

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 *profesorxs* guías, sobre sus trabajos de memoria y/o tesis.

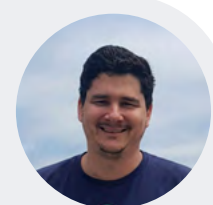
Video-Text Translation for Bridging Vision and Language

Estudiante

Jesus Perez-Martin

Profesores guías

Benjamín Bustos y Jorge Pérez



Soy Jesus Perez-Martin, Ph.D. del Departamento de Ciencias de la Computación de la Universidad de Chile. Mi formación previa es de Licenciado en Ciencias de la Computación por la Universidad de La Habana. Estuve un par de años trabajando entre la industria y la docencia universitaria, y mi interés por el aprendizaje automático aplicado me llevaron a investigar la intersección entre visión y lenguaje. Hoy, además de mi trabajo doctoral, me desempeño como Staff Data Scientist en Walmart, donde desarrollo y escalo soluciones de IA orientadas a contenido y experiencia de e-commerce.

Hay una idea simple detrás de mi doctorado: *si un sistema puede “entender” un video y contarlos con palabras —y, a la vez, transformar texto en contenido visual— entonces podemos construir puentes mucho más sólidos entre visión y lenguaje*. Eso habilita desde accesibilidad (descripciones automáticas) hasta nuevas formas de búsqueda y generación de contenido en contextos reales como el e-commerce.

Mi investigación se centra en tres líneas conectadas entre sí. Primero, *video→texto (video captioning)*: cómo generar descripciones que no sólo sean correctas, sino también naturales. En muchos modelos, el texto termina siendo genérico (“alguien hace algo”) porque es difícil capturar qué está pasando y expresarlo bien. En mi trabajo, el foco fue mejorar la forma en que el modelo combina la evidencia visual con representaciones más “lingüísticas” (semántica y sintaxis) para producir descripciones más informativas y mejor estructuradas.

Segundo, *descripción densa de video (dense video captioning)*: en vez de una sola descripción para todo el video, se busca detectar eventos (qué ocurre y cuándo ocurre) y describirlos con precisión. Aquí el desafío es doble: localizar temporalmente los segmentos relevantes y evitar redundancias, manteniendo coherencia. Parte importante de mi tesis propone hacerlo de forma más integrada, con una visión de “un solo paso”, pensando además en escenarios cercanos al tiempo real.

Y tercero, *texto→video*, con un énfasis aplicado: cómo pasar de una frase o consulta corta a un video que sea coherente con esa intención. En este punto conecto directamente

con una línea que me interesa mucho por su impacto práctico: *traducción automática neuronal (NMT)* centrada en contenido de sitios web de e-commerce. En e-commerce, el lenguaje tiene particularidades: hay marcas, atributos, restricciones de estilo, formatos, y una necesidad permanente de consistencia. En mi caso, exploré un *pipeline* donde la traducción especializada (por ejemplo, español→inglés) ayuda a condicionar mejor la generación, especialmente cuando el modelo generativo está entrenado con datos mayoritariamente en inglés.

Como suele ocurrir en investigación, no todo fue “proponer un modelo y listo”. Uno de los retos más fuertes fue evaluar de manera justa: en generación (tanto de texto como de video), las métricas automáticas no siempre reflejan lo que una persona considera “bueno”. Además, trabajar con video significa lidiar con una señal rica, larga y ruidosa: hay que aprender a seleccionar lo relevante y a representar la temporalidad de forma efectiva. También fue un desafío permanente “traducir” los avances técnicos a una historia clara: explicar la contribución con simpleza sin perder precisión, algo que en la práctica se vuelve una habilidad tan importante como la parte algorítmica.

En lo personal y profesional, mi trayectoria hoy está muy conectada con la industria. Actualmente trabajo como Staff Data Scientist en Walmart, en temas ligados a experiencia de e-commerce y modelos de lenguaje/traducción aplicados a contenido. En ese sentido, mis planes tras el doctorado son seguir profundizando esta intersección entre investigación y producto, empujando soluciones que puedan pasar de un paper a un sistema real: modelos multimodales que entiendan mejor el comportamiento del usuario, que generen contenido útil y consistente, y que mejoren la experiencia en plataformas a gran escala.

