



## Tema 4. Internet como fuente para la consulta y documentación

### 4.1 Introducción

Todos hemos perdido muchas horas tratando de buscar un dato, una dirección, un documento en Internet. En este tema, presenta las principales herramientas de búsqueda en el World Wide Web, como buscadores, directorios, agentes inteligentes o metabuscadores.

#### *Preguntas claves*

- ¿Cómo buscar información en Internet de forma eficaz y eficiente? Esta es la pregunta clave que tratamos de resolver en este tema.
- ¿Cuales son los mejores buscadores?
- ¿Cuales son las últimas tendencias en las herramientas de búsqueda?
- ¿Qué es un multibuscador, un agente de búsqueda?
- ¿Qué es la Internet invisible? ¿Cómo buscar en ella?

#### *Objetivos*

- I) Aprender a buscar información en Internet de forma eficaz y eficiente
- II) Utilizar las herramientas de búsqueda como directorios, buscadores, agentes, metabuscadores, buscadores y directorios especializados

#### Algunos consejos generales

#### Planificación

Es fundamental planificar bien la búsqueda. Tratamos de buscar lo más atinadamente posible en el menor tiempo. Pretendemos ahorrar tiempo, el bien más precioso de que disponemos los seres vivos.

#### Objetivos de búsqueda

Cada búsqueda es diferente y exige estrategias distintas. Definiremos bien el objetivo de la búsqueda. No es lo mismo buscar la dirección de la "Asociación Española de Profesores Universitarios de Contabilidad" que "documentos sobre el impacto del comercio electrónico en la Contabilidad".

En el primer caso hay una única respuesta y se trata de encontrarla en el menor tiempo posible.

En el segundo, se trata de encontrar los documentos clave sobre un tema en concreto.

#### Estrategias y tácticas de búsqueda

Una vez definido el objetivo, plantear las estrategias y tácticas posibles. Reflexionar sobre ellas y tratar de ordenarlas según su eficacia y eficiencia prevista. Por ejemplo, plantear una búsqueda accediendo a los grupos de discusión puede ser eficaz pero lento, es decir poco eficiente. Por el contrario, utilizar un buscador puede ser mucho más rápido pero menos eficaz. Hemos de valorar el tiempo que queremos dedicar a la búsqueda.

#### No todo es Internet

Es importante utilizar fuentes Internet y no Internet (colegas, revistas, etc.). Pero incluso en esto puede encontrar ayuda en Internet, ya que algunos buscadores, como Guiame (<http://www.guiame.net>) incorporan enlaces a fuentes de información que no están en Internet.



### En Internet no todo es el World Wide Web

Dentro de las fuentes Internet, no buscar sólo en el World Wide Web. Por ejemplo utilizar los grupos de debate (más conocidos como "las news" o USENET) o listas de distribución que contienen mensajes de particulares. Conviene leer los objetivos de la lista o grupo de discusión para ver si aceptan este tipo de consultas.

### Combinar varios buscadores y estrategias

Utilizar varios buscadores y estrategias de búsqueda (en directorios, agentes inteligentes, multibuscadores, buscadores especializados, etc.), pero con criterios de eficiencia y eficacia.

Como veremos, hay partes del WWW que no están accesibles a los buscadores, el llamado "Web Invisible".

### Leer las páginas de ayuda

Una recomendación imprescindible es leer las páginas de ayuda de los buscadores. Usualmente se trata de unos sencillos consejos o trucos que se aprenden en unos minutos y cuya lectura nos puede ahorrar muchas horas de búsqueda infructuosa. Experimenta y domina al menos uno de ellos.

### Atención a las páginas personales

Revisa las páginas personales en departamentos universitarios, estudiantes, o el público en general. Puede haber alguien con las mismas necesidades que tu que haya anotado los resultados de sus búsquedas. Puede haber muchas sorpresas como colecciones de enlaces interesantes. Aunque conviene velar por la fiabilidad de la información.

