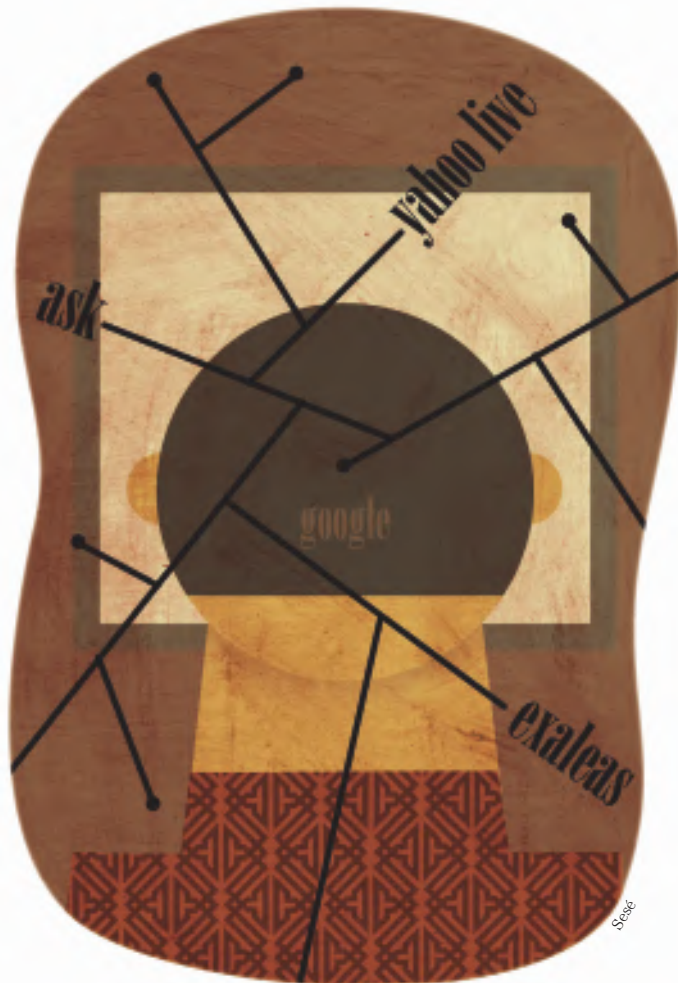


## Búsqueda de información médica en la Web: herramientas avanzadas

Isidro F. Aguillo

Laboratorio de Cibermetría. CSIC. Madrid. España.



En los últimos años, tanto profesionales como público en general han comenzado a reducir significativamente el número de herramientas que utilizan como intermediarias en sus búsquedas de contenidos médicos. Ello significa que, en la práctica, todas las navegaciones en la Web se inicien casi siempre en Google o cualquiera de las implementaciones de la base de datos Medline, normalmente a través de PubMed. Sin menoscabar la calidad y la importancia de ambas bases de datos, y los servicios que ofrecen muchos portales médicos, hay que reconocer que presentan ciertas limitaciones y ciertas lagunas relevantes.

No toda la información médica está en la Web, aunque la mayoría de foros ya tie-

ne su propia sede, y el crecimiento sigue siendo exponencial, ya que es con mucha diferencia la principal fuente de recursos sobre salud. Sin embargo, muchos de los recursos forman parte de la llamada internet invisible, término que identifica a los contenidos que no son indizados por los motores de búsqueda. Hay varias razones para la invisibilidad, pero la principal estriba en la dificultad que tienen los robots de los motores para volcar los contenidos de la base de datos. Sirva como ejemplo que Google, que parece tener una limitación de tiempo, tamaño de ficheros o profundidad de exploración de una sede Web, sólo recoge medio millón de registros de Pubmed (cuya cifra real ya se acerca a los 18 millones a finales de 2007).

### Consejos prácticos

- Google sólo indiza menos de la mitad de los contenidos de la Web. En búsquedas difíciles, es aconsejable utilizar otros motores, como Yahoo, Live, Exalead o Ask, cuyas bases de datos solapan poco con la de Google. Sin embargo, se desaconseja el uso de multibuscadores, dado que sólo recuperan una fracción pequeña de los resultados individuales de cada buscador.
- Los actuales programas de gestión bibliográfica permiten la captura automática de registros desde bases de datos en la Web. Su uso es muy recomendable para construir sistemas personales de referencia que pueden ser muy útiles. Aparte de los ya conocidos comerciales, hay varias alternativas potentes y gratuitas: Zotero, que se instala en el navegador Firefox; Connotea, un servicio de anotación social académica, o Papyrus, que es un programa independiente.