Mirando hacia atrás, lo más valioso del doctorado fue aprender un “oficio”: formular problemas, diseñar soluciones, validarlas con rigor y comunicarlas de forma convincente. Ese aprendizaje queda para siempre, y es el puente que me permite moverme con naturalidad entre la frontera científica y el impacto en aplicaciones reales.

Gradual Retrieval, Ranked Enumeration, and Lazy Evaluation of Graph Database Joins

Estudiante Asunción Gómez
Profesor guía Gonzalo Navarro



En bases de datos de grafos, el *join* es de las operaciones más costosas y, al mismo tiempo, más frecuentes. Los sistemas tradicionales usan algoritmos de *join* de a pares, lo cual se ha demostrado que es subóptimo. Nuevos enfoques consisten en hacer el *join* de múltiples tablas a la vez, lo cual puede reducir el tiempo y espacio, y ha permitido diseñar algoritmos *worst-case optimal*, es decir, óptimos en el peor caso. El uso de estructuras de datos compactas ha permitido ir más allá: no sólo crear algoritmos *worst-case optimal*, sino también reducir considerablemente el espacio requerido, sin afectar la eficiencia en las operaciones. Arroyuelo et al. (2018) diseñaron un nuevo algoritmo de *join* basado en el uso de *qdags*, una versión compacta de los árboles *quadtree*. En su algoritmo buscan optimizar el tiempo de cómputo total del *join* así como el espacio en memoria. Si bien este algoritmo es competitivo e incluso más eficiente que otros algoritmos usados en los actuales sistemas de bases de datos de grafos para calcular todo el *join*, en muchos casos nos basta con encontrar unos pocos resultados.

El objetivo de esta tesis es diseñar e implementar estrategias para disminuir el tiempo de obtención de algunas tuplas del *join*. La primera parte de este trabajo, "Gradual retrieval", busca entregar resultados lo antes posible. La segunda parte, "Ranked enumeration" busca obtener aquellos resultados más importantes ordenados de acuerdo a un peso o prioridad. La tercera parte, "Lazy qdags", tiene como objetivo implementar la versión perezosa de los *qdags*, la cual permite evaluar toda el álgebra relacional (no sólo el *join*) sobre bases de datos y, además, puede disminuir el tiempo de cómputo de los primeros resultados en algunos casos. El código del proyecto se encuentra público en www.github.com/asugomez/qdags.

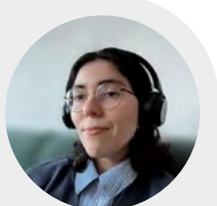
Este trabajo fue presentado en el 2do Congreso de Estudiantes de Postgrado de la Universidad de Chile (su póster quedó en tercer lugar). Un resultado en particular fue luego extendido en conjunto con otros investigadores de A Coruña, España, y publicado en el artículo:

Antonio Fariña, Adrián Gómez-Brandón, Asunción Gómez-Colomer y Gonzalo Navarro. *Cache-Friendly Compressed Boolean Matrices*. Proc. SPIRE'25, págs. 95–108.

Asunción fue una estudiante muy entusiasta, participando en todas las oportunidades que se le presentaron donde compartir su investigación, incluyendo workshops del IMFD, del DCC, y otros. Hoy en día realiza su doctorado en Francia.

An RDF Database Compact Representation for Time- and Space-Efficient Regular Path Queries

Estudiante Josefa Robert
Profesor guía Gonzalo Navarro



La tesis contribuye al problema de manejar y consultar bases de datos tipo Resource Description Framework (RDF) de manera eficiente, con un enfoque en las Two-way Regular Path Queries (2RPQs). La primera contribución de este trabajo es la implementación de una representación compacta para grafos introducida por Navarro, junto con un conjunto de métodos para extraer información valiosa del grafo. Desarrollamos, sobre esta representación, un algoritmo para resolver 2RPQs, que utiliza autómatas de Glushkov para manejar las expresiones regulares y que simula el autómata

en el grafo para identificar los caminos válidos. Para llevar a cabo esto, se utilizan en gran medida los métodos implementados anteriormente para el grafo. Evaluaciones empíricas muestran mejoras en la velocidad de procesamiento de las consultas y en el uso del espacio en comparación con soluciones existentes, particularmente para 2RPQs que involucran una única variable. En general, esta tesis presenta un enfoque competitivo para resolver 2RPQs y establece un rumbo para futuras investigaciones en el área.