### Anotar los resultados de la búsqueda

Es importante anotar los resultados de las búsquedas para no entrar a los sitios dos veces, por ejemplo en un procesador de textos o en los marcadores del programa navegador o bookmarks. Estos bookmarks pueden estar online. Ejemplo:

**Misfavoritos.com** (<http://www.misfavoritos.com>) es un servicio gratuito que permite a cualquier usuario registrado exportar (e importar) los bookmarks de su ordenador a una dirección en Internet a la que, posteriormente, se puede acceder mediante una clave. De esta forma aunque uno acceda a Internet desde diversos ordenadores, siempre tiene disponibles sus bookmarks accediendo a la dirección de Misfavoritos.com.

**BookmarksPlus** (<http://es.bookmarksplus.com>) también ofrece un servicio gratuito en español de favoritos. En inglés existe toda una variedad de sitios que gestionan favoritos, como MyBookmarks (<http://www.mybookmarks.com>) o Favorites Anywhere! (<http://www.favoritesanywhere.com>).

### Seleccione las fuentes

Y, por último, no te creas todo lo que encuentres en Internet. Valora la calidad de las fuentes.



## 4.2 Búsquedas en el World Wide Web

A continuación describimos los principales buscadores y directorios en el World Wide Web. En Internet no hay un índice como en las páginas amarillas de Telefónica. Tampoco podemos acceder directamente a todos los miles de servidores que hay en Internet. En realidad, cuando buscamos en Internet, lo que estamos haciendo es acceder a un servidor o servidores intermedios que contienen un índice parcial de la red. Estos índices son creados por programas de ordenador o por personas.

Nos vamos a centrar en:

- Buscadores puros como Google (<http://www.google.com>)
- Buscadores portal como Altavista (<http://www.altavista.com>)
- Directorios como Yahoo! (<http://www.yahoo.com>)
- Buscadores inteligentes como Askjeeves (<http://www.ask.com>)

Aunque con estos cuatro buscadores puede ser suficiente también veremos buscadores especializados, sitios con solo lo mejor, metabuscadores, espías para estar informado y buscadores de pago.

### ¿Cual es el mejor?

Buena pregunta, pero la respuesta depende de lo que estamos buscando. Una pista es **Searchenginewatch** (<http://www.searchenginewatch.com/awards/>), anteriormente llamado "A Webmaster's Guide To Search Engines" con las principales características y estadísticas sobre los buscadores. Recomendamos la sección "**awards**" (premios) concedidos por votación. En la actualidad, **Google** (<http://www.google.com>) es uno de los mejor valorados.

### ¿Qué busca la gente?

Algunos buscadores como el español **Biwe** (<http://www.biwe.es>) ofrecen la posibilidad de monitorizar en cada momentos las consultas de los usuarios. Acceder a la opción "**El mirón**". (<http://www.biwe.es/miron.html>). [ATENCIÓN: Esta opción pueden mostrar contenidos no adecuados a todas las edades y gustos, al tratarse de una recopilación de las búsquedas de webs y noticias que se están efectuando en Biwe en tiempo real]

En **Hispavista** (<http://buscar.hispavista.com/espia.phtml>).

¿Qué busca la gente en **Google**? En (<http://www.google.com/press/zeitgeist/archive.html>) encontramos un listado de las palabras clave más buscadas cada mes.

Y en **Altavista** (<http://www.altavista.com/sites/about/trends>)

## 4.2.1 Buscadores

### ¿Qué son?

Los buscadores automáticos emplean programas de ordenador llamados robots o arañas, que saltan de una página a otra de la Web recogiendo páginas y almacenando toda la información en una gigantesca base de datos. Esta base de datos contiene, entre otros datos, el título de la página, una descripción, palabras clave y los enlaces.



En algunos casos puede incluir comentarios realizados por los responsables del buscador o una valoración subjetiva -votación- que hagan usuarios que hayan accedido anteriormente al servidor.

El usuario accede a la base de datos rellenando un formulario que ejecuta un mecanismo de búsqueda. Las formas más habituales de hacer esta consulta son con lenguaje natural o mediante combinaciones de palabras clave. Como output recibe un listado con los resultados de la búsqueda.

Una cuestión importante es el orden en que aparecen las consultas, ya que los usuarios solo suelen leer las primeras. Puede haber algoritmos que ubiquen en primer lugar las páginas Web que mejor se adapten a las palabras introducidas, pero en otros casos se cobra por aparecer en primer lugar.

