

CONFERENCIAS:



LA-Web 2008

Por Mari-Carmen Marcos
Universitat Pompeu Fabra
Barcelona-España

Los días 28, 29 y 30 de octubre de 2008 la ciudad de Vila Velha, estado de Espírito Santo de Brasil, fue el punto de encuentro de la última edición de LA-Web, un congreso internacional con sede en Latinoamérica apoyado por el comité de la IW3C2 de la International World Wide Web Conference y que por sexta vez reúne a investigadores de varios continentes.

Vila Velha, unida a Vitória (capital de Espírito Santo) por medio de varios puentes y presidida por un monasterio en lo alto de la colina, nos recibió con clima tropical y una larga playa que sólo algunos, en las últimas horas antes del regreso al aeropuerto, pudimos disfrutar. El comité de organización local, en coordinación con la conferencia WEBMedia 2008, celebrada paralelamente a LA-Web y con la que compartimos algunos ponentes, lo tenía todo preparado para nuestra llegada y estuvo atento a los detalles durante los tres días que duró el evento.

LA-Web se abrió con la conferencia invitada de Manuel Pérez-Quñones, especialista en Interacción Persona-Ordenador en el Center for Human-Computer Interaction de Virginia Tech, Estados Unidos. Su charla abordó un tema de alta actualidad: ecosistemas personales de información, entendidos como el conjunto de dispositivos con los que cada persona interactúa para satisfacer sus necesidades de información, por ejemplo el notebook, el celular, el iPod, la webcam, etc. El tema resultó muy atractivo y, de hecho, todos nos sentimos identificados con los problemas que el expositor presentó con ejemplos muy acertados. Pérez-Quñones apuntó a la situación que vivimos hoy en día: utilizamos distintos dispositivos para obtener información e interactuamos con cada uno por separado, sin posibilidad de poder establecer una relación entre ellos. Mientras los diseñadores se centran en imitar las funcionalidades y el aspecto de los nuevos dispositivos con respecto a los ya existentes, Pérez-Quñones reclamó la necesidad de generar ecosistemas de información donde, al igual que ocurre en los ecosistemas biológicos, se atiende a las relaciones entre los distintos dispositivos para satisfacer las necesidades de información de los usuarios. En su charla presentó múltiples ejemplos de la falta de intercomunicación que se da entre dispositivos ubicuos para poner de

relieve la importancia de que los diseñadores de estas plataformas consideren el ecosistema del usuario con todos los dispositivos que utiliza como un todo, y de esta manera lograr una mejora en la experiencia de uso.

A la charla de Manuel Pérez-Quñones sucedieron otras a lo largo de la mañana en las que los temas estuvieron relacionados con la búsqueda de información y la minería de datos. La tarde estuvo dedicada al tutorial que impartió Ricardo Baeza-Yates, Vicepresidente de Yahoo! Research, con el título "Introducción a la Minería de Web" y que finalizó la tarde del día siguiente. En total fueron 3 horas de clase con uno de los mayores expertos en minería, un tiempo que se hizo corto para todo lo que se puede aprender sobre este tema. El profesor Baeza-Yates dejó clara la necesidad de aplicar técnicas de minería en los distintos tipos de datos de la Web: aquellos que proceden de los contenidos (textuales y multimedia), los relativos a la estructura de hiperenlaces y en los que se generan con el uso que se hace de la Web (las consultas en los buscadores y la navegación de los usuarios, entre otros). Todos estos datos proporcionan una fuente de información muy potente porque nos permiten conocer cómo es la Web y cómo se utiliza, y esta información puede y debe utilizarse para mejorar la Web y hacerla más útil a las personas. La charla estuvo apoyada por muchos gráficos con distintas técnicas de visualización que sólo en sí mismos ya resultan interesantes, además de aclarativos para entender la estructura de la Web. El tutorial centró buena parte de su atención en la Web 2.0 y en cómo utilizar la información que los usuarios generan para mejorar el diseño de los sitios Web e incluso para predecir la intención de los usuarios en sus consultas y así mejorar los resultados de los buscadores.

En la siguiente jornada asistimos a una buena cantidad de charlas sobre temáticas variadas; la mañana comenzó con varias conferencias sobre TIC para el desarrollo, continuó con el tema de la Web semántica y finalizó con ponencias sobre ingeniería Web y comunidades. Esa tarde tuvimos como ponente invitado a Fabrizio Silvestre, investigador del ISTI en el CNR de Pisa, Italia, que dio una charla sobre cómo el análisis de las búsquedas realizadas en el pasado nos ayudan adoptar mejoras para el futuro usando técnicas de minería de datos.

