fue un cambio enorme: adaptarme a un nuevo país, integrarme a una comunidad académica distinta y, al mismo tiempo, retomar la vida de estudiante después de tantos años en el mundo laboral.

Mi tesis, titulada *“Topological Data Analysis for Classification of Noisy and High-Dimensional Datasets”*, se centró en el Análisis Topológico de Datos (TDA), un enfoque emergente que aplica herramientas de la topología y la geometría para estudiar la “forma” de los datos. Me gusta explicarlo con una imagen sencilla: pensemos en un cardumen de peces o en una bandada de estorninos. Si estamos dentro, sólo vemos a los individuos cercanos; lo global se pierde. El aprendizaje automático tradicional funciona un poco así: capta patrones locales. El TDA, en cambio, nos permite mirar desde dentro y desde fuera a la vez, descubriendo tanto estructuras locales como patrones globales que guían el movimiento del conjunto.

Gracias a esta capacidad, el TDA suele usarse como complemento en los procesos de análisis de datos. Sin embargo, en mi investigación quisimos ir más allá y demostrar que puede ser una herramienta central para enfrentar desafíos como datos con ruido, valores faltantes, clases desbalanceadas o etiquetas incorrectas. En muchos de estos problemas, lo que realmente importa es entender bien la estructura de los datos, y ahí el TDA ofrece una perspectiva única.

Un hallazgo clave fue notar que muchas soluciones tradicionales dependen de *grafos de vecinos cercanos*, que representan sólo relaciones de pares y no siempre reflejan la for-

ma real de los datos. Por eso propusimos reemplazarlos por estructuras más expresivas, capaces de capturar relaciones de orden superior, como triángulos o cavidades. Esto abrió nuevas posibilidades para diseñar métodos de clasificación más robustos y aplicables a una amplia gama de escenarios.

Otro gran desafío fue el costo computacional: los objetos que maneja el TDA crecen rápidamente y son difíciles de almacenar y procesar. Para resolverlo, diseñamos estructuras de datos compactas que reducen drásticamente el uso de memoria sin perder eficiencia. También introdujimos nuevas formas de comparar de manera rápida y confiable las “huellas topológicas” de los datos, lo que permitió escalar los métodos a conjuntos mucho más grandes.

Más allá de los resultados técnicos, este doctorado fue para mí un viaje de transformación personal. Aprendí a explicar ideas complejas en un lenguaje sencillo, a trabajar en colaboración con investigadores de distintos países y a no rendirme frente a problemas que parecían imposibles. Hoy miro hacia atrás con gratitud: por mi familia, que me acompañó en cada paso; por mis profesores, que me guiaron con paciencia y rigor; y por la comunidad del DCC, que me recibió con generosidad desde el primer día.

Actualmente me desempeño como Staff R&D Engineer en Synopsys Chile Innovation Center, en el grupo de Soluciones de Procesamiento Distribuido, ayudando a que los productos de la compañía escalen en eficiencia y rendimiento. Además, codicito un curso de Introducción al Análisis Topológico de Datos en la Universidad Católica, y mi meta es poder impartirlo este año acá en nuestro DCC. Con ello espero seguir difundiendo esta área emergente y contribuir a formar nuevas generaciones de investigadores que exploren cómo la topología puede ayudarnos a entender mejor los datos que gobiernan nuestro mundo.

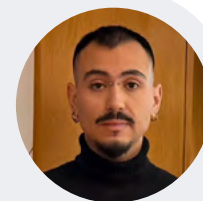
A study on repetitiveness measures for strings

Estudiante

Cristian Urbina

Profesor guía

Gonzalo Navarro



Estudié Licenciatura en Ciencia de la Computación en la Universidad de Santiago de Chile, donde rápidamente desarrollé un gusto por los aspectos teóricos de la Ciencia de la Computación, en especial Teoría de Conjuntos, Matemática Discreta y Teoría de la Computación.

Poco después de terminar mi pregrado, comencé mi doctorado en la Universidad de Chile financiado por una beca ANID Doctorado Nacional, donde decidí realizar mi tesis con el profesor Gonzalo Navarro.

La problemática en que nos concentramos para mi doctorado fue la siguiente: Hoy en día, existen enormes colecciones de texto donde gran parte de los datos son repetidos, o muy similares entre sí. Esto es especialmente evidente en colecciones de genomas en bioinformática, donde dos genomas humanos coinciden en más del 99% de su contenido. Estas colecciones son tan grandes que deben ser comprimidas para poder ser manejables. Idealmente se busca mantener la posibilidad de responder consultas del texto original en espacio comprimido.

