

Capítulo 2

# Funcionamiento de los motores de búsqueda

# Contenido

1. Introducción
2. Directorios
3. Motores de búsqueda
4. Metabuscadores
5. Resumen

## 1. Introducción

En este capítulo se muestra cómo funcionan los sistemas de búsqueda que permiten navegar por la información que se encuentra en Internet.

Es muy importante que se entienda cómo funcionan estos motores de búsqueda, dado que van a ser aliados en la guerra por vender el producto.

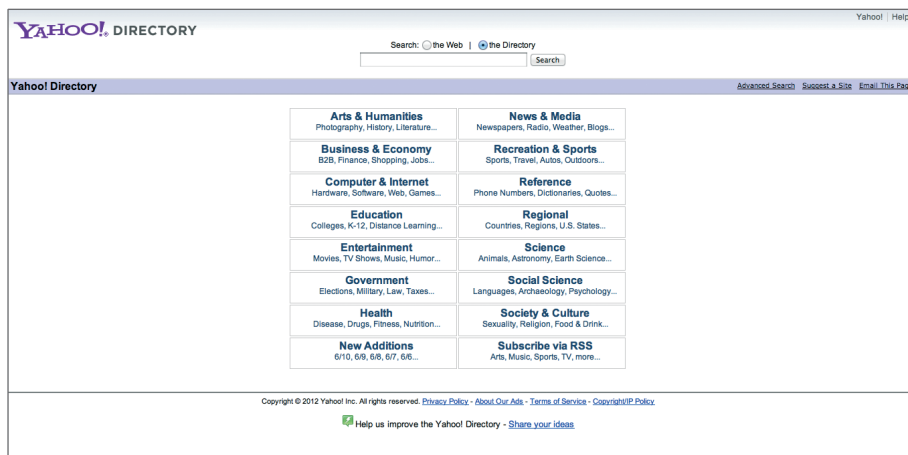
A la hora de poder organizar el contenido en ese amplio sistema de datos, que se llama Internet, se presentan diferentes formas: los directorios, los motores de búsqueda y los metabuscadores.

Inicialmente, los motores de búsqueda fueron gestionados por los propios usuarios, pero conforme el volumen de material fue creciendo las empresas vieron un filón importante en la gestión de este contenido, al cual, en cualquier momento, podrían incluirle algún tipo de publicidad que mantuviera los potentes servidores que necesitan.

Igualmente, los motores de búsqueda evolucionaron a portales de información con noticias y contenidos propios que terminaron por volver a simplificarse, aunque hoy en día ofrecen aplicaciones avanzadas, e incluso sistemas operativos basados en sus bases de datos, como pasa actualmente con *Google*, su *SO Android*, o sus aplicaciones de mapas.

## 2. Directorios

Los directorios, como el de *Yahoo*, son servicios que organizan todas las direcciones web de forma jerarquizada, en categorías y subcategorías, de tal forma que se pueda ir seleccionando lo que más interesa en cada caso. Sería como una gran biblioteca mundial donde se puede acceder a secciones como humanidades, arte, ciencia, noticias, etc.



Directorio de Yahoo.com

Los propios *webmasters* pueden dirigirse a dichos directorios y pedir que sus webs sean insertadas. De tal forma, es la propia persona la que va rellenando los datos e introduciendo su dirección para que puedan ser localizadas por otros internautas.

