



Buscadores

Necesidad de los buscadores

Cuando se conoce el URL donde se encuentra la información buscada, o al menos la dirección del ordenador donde se aloja, es muy sencillo usar la Web. El problema aparece cuando sólo se sabe el tema del que se busca información, pero no dónde está. Los buscadores son la respuesta a este problema, aunque no siempre son la solución. Son índices de los contenidos de Internet, pero ni son exhaustivos ni dan siempre con el punto exacto donde está lo que se busca (puede que ni exista).

Internet está cambiando constantemente: en cada momento aparecen nuevas páginas, desaparecen otras y cambian sus contenidos. Es imposible tener una idea, ni siquiera aproximada, de la distribución de contenidos en cada momento. Aun así, existen empresas que intentan tener continuamente actualizada esa información. Utilizan “robots” (programas buscadores, también se llaman “arañas”) que rastrean continuamente la Red y clasifican sus contenidos; las mejores empresas también utilizan clasificación humana, sin programas. Aparecer en los buscadores es importante para ser conocido, de modo que los creadores de páginas suelen enviar sus referencias a ellos.



Existen, en general, dos modos distintos de usar estos buscadores: por **palabras** y por **temas**. Cuando se busca por temas, hay que ir eligiendo entre distintas categorías, cada vez más específicas, hasta que se llega a una página con enlaces a páginas relacionadas con lo que se busca. Cuando se busca por palabras, se introducen una o más palabras y un motor de búsqueda devuelve un listado de enlaces a páginas que contienen esas palabras; normalmente los listados son muy grandes. También hay búsquedas mixtas: se introducen palabras y los enlaces aparecen clasificados por categorías.

Google

Actualmente es el buscador más utilizado. Sus usuarios valoramos de él, además de ser el buscador que más parte de la red tiene indexada, su rapidez de respuesta y de carga de las páginas y que apenas muestra publicidad (la que aparece lo hace en modo texto y resulta relevante para la búsqueda que ha sido consultada).

Google fue fundada en 1998 por **Larry Page** y **Sergey Brin**. Su éxito fue inmediato y se basó en el boca a boca, ya que los internautas que lo usaban obtenían respuestas mucho mejores que con otros buscadores de la época, y se lo comunicaban a otros internautas conocidos suyos.



Google utiliza un sistema llamado **PageRank** para clasificar las páginas web y así poder dar al usuario en los primeros lugares aquellas páginas que le puedan resultar más interesantes. El sistema consiste en que cada página consigue más votos, por decir así, si otras páginas apuntan a ella.

Se accede a Google en la dirección <http://www.google.com>. El interfaz está traducido a muchos idiomas y probablemente detectará el del usuario. Aquí se ve el aspecto de la página en español:



Búsqueda por palabras

Es, con mucho, la búsqueda más usada en Google. Basta escribir en el cuadro de texto ofrecido las palabras que mejor definan lo que se desea buscar. El orden es importante, hay que poner antes aquellas que parezcan más representativas. Luego se pulsa el botón **Búsqueda en Google** para ver la lista de páginas web que contienen todas las palabras escritas. Si se pulsara el botón **Voy a tener suerte**, Google dirige al usuario a la primera página de la lista, pero sin pasar por ver la lista.

Lista de resultados

Valga como ejemplo que se han escrito las palabras “el elefante africano”. La respuesta de Google podría ser como se ve a continuación:



En ella se puede observar:

- ◆ La palabra “el” se ha omitido para realizar la búsqueda. Es tan habitual que no resulta significativa. En general hay que evitar escribir artículos, preposiciones, pronombres, etcétera en las peticiones.
- ◆ Google ha tardado 18 centésimas de segundo en localizar las 3290 páginas que contenían las palabras “elefante” y “africano”. Evidentemente, es muy rápido; raramente alcanza un segundo.
- ◆ Por defecto, se muestran diez resultados por página, de modo que en la primera página de respuestas aparecen las diez páginas más representativas.

En la ilustración sólo se muestra el primer resultado, el que según PageRank es más importante. Permite explicar el significado de cada parte:

- ◆ La primera línea (en azul en la página original) es el título de la página web encontrada. Es un vínculo a la página, así que pulsando sobre él, se accede. Para un trabajo eficiente, se recomienda abrir la página en otra pestaña, para no perder la página de resultados de Google.
- ◆ En las dos siguientes líneas (en negro) se escribe una parte de la página en la que aparecen las palabras pedidas: son el contexto de las palabras; sirven para que el usuario pueda discernir si la página le resulta de utilidad.
- ◆ En la última línea se ven tres informaciones:
 - ◆ La URL de la página, su tamaño y la fecha en que Google la indexó por última vez.
 - ◆ El vínculo **En caché**, que permite acceder a la versión de la página almacenada en los ordenadores de Google. Esto es especialmente útil en aquellos casos en que la página original ya no está disponible (un vínculo muerto), porque aún así se puede ver su contenido en la caché.
 - ◆ El vínculo **Páginas similares**. Google normalmente no muestra en los resultados páginas que están en el mismo servidor que el resultado mostrado y tengan un contenido similar. Con este vínculo, se puede acceder al listado de esas páginas.

Otros buscadores

Son muy conocidos y usados los siguientes:

- ◆ Yahoo <http://www.yahoo.com/>
- ◆ Altavista <http://www.altavista.com/>
- ◆ Alltheweb <http://www.alltheweb.com/>
- ◆ Mamma <http://www.mamma.com/>
- ◆ Excite <http://www.excite.com/>
- ◆ Lycos <http://www.lycos.com/>



Portales

Con la creciente comercialización de Internet surgió el concepto de portal, que no es más que una sede web que pretende ser el compendio de todas las virtudes de Internet ofrecido a usuarios poco avezados. Las personas que sólo utilizan portales pierden la visión de un Internet como punto de encuentro de todas las culturas y formas de pensamiento que siempre la ha caracterizado, y así la convierten en un medio más de ocio dirigido por las grandes empresas comerciales.

Los servicios que se suelen encontrar en los portales son buscador, noticias y correo, normalmente seleccionados de acuerdo a los intereses de la empresa en cuestión.

En este curso se recomienda la utilización libre de Internet, la búsqueda de información independiente y por tanto utilizar lo menos posible los portales mantenidos por empresas que sólo buscan el beneficio propio.

