

Khipu

Centro para la Investigación de Bases de Datos

Uno de los primeros pueblos en habitar los Andes sudamericanos, los Incas, desarrollaron y utilizaron un medio sofisticado para almacenar información administrativa, el khipu. Un khipu está formado por piezas de cuerdas que están amarradas de algunas formas particulares, y que implementan un modelo complejo de datos (un khipu se asemeja a una estructura de árbol, diríamos en terminología moderna). El khipukamayúq era el experto en la creación, actualización e interpretación de la información codificada en un khipu, algo así como el actual administrador de una base de datos. Desafortunadamente, los invasores españoles del siglo XV ordenaron la destrucción de los khipu y cazaron a los khipukamayúq, destruyendo temporalmente una rica tradición en el almacenamiento de información.

El almacenamiento y manipulación de datos ya significaba un importante desafío para

las antiguas civilizaciones. Hoy en día, las tecnologías digitales de muy bajo costo, permiten almacenar y difundir información en cantidades inimaginables en el pasado. Un ejemplo cotidiano de esto es la Web, que funciona como un gigantesco repositorio de datos. Además, prácticamente todos los grandes proyectos de investigación dedican importantes esfuerzos a la recopilación y análisis de datos. Basta citar ejemplos como el proyecto del Genoma Humano en biología, o proyectos de astronomía como el Observatorio Virtual. Es así como en el mundo moderno el manejo eficiente de información no sólo es un desafío sino más bien una necesidad.

El centro Khipu, inspirado en el trabajo pionero de nuestros antepasados sudamericanos, tiene como objetivo desarrollar, desde nuestra región, investigación de impacto mundial en el área de bases de datos y procesamiento de información. El Khipu



Marcelo Arenas

Profesor Asistente, DCC, Pontificia Universidad Católica de Chile. Doctor en Ciencia de la Computación, University of Toronto, Canadá. Licenciado en Matemáticas, Magister en Ciencias de la Ingeniería e Ingeniero Civil de Industrias mención Computación, Pontificia Universidad Católica de Chile.
marenas@ing.puc.cl



Jorge Pérez

Estudiante de Doctorado en Ciencia de la Computación, Pontificia Universidad Católica de Chile. Ingeniero Civil en Computación y Magister en Ciencia de la Computación, de la misma Universidad.
jperez@ing.puc.cl

nace en el año 2008 como un esfuerzo conjunto de investigadores nacionales por crear un centro de investigación de excelencia, donde se consoliden los logros ya alcanzados, y se generen las interacciones necesarias para potenciar e incrementar la investigación de alto nivel en nuestro país. Hoy Khipu convoca a profesores, investigadores, alumnos y colaboradores de distintas regiones de Chile.

En este artículo resumimos la concepción de Khipu como centro de investigación en bases de datos y manejo de información, quienes lo integramos, los logros que hemos alcanzado y nuestra visión hacia el futuro.

KHIPU: UN ESFUERZO

Nombre	Afiliación actual	Estudios de doctorado	País	Año
Marcelo Arenas	PUC Chile	University of Toronto	Canadá	2005
Pablo Barceló	U. de Chile	University of Toronto	Canadá	2006
Loreto Bravo	U. de Concepción	Carleton University	Canadá	2007
Benjamín Bustos	U. de Chile	University of Konstanz	Alemania	2006
Mónica Caniupan	U. del Bío-Bío	Carleton University	Canadá	2007
Claudio Gutiérrez	U. de Chile	University of Wesleyan	USA	1999
Carlos Hurtado	U. Adolfo Ibáñez	University of Toronto	Canadá	2002
Andrea Rodríguez	U. de Concepción	University of Maine	USA	2000

A NIVEL NACIONAL

Durante los últimos 10 años se ha producido un incremento sustancial en el número de académicos chilenos que han realizado sus estudios de doctorado en el extranjero en el área de bases de datos. Esto permite tener hoy una masa crítica suficiente para instalar en Chile un centro de investigación en el área, que pueda tener impacto a nivel mundial.

El núcleo de Khipu esta formado por ocho investigadores asociados. Como muestra la siguiente tabla, todos estos investigadores obtuvieron sus doctorados en universidades norteamericanas y europeas. Es importante destacar que estos investigadores no sólo están concentrados en Santiago, haciendo de Khipu un esfuerzo a nivel nacional.

Para nuestro centro es también fundamental la formación de alumnos de pregrado, magíster y doctorado en el área de bases de datos. Durante los últimos años, los investigadores de nuestro centro han participado en la formación de un importante número de alumnos, y en la actualidad contamos con 10 alumnos de magíster y 6 alumnos de doctorado.

DESARROLLANDO INVESTIGACIÓN DE IMPACTO MUNDIAL

Una de las premisas de nuestro grupo es generar conocimiento de relevancia mundial. Por esto, parte de nuestro esfuerzo ha estado enfocado en desarrollar áreas de vanguardia y en publicar nuestros resultados en las conferencias y revistas más prestigiosas en nuestra área, y también en computación en general. De hecho, desde el año 2000 el núcleo de investigadores de Khipu ha publicado más de 160 artículos en conferencias y revistas internacionales.