Los compresores e índices de texto tradicionales basados únicamente en la entropía de Shannon son capaces de explotar las frecuencias relativas de los símbolos como factor de compresión, pero no pueden capturar la repetitividad. Por ejemplo, la entropía de Shannon nos dice que para representar la concatenación de n copias del texto 01, necesitamos $2n$ bits. Explotando la repetitividad de este texto, nos damos cuenta de que basta con almacenar 01 con 2 bits, y el entero n usando $\log(n)$ bits. Esto hace evidente que la entropía de Shannon no es una buena medida de compresibilidad para colecciones altamente repetitivas.

Mi tesis se titula "A Study on Repetitiveness Measures for Strings" y estudia, desde un punto de vista principalmente teórico, las distintas medidas de repetitividad que se han propuesto como alternativa a la entropía de Shannon. Muchas de estas medidas están asociadas al tamaño de compresores ampliamente utilizados en la práctica, como variantes del algoritmo de compresión Lempel-Ziv, el *run-length encoding* de la transformada de Burrows-Wheeler, u otras ideas basadas en compresión usando gramáticas libres de contexto.

En concreto, nuestras contribuciones son las siguientes:

1. Estudiamos propiedades combinatorias de varias medidas de repetitividad del estado del arte, especialmente la transformada de Burrows-Wheeler. En específico, estudiamos qué tan robusta es esta medida de repetitividad cuando los textos son dinámicos, es decir, pueden cambiar en el tiempo.
2. Introdujimos nuevas medidas de repetitividad basadas en la noción de morfismo sobre textos. Mostramos que estas medidas son competitivas, e incluso ofrecen la posibilidad de romper lo que muchos consideran cotas inferiores para la repetitividad.
3. Extendimos las gramáticas libres de contexto utilizadas para compresión con nuevos tipos de reglas, obteniendo de esta forma una nueva representación comprimida más poderosa en términos de espacio, y que retiene gran parte de la funcionalidad que ofrecen las gramáticas para indexar textos comprimidos.
4. Generalizamos medidas de repetitividad de textos en una dimensión, a textos en d -dimensiones. Esto es importante porque existen colecciones de datos multidimensionales (como matrices, grafos, coordenadas, entre otros) que son altamente repetitivas, pero el estudio de la repetitividad hasta hace poco se concentraba únicamente en textos de una dimensión.

Mi experiencia en el doctorado fue muy positiva. Tuve la posibilidad de colaborar y conocer personas brillantes de distintas partes del mundo, lo cual me mantiene motivado a seguir mejorando como investigador. También pude realizar dos estancias de investigación en la Universidad de Palermo en Italia, y presentar varios artículos en conferencias internacionales. Recientemente participé en la publicación de dos artículos en las conferencias MFCS 2025 y SPIRE 2025. En esta última, obtuvimos el *Best Paper Award*.

Actualmente estoy postulando y esperando resultados de becas para realizar un postdoctorado en el extranjero. A corto plazo pienso continuar investigando medidas de repetitividad, y de a poco empezar a explorar otras líneas de investigación que me puedan ayudar a desarrollar nuevas ideas.

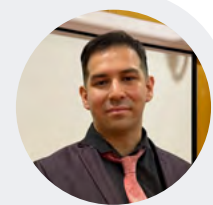
Polygonal/polyhedral mesh generation from Delaunay tessellations

Estudiante

Sergio Salinas Fernández

Profesora guía

Nancy Hitschfeld Kahler



Entré al doctorado por una noticia que me llegó al celular, mientras trabajaba de programador *full stack*. Era una noticia de las investigaciones de la doctora Nancy Hitschfeld, profesora del DCC. Las encontré fascinantes y me contacté por correo para ver si podía cooperar con ella. Amablemente me coordinó una reunión presencial, donde me explicó sus investigaciones y me recomendó entrar al doctorado a trabajar con ella. Fue así como nació mi investigación doctoral, el algoritmo Polylla para la generación de mallas poligonales.

