

Sistemas de información

Nuevos motores para el siglo XXI

Por Mari-Carmen Marcos

A PRINCIPIOS de septiembre Isidro Aguillo nos comunicaba en *IweTel* la reciente aparición de dos nuevos motores “con bases de datos genuinamente propias y con sistemas originales de recuperación y presentación de resultados”: *Wisnut* y *Teoma*.


Wisnut, “el sabio”

Fue creado por **Yeogirl Yun** (cofundador de *mySimon*, un servicio web que compara precios de venta en la Red) en 1999, si bien acaba de ser lanzado a principios de septiembre.

Cuenta con 1.500 millones de páginas indexadas —dato de 29/09/01—, lo que lo sitúa entre los mayores buscadores actuales. Pero no sólo es en esta cifra en lo que se basan sus creadores para esperar el éxito del motor, sino sobre todo en su sistema de ranking de los resultados que responden a la consulta y que coincide en parte con el usado por *Yahoo!* o *Google* (éste último llama a su técnica *Pa-*

geRank). El sistema es muy similar a algunas técnicas bibliométricas: mide las citas recibidas y considera que, cuantas más haya, más relevante puede ser esa página, especialmente si proceden de sitios bien calificados, pues se considerará que la página en cuestión es más interesante que si recibe los enlaces de sitios menos notorios. Esta calificación se calcula lanzando todas las peticiones posibles (tomando como consulta las palabras de la base de datos de términos indexados) a todas las páginas web recogidas; para ello se utilizan las bases de datos de enlaces y de términos indexados. El resultado de relevancia o peso de cada par “palabra-página” también se almacena.

De todas formas, *WiseNut*, no satisfecho con las técnicas habituales para establecer la relevancia de una página web a la búsqueda (preferencia a páginas con los términos de consulta en los primeros párrafos, en el título o en las etiquetas meta) añade al procedimiento ex-



1,495,332,308 Web pages and counting!

WiseNut
Search, Easyly

Search For :
ergonomia cognitiva [Go!]

Get preferences | Try WiseSearch

WiseGuide categories for "ergonomia cognitiva" [What's WiseGuide?]

En: es (it) [search this] Other: 588 Results: 62 documents found

1. Sociedad de Ergonomistas de México A.C.
... SEMAC ERGONOMIA COGNITIVA Actualizado al 20 de octubre de 2000 ERGONOMIA COGNITIVA Los ergonomistas...
... conocidos hasta el momento. Esta área de la **ergonomía** tiene gran aplicación en el diseño y evaluaci...
[See 2 more pages from this site]
<http://www.semac.com.mx/A2/ergonomia/cogniti.html> [Sneak a Peek]
2. Curso de HCI y Ergonomía Cognitiva
... y Ergonomía Cognitiva To: "Distribución" <cicd@vc.ehu.es> Subject: Curso de HCI y Ergonomía...
... Cognitiva From: OCdeUPV/EHU/ok@v-biarp@wv.ehu.es> Date: Thu, 02 Nov 2000 10:04:22 +0100...
[See 4 more pages from this site]
<http://www.vc.ehu.es/tablon/vctvria/mrg00680.html> [Sneak a Peek]
3. Submarino
... Submarino: Euro: Pts € 13 900 83 54 Ergonomía Cognitiva CAÑAS, JOSÉ J. & WAERNS, YVONNE Revisa en...
... forma exhaustiva el concepto de la **ergonomía** con un doble objetivo: servir como libro de texto a...
<http://www.submarino.com/books.asp?Query=NextLevelPage&ProdTypeId=1&CatId=124> [Sneak a Peek]
4. Quiénes Somos!!
... a, Neurociencias, Psicología Conductual y Cognitiva País: Argentina Daissy Ramirez Asesora Fundaci...
... a Asesor Biblió Especialista en Ergonomía (Ergonomía Cognitiva) y Psicología. Prevencionista...
[See 2 more pages from this site]
<http://www.psycologia.com/quien.htm> [Sneak a Peek]

plicado la característica de tener en cuenta el contexto: *WiseRank*. La novedad reside en que tiene muy en cuenta la estructura hipertextual de la web. De esta forma influye tanto la información intrínseca de cada página (la que contiene en el título, el texto y las etiquetas meta) como la extrínseca (el texto que funciona como enlace, las palabras que lo rodean, la que proporcionan las páginas vinculadas a ella y el título y las etiquetas meta de las páginas a las que enlaza). Dos tipos de información que se complementan para mejorar el ranking de relevancia de los resultados.