Para demostrarlo nos presentó varios ejemplos de resultados que han obtenido en los últimos años con el uso de técnicas de *caching*, es decir, almacenar en memoria rápida resultados que se usarán frecuentemente. La tarde finalizó con la segunda parte del tutorial de Ricardo Baeza-Yates.

El último día, la conferencia invitada estuvo a cargo de Mariano Consens, profesor en la Universidad de Toronto y especialista en gestión de datos. Su charla fue más hacia la Web Semántica, pues trató sobre la gestión de datos que están vinculados en la Web refiriéndose a cómo distintos conjuntos de datos están unidos entre sí por enlaces en RDF que se identifican por URIs. La red que se forma con estos vínculos puede ser explotada para mejorar las capacidades de los navegadores y de los buscadores y para crear nuevos escenarios, nuevas aplicaciones y nuevos *dashups*. Consens puso énfasis en algunos de los retos que supone esta concepción de la Web: la gestión de los inter-enlaces (*interlinks*), los metadatos, y la vinculación entre conjuntos de datos. Estos retos han sido abordados por LinkedMDB, una base de datos de películas que tiene un gran número de inter-enlaces a otras bases de datos y a otros sitios Web sobre películas.

A la charla de Consens siguió una sesión sobre la investigación sobre la Web en Latinoamérica. Hubo representación de Chile, Brasil y Argentina, y un interesante debate al final de las ponencias. Con esta sesión se puso fin a tres días intensos. Las actas han sido publicadas en Cd-rom por la IEEE Computer Society Press y las presentaciones de los ponentes invitados están disponibles en abierto en el sitio web de LA-Web 2008: <http://www.cwr.cl/la-web2008/>

El presidente de la conferencia y coordinador del comité de seguimiento de LA-Web, Ricardo Baeza-Yates, nos invitó a participar en el próximo, cuya sede será en México. No dudamos que LA-Web 2009 será de nuevo un éxito, aunque Claudine Badue, Elias Oliveira y Alberto Ferreira, comité local en Vila Velha, han puesto el listón bien alto. Además de la calidad de las charlas, la *moqueca capixaba* y las *caipirinhas* nos dejaron un buen sabor de boca de LA-Web 2008. BITS

III Alberto Mendelzon Workshop on the Foundations of Data Management

Arequipa, Perú, 12 al 15 de mayo, 2009
<http://db.cs.ualberta.ca/amw09/>

Pablo Barceló
 DCC, Universidad de Chile

Este workshop es una iniciativa de la comunidad latinoamericana de investigadores identificados con las áreas de bases de datos, manejo de información y la Web, a la cual nuestro amigo Alberto Mendelzon contribuyó con enorme generosidad. Nuestro objetivo es establecer en el cono sur una instancia científica periódica de alto nivel en los aspectos fundamentales del área. Creemos que esta es una excelente forma de mantener viva la memoria de Alberto, y a la vez de incrementar y solidificar la investigación científica en la región. Como en los dos eventos anteriores, otra de nuestras principales motivaciones es ayudar a que los alumnos latinoamericanos interesados en los fundamentos de bases de datos y la Web (especialmente los alumnos de máster y doctorado) tengan la oportunidad de interactuar con algunos de los más destacados expertos mundiales del área.

Este año hemos decidido realizar el workshop en Arequipa, Perú, desde el 12 al 15 de mayo. El primer workshop (12 al 14 de mayo) consistirá de 3 tutoriales invitados, algunas charlas invitadas, más los artículos aceptados por el comité. Luego habrá un workshop de alumnos graduados (en español, el 15 de mayo). Este también tendrá un par de tutoriales invitados.

Arequipa no es sólo la segunda ciudad más poblada del Perú y un importante centro turístico, sino también el lugar donde se encuentran algunas de las más reconocidas universidades peruanas que realizan docencia e investigación en ciencia de la computación. A través de este workshop esperamos que una buena parte de la gran cantidad de los alumnos peruanos que estudian ciencia de la computación se interesen en los temas relacionados con las bases de datos, manejo de información y la Web.