### ¿Quiénes son?

Listado con los buscadores y directorios internacionales más importantes:

Buscadores puros:

- Google (<http://www.google.com>)
- Wisenut (<http://www.wisenut.com>)
- Altheweb (<http://www.alltheweb.com>)
- Raging (<http://ragingsearch.altavista.com>)

Buscadores tipo portal

- AltaVista (<http://www.altavista.com>)
- Excite (<http://www.excite.com>)
- MSN (<http://www.msn.com>)
- Infoseek (<http://www.infoseek.com>)
- Lycos (<http://www.lycos.com>)
- Northernlight (<http://www.northernlight.com>)
- Webcrawler (<http://webcrawler.com>)
- Directhit (<http://www.directhit.com>)
- Hotboot (<http://www.hotbot.com>)
- Netscape (<http://www.netscape.com>)
- Goto (<http://www.goto.com>)

Directorios:

- El directorio de Google (<http://directory.google.com>).
- About (<http://www.about.com>).
- The Open Directory: (<http://dmoz.org>) realizado por miles de voluntarios.
- Looksmart (<http://www.looksmart.com>)



- Webcrawler (<http://webcrawler.com>)
- Links2go presenta los directorios mediante una estructura muy intuitiva (<http://www.links2go.com>).
- Librarians' Index to the Internet (<http://www.lii.org>)
- Terra (<http://buscador.terra.es/>)
- Ozú (<http://categorias.ozu.es>)
- El índice (<http://elindice.com>)
- Yahoo España (<http://www.yahoo.es>).

Multibuscadores:

- Dogpile (<http://www.dogpile.com>)
- Vivísimo (<http://vivisimo.com/>)
- Kartoo (<http://www.kartoo.com>)
- Ixquick (<http://www.ixquick.com>)
- Beaucoup, (<http://www.beaucoup.com>)
- Obsearch (<http://www.qbsearch.com>)
- Metacrawler: (<http://www.metacrawler.com>)

Buscadores más selectivos o inteligentes

- Ask (<http://www.ask.com>)
- Electric Library (<http://www.elibrary.com>)

*¿Cómo funcionan?. Leer la ayuda.*

Cada uno tiene sus propias formas de buscar, así que lo mejor es leer el documento de ayuda correspondiente. Por ejemplo, el robot de Altavista se llama Scooter y el de Yahoo!, Inktomi. Aunque son similares, los diferencian pequeños matices. A continuación vamos a concentrarnos en Altavista y Google.

Para realizar búsquedas eficientes en Altavista podemos consultar:

- la ayuda en español (<http://es.altavista.com/s?spage=help/introduction.htm&=1>)
- Para la ayuda de Google (<http://www.google.com/intl/es/help.html>).

*Buscadores puros*

Los buscadores puros sólo ofrecen búsquedas. Sus sencillas páginas exclusivamente dedicadas a la búsqueda han logrado un creciente núcleo de seguidores y son muy recomendables. Google, Alltheweb de Lycos, Raging del grupo AltaVista, Wisenut son ejemplos de esta categoría.



## Google!

Google! (<http://www.google.com>) incorpora un algoritmo automático que evalúa los sitios encontrados de tal forma que aparecen solo los mas relevantes. Aparecen en negrita los términos o palabras clave solicitadas, de forma que el usuario sabe por qué se ha seleccionado ese recurso.