Cuando buscaba temas, quería trabajar en algo altamente matemático, pero también práctico; ahí descubrí la generación de mallas poligonales. Para explicarlo fácil: Imagine que se quiere simular si un edificio soporta un terremoto. Este tipo de problemas requiere que se pueda representar una geometría, en este caso un edificio, en una computadora para aplicar algún método numérico sobre el objeto. Para ello se usan varios elementos planos iguales, como triángulos o cuadrados conectados, sin que se superpongan, que representan la geometría—básicamente “un mosaico”—pero siempre repitiendo la misma figura, y sobre esta geometría se le aplica un método numérico que haga la simulación. Pero existe un *trade-off*: si se usan muchos elementos, la simulación va a ser más lenta en tiempo y va a ocupar más memoria, pero va a ser más precisa; si se ocupan pocos elementos va a ser más rápida pero no tan útil.

Justo cuando entré, un método numérico relativamente nuevo estaba destacando: el Virtual Element Method (VEM). Este permite usar cualquier figura como elemento base, no sólo triángulos, y además no tienen que ser todas iguales. Pero no existía ningún algoritmo que hiciera una malla poligonal para el VEM, así que con Nancy nos propusimos resolver el problema creándolo nosotros mismos. De ahí surgió Polylla (*POLYgonal meshing aLgorithm bAsed on terminal-edge regions*). Con Polylla, no sólo desarrollamos un nuevo tipo de generador de mallas poligonales, también ampliamos la investigación proponiendo métodos para generarlas usando aceleración con GPU, para más rapidez, y estructuras compactas, para usar menos memoria.

Sobre el doctorado, fue una experiencia bastante estresante en un inicio. Entré al doctorado sin magíster por lo que no tenía experiencia en investigación. Mi pregrado fue Licenciatura en Ciencias de la Computación en la Universidad de Santiago, cuando me contacté con Nancy, coincidió también que me había ganado varios premios por mi desempeño académico, por lo que lo tomaron en cuenta cuando aceptaron mi postulación.

Al entrar no sabía escribir publicaciones, ni demostrar matemáticamente algoritmos y nunca había trabajado con geometría computacional, ni temas relacionados. Muchas veces intenté renunciar y abandonar todo. Pero la profesora Nancy siempre fue tan amable y cooperativa conmigo; ella me ayudó a mitigar cualquier problema que tuve y me ayudó mucho a crecer como persona y profesionalmente.

Una vez que realizamos nuestra primera publicación todo se aligeró bastante, ya que la investigación captó la atención de investigadores internacionales, que nos dieron un excelente *feedback* sobre cómo seguir investigando, como también de investigadores nacionales que nos ayudaron a sacar más publicaciones. Incluso dentro del departamento la investigación hizo ruido y nos empezaron a llegar alumnos que querían trabajar con nosotros. Ya al final del doctorado me dedicaba más a dirigir investigaciones con los alumnos, que tener que escribirlas. Además, gracias al doctorado, pude hacer pasantías y viajar a conferencias internacionales, donde conocí sobre otras culturas, lo que logró que cambiara toda mi perspectiva de cómo veía el mundo.

Por desgracia, estar demasiado enfocado en el doctorado, me llevó a problemas familiares y de salud, por lo que decidí tomarme un descanso de la academia. Ahora trabajo como ingeniero de datos para la empresa francesa Alstom. Aún estoy decidiendo qué hacer con mi vida, pero agradezco mucho al doctorado y a Nancy por todas las oportunidades que me dieron. Especialmente a Nancy.

Multilingual hate speech detection

Estudiante

Aymé Arango

Profesores guías

Bárbara Poblete y Jorge Pérez



Tras finalizar mis estudios de pregrado en Ciencias de la Computación en la Universidad de Oriente, Cuba, comencé a explorar nuevas oportunidades de formación académica. Fue entonces, a través de compañeros de estudios, que supe de la posibilidad de realizar un doctorado en la Universidad de Chile, una alternativa que en mi entorno en Cuba no era muy conocida. Entre las distintas líneas de investigación que se desarrollan en el Departamento de Ciencias de la Computación (DCC), decidí orientarme hacia el área de Ciencia de Datos donde la profesora Bárbara Poblete me aceptó como su estudiante y, bajo su guía, desarrollé mi tesis doctoral sobre detección de discurso de odio en redes sociales.