Entre los atractivos turísticos de Arequipa se puede mencionar su localización, en las contrafuertes de varios volcanes pertenecientes a los Andes Occidentales. Particularmente impresionante es el Cañón del Colca, el más profundo del mundo, a unas horas de Arequipa. A lo largo del valle se pueden ver innumerables construcciones prehispánicas, y en lo alto del cañón es usual observar cóndores. Además, el centro de la ciudad de Arequipa, declarado Herencia Cultural por la UNESCO el año 2000,

combina iglesias barrocas con mansiones de la época colonial. Destacan la catedral, un imponente edificio del siglo XVI, y el Monasterio de Santa Catalina. La ciudad es además famosa porque la mayoría de sus casas está construida en piedra volcánica de color blanco, por lo cual a veces recibe el nombre de "Ciudad Blanca". Además Arequipa está a corta distancia de dos de los principales centros turísticos del Perú: Cusco y Machu-Picchu, y Nasca.

ALBERTO MENDELZON

Pero, ¿quién es Alberto Mendelzon y por qué el workshop tiene su nombre? Alberto Mendelzon, sin lugar a dudas, es uno de los más influyentes científicos en el área de las bases de datos. Esto, ciertamente, no es un logro menor considerando la gran cantidad de "superestrellas" que han realizado investigación en el área. Por ejemplo, de los 25 científicos de la computación con mayor índice-*h* (*h-index*) al menos cinco han realizado investigación prolongada en el área: Héctor García-Molina, Jeff Ullman, Christos Papadimitriou, Alon Halevy y Jennifer Widom. Además, García-Molina y Ullman son los dos científicos de la computación con mayor índice-*h* (89 y 87 ¡respectivamente!). El índice *h* de Alberto es también impresionante: 43 a la fecha, lo que significa que 43 de sus artículos tienen al menos 43 citaciones. Además, cada uno de los cinco artículos más influyentes de Alberto tiene al menos 400 citaciones, mientras que su artículo más citado tiene 640.

Pero claro, todo esto son sólo fríos números, y a nosotros nos gustaría transmitir en unos pocos párrafos, al menos, una parte de lo que fue Alberto como investigador y persona. Primero, los datos biográficos: Alberto nació en Buenos Aires, Argentina, casi 58 años atrás. Recibió su grado de Ph.D. en Princeton en 1979, y tras pasar un año como pos-doc en IBM Watson las emprendió hacia la Universidad de Toronto, lugar que nunca dejaría. Alberto murió en junio de 2005 producto de un cáncer.

Quizá el trabajo más famoso de Alberto es el que realizó durante sus años de Ph.D. (supervisado por el mismísimo Ullman), donde delineó de manera definitiva el posterior desarrollo de la teoría y la práctica de las bases de datos

relacionales. El aporte más significativo de este trabajo es haber presentado por primera vez un método –llamado *chase*– para verificar implicación entre restricciones de integridad. Tal método se halla actualmente diseminado a través de toda la literatura en bases de datos, y es utilizado diariamente de forma directa o indirecta por cada persona que diseña bases de datos. Es además el método utilizado por los sistemas comerciales para verificar la consistencia y corrección del diseño de los datos. Pero ciertamente el aporte de Alberto a las bases de datos no se detuvo ahí: sus artículos han delineado e influido fuertemente en muchas otras áreas, incluyendo revisión de conocimiento y actualización de bases de datos, lenguajes de consultas para bases de datos de grafos, consultas en la Web, utilización de vistas en integración de datos, etc. En 2007, ACM PODS, la conferencia más prestigiosa en el área de los fundamentos de las bases de datos, estableció el Alberto Mendelzon *test-of-time-award* en memoria de Alberto.

Por cierto, Alberto no sólo fue un gran científico sino también una persona muy especial para todos aquellos que compartieron con él. Famoso por su modestia y sentido del humor, y admirado por su generosidad, no puede ser mejor descrito que por sus propios colegas y alumnos. Entre ellos me gustaría destacar el testimonio de Moshe Vardi, otra "superestrella": "Alberto no sólo fue tremendamente respetado por sus colegas debido a su trabajo científico, sino también extremadamente querido por sus modos sencillos, sentido del humor y cálida personalidad. Aunque sólo nos veíamos una o dos veces al año, para mí no era sólo un buen colega sino que también un amigo. Lo extrañaré mucho".

Alberto fue particularmente generoso con la comunidad científica latinoamericana dedicada a la computación, de la cual él sin duda se sentía parte. Cada vez que pudo, Alberto visitó países de la región (Argentina, Brasil, Chile, entre otros) para establecer colaboración científica o tan sólo para dictar una charla. Además Alberto ayudó a una larga lista de estudiantes y profesores a llegar a Toronto, ya sea como alumnos de máster, doctorado o posdoctorado. Esta es la razón por la que el workshop recibe su nombre. BITS