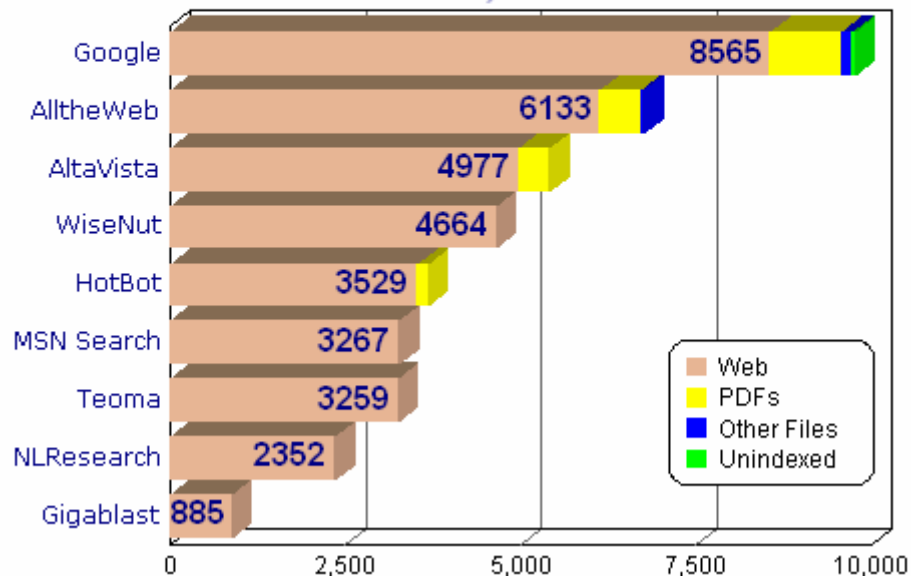
Google utiliza el número de enlaces de una página concreta como medida para evaluar su calidad informativa. De este modo, cada vínculo de una página a otra funciona como un voto a favor de la página receptora. Pero Google no valora todos los votos por igual: valen más aquellos vínculos, o votos, que provengan de páginas que a su vez reciban más enlaces de otras páginas. Mediante este "democrático" sistema el buscador Google ordena la lista de los resultados mediante un escrutinio de votos que sitúa en los primeros lugares de la lista a las webs que más enlaces reciben. Este sistema es el llamado "Page Rank".

- Google tiene una opción, "**I'm feeling lucky**", que lleva directamente al recurso que considera más relevante.
- Tiene también la función "**related**" que permite identificar recursos parecidos al solicitado.
- Google es capaz de buscar documentos **.pdf**.
- De acuerdo con (<http://www.searchengineshowdown.com/stats/size.shtml>) es el más grande en tamaño.

Data from search engine analysis run on Dec. 31, 2002.  
by Greg R. Notess.

## Total Hits from 25 Searches

Dec. 31, 2002 ©2003 G. Notess

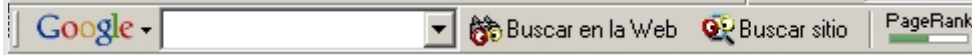


- Y la función "link" muestra todos los vínculos de retorno para una dirección URL determinada.



## La barra de Google

Si dedicamos mucho tiempo a buscar en Internet, una opción muy interesante es añadir al navegador "la barra de Google". (<http://toolbar.google.com>).



Una vez instalada la barra, aparecerá en nuestro navegador. Por cierto recomendamos personalizar las barras de navegación, quitando los elementos que no usemos para no llenar la pantalla del ordenador de barras. Para ello, hacer click con el botón derecho del ratón sobre cualquiera de ellas y ejecutar la opción [personalizar].

Entre las utilidades de la barra de Google!, aparte de que nos ahorramos tener que entrar al buscador, permite buscar sólo en un determinado servidor, obtener su "Page Rank", resaltar las palabras buscadas en cada página, y haciendo click donde pone Google acceso directo a las opciones de búsqueda avanzada.

### Otros buscadores puros

**Alltheweb** (<http://www.alltheweb.com>) de Lycos, también similar a los anteriores en velocidad y sencillez.

**Wisnut** (<http://www.wisnut.com>) es el último en llegar y se habla de él como alternativa a Google!.

**Raging** (<http://www.raging.com>) es del grupo Altavista. Su objetivo es proporcionar al usuario los mejores resultados, con gran rapidez y facilidad de uso. Es interesante usar la opción de personalización (customize) que permite elegir la forma en que se visualizan los resultados.

## 4.2.2 Búsquedas en lenguaje natural

### ¿En qué consiste?

Una forma rápida es simplemente plantear la búsqueda en lenguaje natural. Es decir, si queremos buscar sobre Comercio Electrónico y Contabilidad, en la mayoría de los buscadores podemos simplemente escribir eso:

**comercio electrónico y contabilidad**

Los principales buscadores admiten búsquedas en lenguaje natural. En Altavista (<http://www.altavista.com>) o Google (<http://www.google.com>) veremos que aparecen miles de documentos. Ambos buscadores disponen de algoritmos que tratan de ordenar los documentos según su importancia. Pero es posible mejorar los resultados de esta búsqueda.