Los resultados de este trabajo fueron publicados en una conferencia:

Gonzalo Navarro y Josefa Robert. *Compressed Graph Representations for Evaluating Regular Path Queries*. Proc. SPIRE'24, págs. 218–232.

Josefa fue una alumna muy independiente, que demostró gran madurez académica. En particular, se encargó ella de buena parte de la escritura del artículo.

MobilityHexVis: Analyzing Socioeconomic Urban Mobility Patterns through Small Multiples Visualization



Estudiante Alan Acevedo

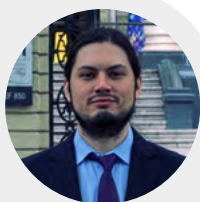
Profesores guías Eduardo Graells-Garrido y Horacio Samaniego

El estudio “MobilityHexVis: Analyzing Socioeconomic Urban Mobility Patterns through Small Multiples Visualization”, desarrollado por Alan Acevedo en su tesis de magíster en el DCC junto a Eduardo Graells-Garrido (DCC) y Horacio Samaniego (Universidad Austral de Chile), aborda la dificultad de analizar fenómenos socioespaciales con datos masivos de telefonía móvil. Entender el desplazamiento de grupos según su nivel socioeconómico es necesario para la plani-

ficación urbana, pero el volumen de datos complica su análisis. MobilityHexVis es una plataforma interactiva basada en *small multiples* y glifos hexagonales. Esta herramienta permite comparar patrones de movilidad por nivel socioeconómico y distancia. Cada hexágono usa una división bicolor para mostrar proporciones de origen y destino, e integra el índice de Gini para medir la mixtura social urbana.

La relevancia de esta investigación está en la integración de visualización de datos y ciencia urbana, con una publicación en la revista “Environment and Planning B”. El sistema, aplicado en Santiago, identificó centros de empleo y brechas en la movilidad según el estrato social. El proyecto destaca por su enfoque multidisciplinario, donde la computación aporta evidencia para el diseño de ciudades inclusivas. Esta labor, vinculada a un proyecto Fondecyt Regular, valida la calidad del postgrado en el DCC. El trabajo conjunto entre estudiantes y académicos genera tecnología con impacto social y científico, al proponer herramientas computacionales para resolver problemas urbanos complejos.

Desarrollo de lineamientos para el deployment personalizado del producto de software TIMA



Estudiante Cristóbal Carmona

Profesor guía Sergio Ochoa

El objetivo de esta tesis fue generar lineamientos específicos para mejorar el proceso de *deployment* personalizado del software TIMA (TIMining Aware), un gemelo digital 3D de apoyo a faenas mineras. Cada faena minera es particular y, debido a esto, cada proceso de *deployment* de TIMA se vuelve particularmente complejo. La complejidad surge a partir de la cantidad y diversidad de actividades que el gemelo digital debe apoyar, y de la variedad de datos que debe manejar en cada faena. Esta heterogeneidad genera alta in-

certidumbre e ineficiencia, y también una alta dependencia de personal calificado.

Los lineamientos generados en el marco de esta tesis estuvieron enfocados en abordar esta problemática. Particularmente, se definieron guías específicas para abordar tres aspectos clave: 1) automatización de subprocesos “cuello de botella”, 2) testing automatizado de funcionalidades, y 3) uso de buenas prácticas para la definición de conectores y procesos.

Los resultados de aplicar estos lineamientos fueron exitosos, facilitando el proceso de *deployment* personalizado de TIMA según las características de cada faena minera. También redujo significativamente la dependencia de personal calificado para llevar a cabo este proceso. Aunque este trabajo logró los objetivos planteados, la diversidad de faenas mineras es tan amplia que hace que se requieran más pruebas en producción para robustecer las capacidades de este gemelo digital.