

Domingo 25 de febrero 2007 | Actualizado a las 18:33h

Internet

Entrevista a Ricardo Baeza, director del Laboratorio Yahoo!

"En el futuro vamos a tener un ángel de la guarda digital"

23/02/2007

Karma Peiró | Yahoo! es el portal web más consultado de Internet. Cada mes lo visitan 500 millones de personas para buscar información, aprovecharse de alguno de sus servicios, consultar el correo o crear comunidades virtuales en red. Desde octubre del 2006, Barcelona cuenta con el único laboratorio de investigación que la multinacional norteamericana tiene en Europa. Su director, Ricardo Baeza, está convencido que los buscadores en Internet van a ser siempre necesarios y prevé un futuro donde accederemos a nuestro mundo virtual con la voz desde cualquier lugar.

-Todo el mundo conoce el buscador Yahoo! Pero, ¿cómo es por detrás? ¿Cómo hace para arrojarnos información en segundos?

- Es una cantidad de tecnología inmensa, de búsqueda, de minería de datos, de detección de spam web, de filtrado de contenidos sensibles, etc. Es complicadísimo. Todo son algoritmos soportados por decenas de miles de ordenadores, distribuidos por todo el mundo, para que sea muy rápido.

- ¿Qué investigan en el Laboratorio Yahoo! de Barcelona?

- Los objetivos principales de este centro son tres: atraer talento europeo, mejorar la búsqueda de información en la web y aprovechar todos los datos para encontrar nuevos servicios, lo que se llama "minería de datos".

- O sea, que como los mineros se dedican a excavar pero entre millones y millones de datos...

- Si. Excavamos el contenido, la estructura (que son los enlaces entre las páginas) y el uso que hace la gente en nuestros sitios Web. Con la minería de datos también localizamos a los que se dedican al spam de web y crean páginas con contenido falso o enlaces mentirosos. Excavar el uso también sirve para saber las consultas más frecuentes que se le hacen al buscador y estudiar la calidad de los resultados.

- ¿Y qué buscamos más en España?

- Las consultas de aquí no son muy distintas de las de otros países.

- Los grandes temas son sobre política, ciencia, futuro..

- No, es mucho más simple. Software gratis, imágenes, vídeo, sexo, etc. Básicamente, se sigue una ley de Zipf o "Ley de potencias". Es decir: hay mucha gente preguntando pocas cosas y mucha gente preguntando cosas distintas. Lo que tiene que ver con información especializada pertenece a esa categoría de preguntas distintas.

'Granjas de enlaces' para posicionarse

- Los internautas estamos hartos del spam de correo, ¿cómo es el de web?

- Igual que el del correo, utilizan técnicas poco éticas. Hay páginas que tienen palabras clave que

atraen a la gente en una búsqueda pero cuyo contenido no tiene nada que ver. Por ejemplo, puedo buscar "MP3" pero la página resultante es de subastas. Son páginas engañosas. También están las "granjas de enlaces" ("link farms"): sitios web enlazados entre ellos para quedar mejor ubicados en los buscadores. Los casinos y lugares de apuestas suelen usar "granjas de enlaces".

- Y cuando encuentran una de estas "granjas" ¿qué hacen?

- Darle menos peso a esas páginas en el buscador y valorarlas con un contenido menos rico que el de un diario, por ejemplo.

- ¿Qué otras técnicas utiliza el internauta para posicionar su página?

- Usar texto transparente. Se escribe en blanco sobre un fondo blanco. El buscador es ciego, no entiende de colores, sólo lee el lenguaje HTML. Las páginas web accesibles para personas con discapacidad visual son muy buenas para ser procesadas por los buscadores. Otro truco es escribir muy pequeño, en cuerpo 5, que se lee como una línea pero el buscador la registra igualmente.

- Después de todo lo dicho... los buscadores no son perfectos...

- No, en absoluto. Pero considerando el tamaño de la Web y que las consultas que recibe un buscador no son muy precisas, los resultados son muy buenos.

- ¿Lo serán alguna vez?

- Si la gente dijera siempre la verdad...

- Eso es imposible... ¿alguna idea para mejorarlos?

- Si la web semántica existe algún día y las páginas tienen información temática, se podrían mejorar. Si además hubiera alguna forma de certificar que el contenido es real y verdadero, entonces serían perfectos. Pero actualmente los buscadores están luchando con un adversario enorme: la gente que intenta engañarlos y contra muchos otros que, por ignorancia, ponen información falsa, no actualizan enlaces, etc. Hay miles de razones por las cuales no pueden ser perfectos que dependen de las personas.

El futuro cercano

- Se dice que cuando llegue la web semántica o inteligente, será el fin de los buscadores. ¿Qué piensa?

- Con la Web semántica cambiarán muchas cosas pero los buscadores seguirán siendo necesarios. Se harán consultas y se obtendrán respuestas muy diferentes. Si alguien quiere irse de vacaciones, el internauta consultará: "Me quiero ir a las Islas Canarias. Organízamelo". Y el buscador localizará el hotel, los billetes de avión, el coche de alquiler, un restaurante para ir a cenar, etc. En un maremagnum de información se necesita algún mecanismo directo para acceder a los datos y esos son y serán los buscadores.

- ¿Cuándo serán tan inteligentes?

- Ya estamos investigando la integración de información, es decir, juntar todas las respuestas de una consulta similar de manera coherente. De aquí a unos años serán inteligentes. Quizás no sean tan sofisticados pero empezarán a funcionar así.

- ¿Qué podremos hacer en breve con la tecnología que ahora no imaginamos?

- Internet estará en todas partes, podremos interactuar con nuestra información digital desde cualquier lugar y sin usar ningún dispositivo. Con la voz le pediremos a una pared que despliegue una pantalla de ordenador y accederemos a nuestro mundo virtual. Tecnológicamente ya se puede hacer, es un problema de costos. Imagina que me olvidé comentarte algo. Le podría decir a mi asistente digital: "Envíale a Karma tal documento". Eso va a ser el futuro, vamos a tener nuestro ángel de la guarda digital.

- Lo ve como algo positivo. Hay gente que se agobia con la información que tenemos...

- Sólo se usará la que necesitemos en cada momento. Primero se tienen que resolver los

problemas de seguridad y confidencialidad, pero la información digital será tuya y nadie más que tú, con tu voz, podrá acceder a ella. Esto aumentaría mucho la productividad laboral. Quizás veamos a gente hablando sola por la calle pero estarán haciendo cosas. Lo mismo para el entretenimiento. ¿Por qué tengo que ocuparme en meter en mi reproductor MP3 la música que me gusta escuchar? Bastará pasear por una calle y pedir lo que quiero que suene.

- Además de superar el coste tecnológico, hará falta un cambio de mentalidad para adaptarse a esta vida ...

- Sí, hace falta un cambio de mentalidad. Internet se utiliza de forma masiva desde hace muy poco tiempo. Y no es lo mismo ser niño con internet que sin ella. En 20 años, los niños de finales del siglo XX protagonizarán un cambio revolucionario cuando tengan 30 o 40 años.

LA VANGUARDIA, el diario más vendido en Catalunya. Copyright La Vanguardia Ediciones S.L.
All Rights Reserved - Aviso Legal - Contacte con nosotros - Publicidad

Enlaces Grupo Godó

lavanguardia.es
elmundodeportivo.es
rac1.cat

rac105.cat
8tv.cat
servijob.com

lavanguardiainmo.com
diaridelsestudiants.com

vanguardiadossier.com
lavanguardiatiendas.com

magazinedigital.com
publipressmedia.com