Consejos prácticos

- Algunas herramientas sociales pueden ser muy útiles, aunque no estén dedicadas específicamente a temas médicos o de salud. El servicio Yahoo Answers es el de más éxito en su sector y se pueden formular preguntas en cualquier idioma. La familia de wikis de la Wikipedia no sólo ofrece la ya famosa enciclopedia, sino la Wikiversity, una universidad virtual, o el Wiktionary, un diccionario multilingüe, entre otros.

La existencia de una importante parte de la Web que no se recoge en las bases de datos de los motores de búsqueda pasa inadvertida a la mayoría de los usuarios, e incluso los que conocen este problema encuentran dificultades para solucionarla o paliarla con herramientas adicionales. A continuación, ofrecemos una pequeña serie de recursos que pueden ser puntualmente interesantes y que abren el abanico de la oferta.

Web académica

Además de PubMed, están disponibles otras bases de datos bibliográficas de forma abierta en la Web. La mayoría de ellas presenta una cobertura biomédica limitada y las prestaciones en la recuperación son claramente inferiores. Alguna innovación aparece en Scirus (www.scirus.com), donde se puede buscar simultáneamente en recursos web y bibliográficos, pero no hay una integración verdadera. Desafortunadamente, el buscador académico de Microsoft, Windows Live Academic, tampoco aporta nada nuevo y la base de datos es realmente pequeña.

Google Scholar (Scholar.google.com) es el primer buscador que realmente integra resultados de la Web y bibliográficos, el cual presenta importantes prestaciones. En primer lugar, soluciona, al menos en parte, el problema de la web invisible, al negociar independientemente del comportamiento del robot, la incorporación de registros de otras bases de datos, muchas de ellas propietarias. Esto incrementa sustancialmente la cobertura global, aunque sólo un millón de registros de PubMed aparecen en su base de datos. En segundo lugar, es el primer gran sistema en abierto con enlaces de citas, lo que tiene importantes consecuencias, tanto en la recuperación de información, como en el uso de estos registros para el análisis bibliométrico.

Entre las limitaciones de Google Scholar, todavía en versión beta, se encuentran la heterogeneidad de los objetos incorporados (no sólo artículos, sino también libros, actas, o incluso meras referencias); la falta de un mínimo control documental, lo que dificulta cierto tipo de búsquedas (autor, revista), y los importantes sesgos de cobertura, ya que la incorporación de fuentes se realiza en su mayor parte a petición del interesado en aparecer en este servicio. Se trata de un producto llamado a tener un peso importante en el futuro, pero en la actualidad adolece de notables defectos que deben tenerse en cuenta.

La extracción de registros de la mayoría de estas bases de datos hacia sistemas de gestión bibliográfica, tipo Reference Manager, EndNote o similares, está resuelta respecto a formatos de intercambio, pero no en cuanto a la automatización de descargas de un gran número de ítems. Para Google Scholar hay un programa que, aunque no se ha diseñado originalmente para este fin, puede ayudar en este tipo de penosas tareas. Se trata de Publish or Perish (www.harzing.com/pop.htm), un programa bibliométrico que, a partir de los datos de Google Scholar, construye perfiles de actividad de investigadores. El programa se distribuye gratuitamente y permite la exportación de registros en múltiples formatos, por lo que es una herramienta ideal para compilar pequeñas biblio/webliografías (fig. 1).

En versión beta, también se encuentra HealthVault, el buscador de salud de Microsoft (health.live.com) que se adelanta a un producto similar de Google. Se trata de un mash-up (integrador de servicios) personalizable, que tiene la opción de compartir información como en una red social, previo registro.



Figura 1. Programa bibliométrico Publish or Perish.



Figura 2. Ejemplo de gráfico de BLOGPULSE.