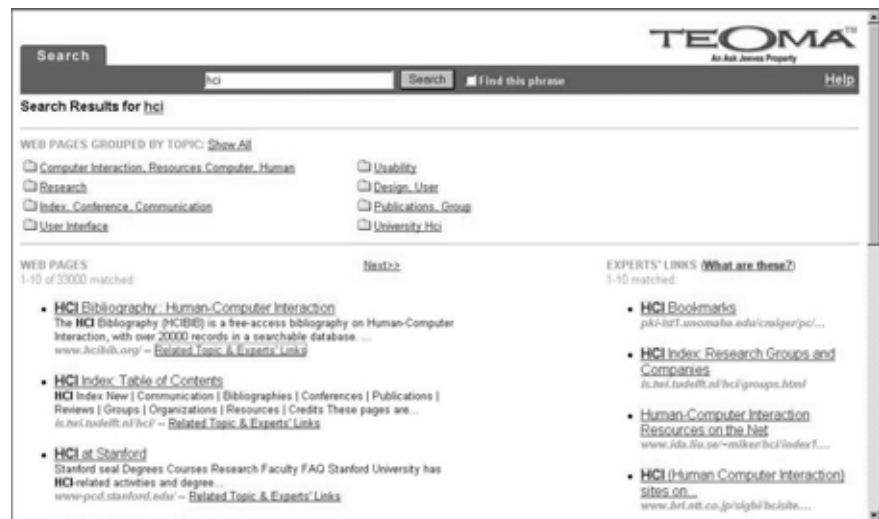
El robot de búsqueda que utiliza, *Zyborg*, rastrea cincuenta millones de páginas al día, recogiendo las URLs y las envía a una base de datos que asigna un número de identificación a cada una. Se comprueba si cada URL recibida está ya recogida en la base de datos y, si no es así, se pide al robot que la indexe y la almacene.

Por último, la sintaxis de consulta permite introducir frases, “+” para inclusión y “-“ para exclusión, pero no operadores booleanos propiamente dichos (and, or, not).

A la hora de ofrecer los resultados proporciona un listado ordenado por relevancia, donde se indican las primeras palabras de las páginas recuperadas en lugar de las frases donde aparecen los términos de búsqueda. Además crea, de forma automática, unas categorías y subcategorías en las que se incluyen los resultados obtenidos (*Wise-Guide*) y que aparecen en la parte superior. Estas ordenaciones se generan basándose en la semántica de los términos incluidos en las páginas recuperadas.

Teoma, “el experto”

Este motor, calificado de nueva generación, hizo su aparición en el mes de junio y fue comprado por *Ask Jeeves Inc.* a primeros de sep-



tiembre. La adquisición no ha supuesto que se cambie su dirección web sino que la intención es mantenerla, además de trabajar para que su tecnología de búsqueda se integre en el sistema *DirectHit* de *AskJeeves* —comprado en enero de 2000—.

<http://www.directhit.com>

<http://www.askjeeves.com>

Este motor nació en 1998 como un proyecto de investigación en *Rutgers*, universidad del estado de New Jersey, y por el momento continúa en su versión beta; de hecho, su base de datos es sustancialmente menor que otras, pues cuenta con alrededor de 100 millones de URLs frente a las más de 1.600 millones de *Google*, si bien tiene previsto aumentar.

El anunciado éxito de *Teoma* se basa en que determina los resultados estableciendo un ranking de los sitios web, para lo que se basa en su “popularidad específica” en cuanto al tema de que trata, es decir, el número de páginas web sobre el concepto a las que se hace referencia en esa página; y en su “popularidad general”, es decir, el número de páginas web que hacen referencia a esa página en concreto. Esto hace a la vez que se identifiquen los sitios que se vinculan entre sí, que a menudo coinciden con los sitios expertos en el tema, lo que aporta más relevancia y au-

toridad a los resultados de búsqueda. Este proceso se ejecuta para cada búsqueda en tiempo real.

Los resultados se presentan en tres formas: las páginas web recuperadas, las categorías temáticas en las que se han agrupado aquellas encontradas (asignadas automáticamente en cada búsqueda basándose en las palabras que aparecen en ellas) y, finalmente, enlaces a sitios denominados “expertos” que recogen a su vez listados de recursos web sobre el tema.

Los resultados que se obtienen en esta versión beta son bastante relevantes, si bien no tanto como los que otros motores ya consagrados como el citado *Google*.

En cuanto a la creación de categorías temáticas, tanto *Teoma* como *Wisenuit* están por el momento centrados en las peculiaridades de la lengua inglesa. Si probamos a efectuar consultas en español nos podemos encontrar categorías denominadas con expresiones del tipo “En el” (ver ejemplo en la pantalla capturada de *Wisenuit*). Este hecho se produce por estar creadas de forma automática basándose en la repetición de las palabras en los documentos web y por no tener un registro completo de las consideradas vacías de nuestro idioma.

Teoma tampoco ofrece una amplia variedad en cuanto a opciones de búsqueda, ni cuenta con una

consulta avanzada. Tan sólo permite indicar que la consulta se trata de una frase uniendo las palabras con un guión o marcando la casilla

que aparece junto a la ventana de búsqueda; y en lugar de operadores booleanos contempla la utilización de los signos “+” y “-“ menos para

incluir y excluir respectivamente los términos seleccionados.

<http://www.wisenut.com>

<http://www.teoma.com>