La investigación que realiza nuestro grupo abarca diversas áreas sobre el procesamiento de información. Algunas de las áreas que se están desarrollado son: bases de datos multimedia, bases de datos geoespaciales,



El centro Khipu, inspirado en el trabajo pionero de nuestros antepasados sudamericanos, tiene como objetivo desarrollar, desde nuestra región, investigación de impacto mundial en el área de bases de datos y procesamiento de información.



integración e intercambio de datos, manejo de datos en la Web Semántica, manejo de información inconsistente. A modo de ejemplo, detallamos algunas de estas líneas de investigación y parte de los logros obtenidos.

- **Manejo de información inconsistente.** Usualmente a la información almacenada en una base de datos se le exige que satisfaga ciertas restricciones de sanidad (por ejemplo, que una persona tenga sólo un número de RUT asociado). En muchas aplicaciones estas restricciones se violan dejando a la base de datos en un estado *inconsistente*. La pregunta entonces es cómo poder razonar sobre datos con este tipo de inconsistencias, de manera eficiente y limpia. Esta línea de investigación fue iniciada por miembros de nuestro grupo quienes desarrollaron los fundamentos teóricos necesarios para estudiar el problema. Hoy diversos grupos de investigación a nivel mundial están desarrollando las distintas aristas del tema basándose en los resultados generados por nuestro grupo, aplicando también los resultados en áreas como bioinformática y sistemas de información geoespaciales.
- **Manejo de datos multimedia.** Las bases de datos multimedia están compuestas por imágenes, gráficos en 2 y 3 dimensiones, audio, y videos digitales, en conjunto con datos de texto. En la actualidad existen diversas aplicaciones que necesitan de datos multimedia, como por ejemplo aplicaciones de

manufactura, arte, y cine digital. Una de las áreas desarrolladas por integrantes de nuestro grupo es la *búsqueda por similitud* en bases de datos multimedia. Esta técnica permite obtener mejoras considerables en la efectividad de las búsquedas en este tipo de bases de datos. Muchos desafíos existen en esta área, como por ejemplo, el diseño de nuevos algoritmos y estructuras de índices adecuadas para acelerar las búsquedas de multimedia.

- **Manejo de datos en la Web Semántica.** La Web Semántica es una extensión de la Web tradicional, pensada en la inclusión de metadatos que permitan, tanto a personas como a máquinas, entender el *significado* y las *relaciones* de la información almacenada en la Web. Una de las líneas de investigación de nuestro grupo tiene que ver con desarrollar las herramientas teóricas necesarias para poder manipular los metadatos semánticos en la Web. Uno de los principales logros alcanzados en esta área fue la formalización de la semántica y subsiguiente estudio del lenguaje de consulta SPARQL, que hoy es el lenguaje estándar para consultar los metadatos de la Web Semántica.

Parte del impacto de la investigación generada por Khipu se puede medir por los reconocimientos internacionales que han obtenido nuestros miembros. Dos artículos escritos por miembros de Khipu han sido reconocidos con el *Best Paper*

Award (premio al mejor artículo) en la conferencia más importante de Teoría de Bases de Datos (*ACM Principles of Database Systems*). Adicionalmente, tres de nuestros artículos han sido reconocidos con el *Best Paper Award* en las conferencias más importantes del área de Web Semántica (*International Semantic Web Conference* y *European Semantic Web Conference*). Nuestros alumnos han obtenido también importantes premios. Uno de ellos se convirtió recientemente en el primer alumno chileno en recibir la *Microsoft Research Fellowship*, el reconocimiento académico más importante que entrega Microsoft a alumnos de doctorado.

NUESTRA VISIÓN DEL FUTURO

Las habilidades para almacenar, procesar y analizar información de manera eficiente son fundamentales para el desarrollo de nuestra sociedad. Es por esto importante contar con centros de investigación que puedan proponer soluciones a estas problemáticas. Chile tiene hoy una oportunidad de contar con uno de estos centros, por el aumento significativo en el número de investigadores en el área de bases de datos que residen en nuestro país. Es por esto que decidimos crear Khipu, un esfuerzo nacional para formar un centro de investigación en el área de bases de datos y manejo de conocimiento. En el futuro cercano, esperamos que Khipu se transforme en un referente latinoamericano en el área, tanto por la investigación en bases de datos, como por la formación de alumnos de pregrado, magíster y doctorado que colaboren con el desarrollo de nuestra región. También esperamos que nuestra investigación continúe teniendo impacto a nivel mundial, y que nuestro centro se convierta en un lugar a visitar al momento de buscar nuevas y mejores tecnologías de bases de datos.^{BITS}

Más información sobre nuestro centro puede ser obtenida en www.khipu.cl