## Web 2.0

Aunque muchas de las herramientas de la llamada Web 2.0 son ya propiedad de las empresas que producen también los motores de búsqueda, no se ha producido una integración plena y, en la mayoría de los casos, los servicios de la web social tienen una existencia propia independiente.

El término Web 2.0 es desafortunado, ya que no es una evolución tecnológica de la Web y se utiliza para designar a un conjunto bastante laxo de servicios con relaciones bastante difusas entre ellos y a veces de difícil delimitación. Básicamente, se trata de sedes que priman la participación de los usuarios en la construcción de éstas, especialmente en lo que se refiere a aportación o edición de contenidos. El propio sistema puede construir, a partir de esa información, redes entre los diferentes usuarios, y utilizar sus perfiles personales, comportamientos o aportaciones, por lo que a menudo se denomina a la Web 2.0 como la web social.

En el mundo médico, con comunidades tanto abiertas como cerradas, el impacto de la Web 2.0 ha sido notable. Puesto que los motores aún no indizan correctamente muchos de sus contenidos, son necesarias herramientas específicas de búsqueda y localización de información.

La blogosfera se caracteriza por una gran heterogeneidad, con bitácoras de gran interés, frecuente actualización y con material original frente a otras muchas apenas mantenidas, y con escaso material propio. Por ello, es necesario contar con un buscador específico que separe el grano de la paja, más aún cuando Google no cubre adecuadamente este sector de la Web (muchos blogs son granjas de enlaces que alteran el correcto funcionamiento del algoritmo PageRank).

Los blog-buscaadores más conocidos, aparte de las bases de datos de blogs de los grandes buscadores comerciales, son Technorati ([www.technorati.com](http://www.technorati.com)) y BlogPulse ([www.blogpulse.com](http://www.blogpulse.com)). Este último es particularmente interesante, puesto que realiza estudios cuantitativos para medir la popularidad de temas y blogs y puede servir como indicador de impacto (fig. 2). Aunque no siempre el impacto va ligado a la calidad de los contenidos, se trata de una herramienta práctica en la selección y la evaluación de mensajes.

Los wikis permiten la edición de sistemas de información gigantes, generados por aportación directa de colaboradores a través de las propias páginas web. Aunque el control de calidad puede resentir-

se por este tipo de prácticas abiertas, una adecuada moderación sirve para solventar la mayoría de los problemas. Para los wikis hay una interesante lista de wikis médicos ([davidrothman.net/list-of-medical-wikis](http://davidrothman.net/list-of-medical-wikis)) que reúne más de 50.

El mundo multimedia ha experimentado una revolución cuantitativa con la Web 2.0. Millones de ficheros de vídeo, audio, fotografías o gráficos aparecen ahora en las web sociales. El podcast (audio o vídeo) es ya una realidad y la Web se convierte en móvil. Un buen ejemplo es la colección de la Clínica Cleveland ([www.clevelandclinic.org/healthedge](http://www.clevelandclinic.org/healthedge)).

Hasta la fecha, ninguno de los buscadores explora adecuadamente nuestro sector, pero eso puede cambiar en un futuro próximo. ]

## Consejos prácticos

- Las herramientas sociales invitan a la participación, de modo que cualquiera puede mantener un blog o colaborar en un wiki. Sin embargo, para recuperar información hay que seguir dependiendo de buscadores convencionales. Una manera fácil de reducir el ruido de los grandes motores es construir un motor de área local, es decir, un buscador que sólo interroge una serie de sedes Web seleccionadas. Esto es lo que permite hacer de forma gratuita el Google Custom Search Engine ([www.google.com/coop/cse](http://www.google.com/coop/cse)) que puede diseñarse para interrogar simultáneamente blogs, wikis y páginas web.

## Para saber más

Aguillo I. A new generation of tools for search, recovery and quality evaluation of World Wide Web medical resources. *J Manag Med.* 2000;14:240-8.

Barsky E. Introducing Web 2.0: weblogs and podcasting for health librarians. *J Can Health Libr Assoc.* 2006;27:33-4.

Boulos MNK, Maramba I, Wheeler S. Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. *BMC Medical Education.* 2006;6:41.

Giustini D. How Web 2.0 is changing medicine. *BMJ.* 2006;333:1283-4.