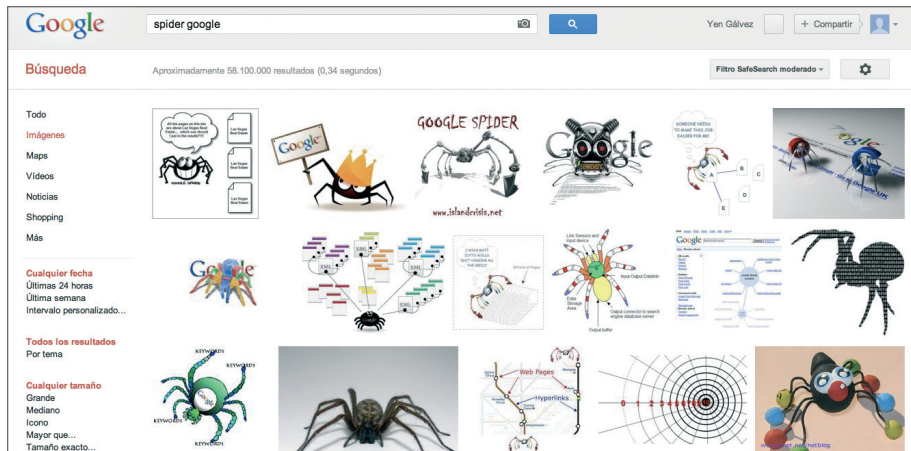


## Actividades

1. Navegar en <http://dir.yahoo.com/> para descubrir distintas webs conforme a los gustos propios.

## 3. Motores de búsqueda

El caso de los motores de búsqueda es bastante diferente. Estos se dedican a almacenar toda la información que pueden de todos los datos que conforman Internet y luego indexarlos mediante un algoritmo propio.



Búsqueda de imágenes sobre Spider

El sistema que se encarga de buscar todos estos datos, para después ser enviados a analizar, se llama *Spider*. Es un robot (código informático) que analiza cada web y los enlaces hasta que consigue convertir toda esa trama de información en algo que pueda ser procesado por el motor de búsqueda.

Una vez el *Spider* ha realizado su trabajo se indexa esa información de tal forma que se pueda realizar una búsqueda para mostrar los datos de forma rápida al usuario.

Sin embargo, el robot no siempre se lleva consigo todos los datos de un sitio web. Determina la importancia de este, de tal forma que si analiza que los datos no son demasiado importantes para la comunidad, simplemente envía una o dos páginas de información para el indexado. Por el contrario, si evalúa y decide que la web es de alto valor en contenido, entonces envían toda la información que puedan encontrar en su directorio. A veces los motores de búsqueda analizan FTP y otros sistemas de archivos.

Este sistema no tiene una actualización inmediata, si no periódica. Esto significa que es muy probable que la información que se encuentre esté desactualizada en ese momento. Sin embargo, el robot se encarga de volver a visitar los sitios para compararlos con los datos almacenados, y en caso de ser necesario,

volver a actualizar todo lo indexado para mantenerse al día. El programa de computación trabajará para crear una base de datos lo más fiable posible.

De cara al usuario el trabajo de indexación es primordial. A la hora de buscar cualquier elemento se hará por palabras claves, estas son las palabras más relevantes de un tema.

Añadir a una búsqueda artículos o determinantes no modifica la búsqueda, dado que lo realmente importante son las palabras clave o *keywords* que se introducen. Esto significa que daría igual que se buscara “Quiero la fecha de nacimiento de Velázquez” o se pusiera directamente “fecha nacimiento Velázquez”. Sin embargo, sí es cierto que a la hora de realizar la búsqueda se pueden usar símbolos que cierren los resultados a recibir. Estos son algunos ejemplos:

- Si la frase que se introduce se encuentra entre “” (comillas) se buscará la totalidad del contenido y no las palabras clave sueltas. Esto significa que a la hora de buscar “En un lugar de la Mancha” mostrará esa frase en los sitios webs que muy probablemente estén referidos al libro de Don Quijote. Sin embargo, si se pone lugar y mancha, como palabras clave, es muy probable que muestre información sobre Castilla la Mancha.
- Se pueden negar palabras, es decir, convertirlas en negativas. De esta forma se consigue que devuelva resultados donde dicha palabra no se encuentre. Por ejemplo, poniendo “-Castilla” en la anterior búsqueda de “lugar” y “mancha”, ya no se recibirían resultados sobre Castilla la Mancha.
- Si a las palabras clave se le añade un + delante devolverá una búsqueda más cerrada, eliminando aquellas otras palabras que puedan ser sinónimos o búsquedas relacionadas.



### Actividades

---

2. Utilizar cualquier buscador y practicar las diferentes opciones de búsqueda. Señalar cuál encuentra más precisa.
-

Los motores de búsqueda poseen un servicio comercial que puede resultar interesante. Es el SEM (*Search Engine Marketing*). El SEM es un sistema por el que cualquier empresa puede forzar a un motor de búsqueda a seleccionar su contenido como privilegiado. Como ya se ha dicho, es el robot quien decide lo que tiene valor en función de si encuentra enlaces con otras webs, si se habla del sitio, etc. En caso de que se haya creado una web nueva es muy posible que se tenga un tráfico muy bajo. Es por ello que realizar una campaña SEM para forzar al buscador puede ser muy provechoso.