Alguno como Ask Jeeves (<http://www.ask.com>) que presentaremos más adelante se ha especializado en búsquedas en lenguaje natural.

Altavista (<http://www.altavista.com>) recomienda añadir más términos a la búsqueda, de forma que se vuelva más específica. Esto se consigue con los símbolos + y -. Por ejemplo, en la búsqueda anterior, nos han aparecido todos los documentos en los que se aparece la palabra comercio, la palabra electrónico o la palabra contabilidad. Por eso aparecían cientos de miles, lo que, en nuestra opinión desaconseja las búsquedas en lenguaje natural. Si probamos con la siguiente búsqueda:

**+comercio +electrónico +contabilidad**



Aparecerá un número más razonable de documentos, aunque todavía muy elevado. Nótese que también la primera palabra (comercio) lleva delante el signo +. [Nota: Aunque en la ayuda de Altavista recomienda esta forma de operar, en la realidad está haciendo como Google, de forma que ya no es necesario poner los +]

Mediante el signo menos, excluimos términos. Por ejemplo, podemos decirle que no queremos que nos muestre documentos en los que aparece la palabra Microsoft.

En este caso escribiríamos:

**+comercio +electrónico +contabilidad -Microsoft**

## Google es diferente

La mayoría de los buscadores (Altavista en la búsqueda simple, Yahoo, Excite, etc.) acepta esta terminología. Una excepción es **Google**, (<http://www.google.com>) que agrega automáticamente "+" entre las palabras, de modo que devuelve las páginas que incluyan **todos** los términos de búsqueda. Google admite el uso de -. **Esta es una de las razones por las que Google ofrece mejores resultados en las primeras posiciones.** El resto de buscadores ya empiezan a utilizar esto.

### Frases literales

Si quiere que una determinada frase forme parte del documento o documentos en que está interesado, tecléela precedida y seguida por el símbolo "dobles comillas". En el ejemplo anterior sería interesante escribir "comercio electrónico".

**+"comercio electrónico" +contabilidad**

El número de documentos que encuentran los buscadores es ya más reducido. El uso de las comillas para expresar frase literal es aceptado por casi todos los buscadores.

## Casi siempre es útil entrecomillar

**En la mayoría de las búsquedas es recomendable utilizar las comillas.** De hecho, aunque no las pongamos los propios buscadores como Altavista "deducen" si se trata de una expresión literal. Por ejemplo, aunque no pongamos entre comillas George Bush, Altavista deduce que te refieres a "George Bush" y devuelve documentos en los que aparece la expresión literal "George Bush".

### Mayúsculas y minúsculas

Algunos buscadores como Google simplemente no distinguen entre mayúsculas o minúsculas y es indiferente ponerlo de una u otra forma.

Pero para otros, como Altavista, las minúsculas es el caso general. Si se pone la palabra en minúsculas Altavista busca todas las palabras, tanto en mayúsculas como en minúsculas. Por el contrario, si se pone la palabra en mayúsculas (o solo la primera letra) solo busca aquellas palabras que se hayan escrito con mayúsculas. Por lo tanto, en general es recomendable en Altavista usar solamente letras minúsculas. Haga la prueba de escribir:

**+"Comercio Electrónico" +Contabilidad**

### Comodines

El uso de comodines (\*) es muy importante y puede usarse en buscadores como Altavista y Yahoo!, pero no en Google. Por ejemplo, para que busque tanto las palabras en singular como plural o raíces de palabras. En nuestra búsqueda anterior, un documento que hablara del impacto del comercio electrónico sobre los contables, pero que no citara la palabra contabilidad,





no sería encontrado. También, en muchos países de Latinoamérica se dice Contaduría y contador. Por lo tanto, pruebe con:

**+"comercio electrónico" +conta\***

### Caracteres especiales y multinacionales

Es mejor consultar la ayuda de cada buscador, pero sabemos, por ejemplo, que en Altavista no es necesario poner acentos o ñes. Incluso es mejor no hacerlo, pues si se emplean acentos u otras marcas diacríticas, AltaVista sólo mostrará los documentos cuya identificación coincida exactamente con los términos empleados en la búsqueda: al poner "electrónico" buscará solo si aparece "electrónico". En cambio "electronico", busca tanto con acento como sin acento.