La detección de discurso de odio en texto es un problema complejo de Procesamiento de Lenguaje Natural, en el que los modelos de aprendizaje automático deben ser capaces de sortear la ironía y las estrategias de oclusión léxica utilizadas para ocultar términos ofensivos. La mayoría de las soluciones y recursos disponibles en ese momento habían sido desarrollados para el idioma inglés, mientras que otros idiomas, como el español, estaban poco explorados.

En primer lugar, realizamos un análisis crítico del estado del arte en inglés, comprobando que los mejores resultados reportados estaban sobreestimados. Mostramos su limitada capacidad de generalización y las causas de este problema, como el uso de datos sesgados, lo que los hacía poco útiles en escenarios reales. Consideramos que el escenario *cross-lingual* puede entenderse como un caso extremo de generalización de modelos.

Diseñamos distintos modelos de aprendizaje automático capaces de clasificar texto en un idioma diferente al utilizado

en el entrenamiento. Para ello, propusimos conjuntos de características multilingües específicamente diseñados para la tarea. El primero consistió en un conjunto de características extraídas de la metainformación de las redes sociales, como la popularidad de la publicación, el número de veces compartida o su alcance. Asimismo, propusimos un conjunto de *word embeddings* específicos para el dominio, y mostramos que estas representaciones, aunque simples, pueden capturar información más significativa en la detección de discurso de odio que modelos más complejos. El profesor Jorge Pérez aportó su experiencia en *deep learning* para aplicar técnicas avanzadas de modelamiento del lenguaje.

Uno de los principales desafíos para validar estos modelos fue la escasez de datos etiquetados. Por ejemplo, para el español sólo existían un par de conjuntos de datos centrados en la variante de España, sin considerar otras variedades. Ante esto, demostramos la importancia de la representación multicultural y construimos el primer conjunto de datos de discurso de odio originado en Chile, con la colaboración de un equipo multidisciplinario.

Aunque hubo momentos difíciles, el doctorado en la Universidad de Chile fue una experiencia enriquecedora, llena de aprendizaje y crecimiento personal. Tuve la oportunidad de conocer a muchas personas y una nueva cultura. Siempre conté con la ayuda de los profesores del DCC y de los investigadores del Instituto Milenio Fundamentos de los Datos (IMFD), del cual formé parte durante varios años.

Reconocimiento de patrones repetitivos en imágenes de motivos de herencia cultural

Estudiante Sebastián Sepúlveda
Profesores guías Benjamín Bustos e Iván Sipirán



En este trabajo estudiamos y evaluamos herramientas computacionales para identificar automáticamente la posición en donde aparecen patrones de motivos de herencia cultural en la superficie de cerámica antigua. Para esto, utilizamos un conjunto digitalizado de 82 piezas de cerámica antigua pertenecientes al Museo Josefina Ramos de Cox, ubicado en Lima, Perú, para las cuales se anotaron manualmente la aparición de cada patrón relevante en la superficie de la cerámica. La tarea consiste en, dada la imagen de la superficie de una cerámica, detectar la aparición de cada instancia de

algún patrón relevante. Una característica que tienen estos patrones es que, usualmente, aparecen en forma repetitiva a lo largo del contorno de la cerámica.

Para resolver este problema, evaluamos distintos algoritmos y métodos de detección y segmentación de objetos como YOLOv8, Retina-Net, Mask-RCNN, Faster-RCNN, Template Matching y Segment Anything. La evaluación consistió en dos estrategias distintas: una estándar, en donde los patrones relevantes se encontraban en el conjunto de entrenamiento, y una *zero-shot*, en donde los patrones relevantes no estaban en el conjunto de entrenamiento. Los resultados de la evaluación experimental muestran que YOLOv8 obtiene la mejor eficacia en el caso de la estrategia estándar, mientras que Retina-Net obtuvo el mejor resultado en el caso de la estrategia *zero-shot*.

Los principales resultados de esta investigación fueron publicados en el ACM Journal on Computing and Cultural Heritage. Este trabajo fue parcialmente financiado por el Proyecto ANID - Fondecyt Regular – N° 1230448.

Análisis de la comprensión de modelos de lenguaje generativo en el comportamiento político chileno

Estudiante Vanessa Gaete
Profesores guías Andrés Abeliuk y Naim Bro



Los modelos de inteligencia artificial como ChatGPT han sorprendido por su capacidad para predecir comportamientos políticos en Estados Unidos. A partir de simples datos demográficos (como edad, género o nivel educativo), estos sistemas pueden generar respuestas que imitan de manera realista las opiniones humanas, lo que abre la posibilidad de reducir drásticamente los costos de las encuestas y transformar la investigación en ciencias sociales. Sin embargo, casi todos estos avances se han desarrollado y probado en contextos estadounidenses, dejando abierta una pregunta clave: ¿funcionan igual en otras sociedades?