Sin embargo, no es la única forma. También se puede utilizar el SEO (*Search Engine Optimization*). Esto es básicamente conseguir que la web sea tenida en cuenta como de alto valor. Para puntuar una web se utilizará una herramienta llamada *PageRank*.

Para conseguir una puntuación elevada es necesario que el tráfico hacia la web aumente, ya sea creando *backlinks* (conectar la web desde un sitio exterior) o enlazándola con redes sociales. Además se debe tener bien organizada para que el robot pueda indexarla cómodamente.

Se debe entender que el mejor método para crear *backlinks* o enlaces externos a la web es el uso de foros.

Incluso llegado el momento se puede decidir hacer parte del trabajo del robot y facilitarle su búsqueda creando una indexación del sitio web.



### Aplicación práctica

---

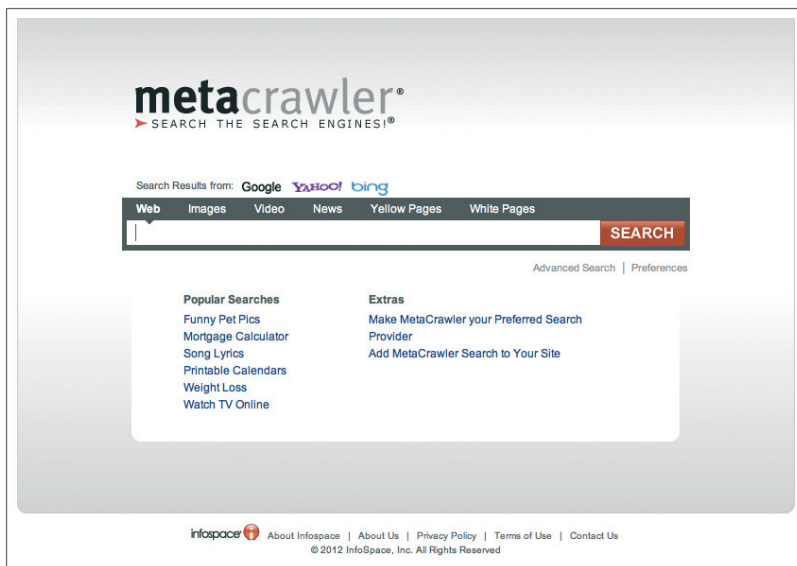
**¿Qué otro método común suele ser usado en Internet para promocionar páginas web?**

#### **SOLUCIÓN**

Los foros suelen ser un método muy útil para compartir con usuarios específicos direcciones de interés, lo cual sirve para subir el PageRank.

## 4. Metabuscadores

Los metabuscadores son sistemas que albergan en una misma herramienta diferentes buscadores. De esta forma se pueden poner varias palabras clave que devolverán los resultados de diferentes motores, como *Yahoo*, *Bing*, *Google*, etc. Son menos usados que los buscadores en sí, pero son interesantes si se pretende encontrar algo difícil de localizar.



Pantalla del buscador Metacrawler



### Actividades

---

3. Comparar los resultados obtenidos entre un buscador sencillo y un metabuscador. Establecer el más exacto.
-

## 5. Resumen

Un motor de búsqueda permite navegar por la información contenida en Internet de una forma más rápida y cómoda.

Los directorios albergan los diferentes sitios web por categorías y los propios *webmasters* pueden dirigirse directamente a dichos directorios para que alberguen su propia web.

Un *Spider* es un *bot* (robot informático) que recopila información para crear la base de datos del motor de búsqueda. Dicha labor permitirá la elaboración de un índice de información que usará el motor para mostrar correctamente los resultados.

Las *keywords* son las palabras que definen cómo el motor de búsqueda va a clasificar el sitio web en Internet. Es importante tenerlas bien definidas, dado que si se está correctamente posicionado el cliente podrá encontrar la web a partir de ellas.

Gracias al SEM (*Search Engine Marketing*) se pueden crear campañas de publicidad en los resultados de búsqueda, posicionando anuncios en los laterales de los resultados o como enlaces patrocinados, y pagando por el *Coste Por Click*.

Mediante un buen SEO (*Search Engine Optimization*) se puede colocar la web en mejores posiciones en el listado de resultados, pero para ello hay que permitir que el *Spider* pueda indexar correctamente la web con palabras clave adecuadas e información importante.

Los metabuscadores permiten realizar búsquedas en varios motores al mismo tiempo, de tal forma que se puede tener más fácil el acceso a la información que se busca.