**+"comercio electronico" +conta\***

### Algunas funciones especiales y trucos

Los buscadores no se limitan a buscar textos, sino que mediante el uso de funciones especiales permite identificar:

- las páginas con la palabra o frase especificada en el nombre de dominio,
- las páginas que residan en un servidor web específico
- las páginas que contienen alguna imagen
- las páginas que contengan algún enlace a un servidor determinado, y muchas más.

## Altavista

Vemos las más importantes en la tabla siguiente, para Altavista. En la ayuda de Altavista hay otras más específicas.

Al solicitar:	Localiza:
<b>+"cuentas anuales"+domain:.es</b>	Documentos con la expresión literal "cuentas anuales", en servidores cuyo dominio sea .es, que corresponde a España
<b>+"cuentas anuales"+host:www.telefonica.es</b>	Documentos con la expresión literal "cuentas anuales", en el servidor www.telefonica.es
<b>image:pacioli</b>	Documentos que contengan un gráfico de nombre pacioli; también en ( <a href="http://images.altavista.com">http://images.altavista.com</a> )!
<b>title:"cuentas anuales"</b>	Para buscar documentos en cuyo título esta la expresión literal "cuentas anuales"
<b>like:www.ugr.es</b>	Busca páginas similares o relacionadas con la dirección URL especificada
<b>link:www.ugr.es</b>	Encuentra las páginas que contengan algún enlace a www.ugr.es
<b>url:ugr</b>	Busca páginas con una expresión o palabra específica en la dirección URL.

## Google

Google, (<http://www.google.com>) incorpora la interesante función site: busca en un servidor concreto (o dominio). Por ejemplo para buscar en ugr.es documentos que tengan la palabra examen:

**examen site: ugr.es**

También incorpora la función **related:** busca páginas semejantes a la localizada.

**related:ugr.es/inicio.html**



También puede usarse la opción "Me siento afortunado", que te lleva directamente al sitio que considera mejor.

## Consejo: usar las opciones avanzadas del buscador

Lo mejor es acudir directamente a las opciones de búsqueda avanzadas:

**Google** ([http://www.google.com/advanced\\_search](http://www.google.com/advanced_search)). Allí podemos decirle que busque sólo documentos pdf, documentos recientes, que busque sólo en el título, etc. **Muy recomendable.**

## Yahoo!

Otros buscadores como Yahoo!, (<http://www.yahoo.com>) o (<http://www.yahoo.es>) también incorporan funciones similares.

## Sintaxis de Yahoo!

Otros buscadores como **Yahoo!**, también incorporan funciones similares, aunque con diferente sintaxis. Por ejemplo, si ponemos u: delante de la palabra buscada, Yahoo! restringe sus búsquedas a las URL del documento.

**Ejercicio:** Prueba como ejemplo, **universidad u:ugr**

## Northernlight

Veamos en Northernlight (<http://www.northernlight.com>) la búsqueda de información de empresas.

## Northernlight, buscar información de empresas

Northernlight tiene una opción especial para buscar información de empresas, añadiendo la expresión COMPANY: o TICKER:.

Por ejemplo, para información de la Telefónica de España, SA.  
COMPANY:telefonica o TICKER:TEF

En Northernlight (<http://www.northernlight.com>) cada sitio aparece una sola vez. Posteriormente se le puede indicar que muestre todas las páginas encontradas de ese servidor. También resalta la palabra buscada, es decir indica por qué ha seleccionado ese servidor. Es muy interesante para evitar confusiones. [Nota: desde 16 enero 2002 muchas búsquedas son de pago]

## Las páginas no existen

Muchas veces nos pasa que una página interesante a priori en el buscador ha cambiado de dirección o no nos deja entrar. Google, (<http://www.google.com>) incorpora la posibilidad de leerla en su caché. La caché es la instantánea de la página que tomó Google cuando entró al servidor.