Esta tesis es el primer estudio en Chile que evalúa si los modelos de lenguaje grandes pueden reproducir patrones de opinión y comportamiento político en la población chilena, comparando su desempeño con el obtenido en Estados Unidos. El objetivo fue determinar si estas herramientas podrían servir como métodos de predicción confiables para eventos políticos locales y qué estrategias permiten mejorar su rendimiento.

Se pusieron a prueba cuatro modelos (ChatGPT-4, ChatGPT-3.5, Llama-2-13b y Mistral) en tres escenarios: la elección presidencial chilena de 2021, el plebiscito constitucional de 2022 y las actitudes frente al aborto. Los resultados se contrastaron con experimentos equivalentes realizados con datos estadounidenses. Para ello se utilizaron encuestas públicas: las del Centro de Estudios Públicos (CEP) en Chile y las del American National Election Studies (ANES) en Estados Unidos.

Los resultados fueron claros: ninguno de los modelos logró realizar predicciones precisas en el contexto chileno, mostrando un desempeño consistentemente inferior al observado en Estados Unidos. Además, se detectaron sesgos significativos: en Chile, los modelos tuvieron más dificultad para predecir las opiniones de las mujeres que de los hombres, algo que no ocurrió en el caso estadounidense. En ambos países, las respuestas fueron más acertadas entre personas no religiosas (ateas o agnósticas) que entre quienes se identifican con una religión.

En conjunto, los resultados evidencian que, aunque la inteligencia artificial promete transformar la investigación social, su aplicabilidad fuera del contexto estadounidense sigue siendo limitada. Por lo tanto, es necesario desarrollar modelos y metodologías más inclusivos y contextualizados a la diversidad cultural y social antes de confiarles la tarea de predecir el pulso político de nuestras sociedades.

Aplicación web para diseñar bases de datos



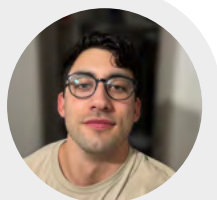
Estudiante Matías López
Profesores guías Aidan Hogan y Sebastián Ferrada

Un tema desafiante para los estudiantes de Bases de Datos (CC3201) es aprender cómo diseñar un modelo conceptual de un dominio usando diagramas Entidad–Relación (E–R). Tales diagramas capturan las entidades del dominio (p.ej., usuario, libros), sus atributos (p.ej., RUTs, nombres, títulos) y las relaciones entre entidades (p.ej., lee, califica, vende), para luego definir el esquema de la base de datos. La forma tradicional de hacer esta tarea era dibujar diagramas en papel, algo que no permitía recibir *feedback* hasta la evaluación de entrega final.

El objetivo del trabajo de título de Matías López —coguiado por Sebastián Ferrada y Aidan Hogan— fue diseñar, implementar y evaluar un sistema en línea que permite a los estudiantes de Bases de Datos definir, validar y visualizar diagramas E–R y así, recibir *feedback* inmediato sobre sus avances. Se diseñó un nuevo lenguaje y sintaxis para definir diagramas E–R, y se implementó una aplicación web que permite validar, parsear, detectar errores y visualizar estos diagramas.

Este trabajo fue publicado en el *International Workshop on Data Systems Education*, que forma parte de SIGMOD 2024: la conferencia más importante en bases de datos internacionalmente. Otra memorista, Kathleen Köhler, guiada por Matías Toro y Aidan Hogan, siguió el trabajo: ha implementado varias extensiones que incluyen la edición colaborativa en vivo de los diagramas E–R. La aplicación web (ERDoc) está disponible en línea (<https://erdoc.dcc.uchile.cl/>), ha sido usada por cientos de estudiantes del curso CC3201, y puede ser usada en otras universidades u otros contextos.

Implementación de algoritmos subóptimos para reordering en síntesis de circuitos integrados para Synopsys



Estudiante Diego Ruiz
Profesor guía Gonzalo Navarro

La industria de los circuitos integrados experimenta un rápido crecimiento y desempeña un papel fundamental en el avance tecnológico mundial. Synopsys, empresa líder en la industria, emplea la técnica de Scan Testing para verificar la funcionalidad de los circuitos integrados después de la fabricación, a través del proceso de Scan Insertion. Este proceso implica la incorporación de circuitos adicionales, incluida la conformación de una cadena entre componentes del circuito. En particular, el componente esencial de Scan Insertion, conocido como Reordering, se centra en optimizar la conectividad de esta cadena, buscando minimizar

su largo. Este desafío se aborda como una instancia particular del Problema del Vendedor Viajero, conocido por ser NP-Completo.

Aunque los algoritmos actuales de Synopsys son funcionales, tienen un amplio margen de mejora y distan del estándar académico. Tras revisar heurísticas desarrolladas en la academia para el Problema del Vendedor Viajero, identificamos el algoritmo Lin-Kernighan con mejoras de Keld Helsgaun como prometedor para implementar en Synopsys. La implementación del algoritmo muestra resultados notables en la mejora de la optimalidad y tiempos de ejecución prometedoros. Se planea que el nuevo algoritmo sea adoptado como el estándar en la herramienta de diseño de Synopsys.

Esta memoria resulta particularmente interesante por resolver un problema real, que se traduce en un problema algorítmico abstracto. Este problema, no trivial de resolver directamente, se aborda con heurísticas conocidas pero desafiantes para implementar. Finalmente, la solución desarrollada resulta exitosa en la práctica. Es un perfecto ejemplo de transferencia tecnológica de la académica a la industria.

Requerimientos geométricos de modelos hidrogeológicos de cuencas afectadas por megasequía: Caso de estudio cuenca del Limarí



Estudiante Antonio Torga

Profesores guías Nancy Hitschfeld Kahler, Pedro Sanzana y Felipe Troncoso

En el contexto del proyecto Fondecyt 1241596 dedicado al desarrollo de nuevos algoritmos para aplicaciones en ciencia e ingeniería computacional se genera una colaboración interdisciplinaria entre las ciencias de la computación e hidrología. Un primer resultado fue el software GeoLinkage, plugin del sistema de información geográfica GRASS GIS, encargado de generar automáticamente el archivo de “enlace” para la integración de un modelo hidrológico superficial, WEAP, con uno subterráneo, MODFLOW. La creación manual de estos archivos es un proceso engorroso, que toma a un modelador horas en un software GIS. Además, la creación “a mano” es propensa a errores, por lo que la asistencia computacional es crucial para generar resultados correctos que no provoquen una pérdida de flujo de agua entre los modelos.

GeoLinkage automatiza correctamente el proceso de creación de este archivo de enlace, reduciendo de horas a minutos, permitiendo además la iteración rápida de este archivo. Sin embargo, GeoLinkage no es capaz de diagnosticar erro-

res en el archivo de enlace resultante, los cuales pueden ser generados por GeoLinkage, ya que hereda estos errores de sus archivos de entrada. Con el fin de llenar ese vacío, este proyecto de memoria actualiza el software GeoLinkage y lo extiende con el nuevo módulo GeoChecker, que ejecuta un chequeo automático del archivo resultante proveyendo visualizaciones de los errores encontrados y reportes detallando su magnitud y su causa.

El error geométrico que se buscaba diagnosticar en este proyecto es la superposición de elementos del modelo subterráneo (MODFLOW) en el archivo de enlace que no estuviera respaldada por una conexión en el modelo superficial (WEAP). Un ejemplo de este tipo de errores sería la superposición de una cuenca hidrológica o sectores de riego sobre un acuífero. Generalmente esto implicaría una comunicación de flujo entre ambos sistemas, sin embargo, si no existe una conexión entre ellos en WEAP el flujo se pierde, lo que induce errores en el balance hídrico completo. La actualización llevada a cabo, que incluye el nuevo módulo GeoChecker, diagnostica estos problemas, entregando al equipo de modelación la información necesaria para poder rectificar posibles problemas de enlace.

GeoChecker es un módulo añadido a GeoLinkage, y puede ser activado desde cualquiera de las interfaces de este programa (posee una interfaz en GRASS GIS y una de línea de comandos). Este trabajo fue presentado en la sesión de Hidroinformática de la asamblea general de la European Geosciences Union 2025, y también fue presentado en el Congreso de Hidráulica 2025 de la Sociedad Chilena de Ingeniería Hidráulica.