

INTERNET Y EL "WORLD WIDE WEB": HERRAMIENTAS DE BÚSQUEDA

Prof. Edgar Lopategui Corsino

CONSIDERACIONES GENERALES

Las predicciones de NUA apuntan a un problema de exceso de información y menos efectividad en los mecanismos de búsqueda. Se estima que existen cerca de 320 millones de páginas en el WWW (NEC Research Institute). Al presente, hay alrededor de 1,800 diferentes tipos de mecanismos de búsqueda. Es importante notar que los mecanismos de búsqueda solo abarcan un porcentaje limitado de la totalidad de los datos en el Web.

La internet/WWW no posee toda la información del mundo, de manera que no siempre se encuentra lo que se desea. Además, algunos resultados de búsquedas no son útiles. Más aún, la búsqueda puede ser imprecisa e inconsistente. Con esto en mente, es imprescindible evaluar la calidad de los resultados, cuando se trate de información.

La información en el Web se encuentra organizada en la forma de **base** (o **banco**) **de datos**. Un banco de datos representa una colección de información, la cual se encuentra organizado para aumentar su valor. La guía telefónica es un ejemplo. Un banco de datos en el Web se compone de **registros** (e.g., entradas en una guía telefónica), **campos** (e.g., nombre, dirección, número telefónico) y de una **organización** (alfabético, jerárquico, numérico, entre otros). En el Web, las páginas representan un registro. Los campos pueden ser el título, URL, encabezado, texto y gráficos.

Una búsqueda de información en la internet/Web se puede llevar a cabo de diversas maneras. Por ejemplo, el usuario simplemente puede hojear/examinar (curiosando o "browsing") diferentes páginas que encuentre en el Web. Otra manera de encontrar la información deseada es conociendo de ante mano la dirección o sitio Web (URL). Esta dirección puede ser obtenida a través de otra persona o encontrada en un libro o registro de direcciones de Internet/Web. El último recurso sería utilizar alguna herramienta de búsqueda.

Estrategias de Búsqueda

Para lograr búsquedas efectivas siga las siguientes recomendaciones

- Seleccione dos o tres herramientas de búsqueda y aprenda a usarlas correctamente.
- Identifique sus términos antes de iniciar la búsqueda.
- Considere sinónimos, palabras relacionadas y variaciones tales como plurales.
- Aprenda a utilizar los conectores booleanos.
- Utilice comandos de aproximación y símbolos de truncada cuando estén disponibles.
- No se limite a usar una sola herramienta de búsqueda.

Refinar una Búsqueda

Para efectuar búsquedas que resulten en los pareos más relevantes posible, aplique las siguientes tácticas de búsqueda:

Ser Preciso:

Términos de búsqueda específicos resultarán en pareos reducidos que tienen una mayor posibilidad de que sean relevantes al tópico desado. Antes de entrar una palabra clave, piense **exactamente** qué tipo

de información se desea. Si es posible, determinar la cadena de búsqueda más específica posible ante de realizar la búsqueda, entonces se podrá ahorrar mucho tiempo y frustraciones cuando los resultados exhiben un limitado y muy relevantes sitios Web.

Añadir una Palabra Clave:

Algunas veces una sola palabra clave puede ser muy vaga y como resultado devuelva un gran número de hits, muchos de los cuales pueden no estar relacionados con el tópico de interez. Se puede reducir un tópico de búsqueda al añadir, pues, otras palabras claves en la cadena de búsqueda. No obstante, entrando palabras claves adicionales no siempre arroja resultados precisos. Si entrando palabras claves adicionales reduce el número de pareos, pero los pareos aún cubren muchos tópicos no relacionados, usted deberá de sustituir su cadena de búsqueda con menos palabras claves pero que sean más precisas. En adición, entrando varias palabras claves adicionales puede hacer un tópico de búsqueda tan reducido que usted obtendrá muy pocos o ningún resultado (sitios Web) sobre su búsqueda. Si este es el caso, entonces se recomienda eliminar algunas de tus palabras claves o sustituírlas con otros términos más amplios.

Utilice Operadores Lógicos:

Cuando se emplean múltiples palabras claves, se aumenta la posibilidad de obtener hits relevantes si se especifica que usted solo quiere acceder sitios que contienen **todas** las palabras y no aquellas que contienen **cualquiera** de ellas. Para asegurar que una cadena de palabras claves múltiples resulten en un listado de pareo más corto, usted necesita separar los términos con un operador **and**, el cual le instruye al proveedor de búsqueda de buscar solamente aquellos sitios que contienen **todas** las palabras. El operador comunmente se representa con un signo de más (+) o por la palabra **AND** escrita en mayúscula. Como resultado, el número de documentos enumerados se reducen drásticamente y los documentos exhibidos contienen información más relacionada con el tópico de interez. Usted también puede emplear el operador de **not** para **excluir** palabras de su búsqueda y eliminar documentos no deseados de los resultados. El operador not se representa comunmemte con un signo de menos (-) o la palabra **NOT** escrita en letras mayúsculas. Diferentes sitios Web de búsqueda emplean diferentes palabras o símbolos para representar los operadores, de manera que se debe estar seguro de cotejar la sección de Ayuda ("**Help**") del Buscador para determinar los operadores que se pueden emplear.

Se recomienda que aprendas cómo cada sitio específico trabaja. Cada buscador son levemente diferentes una al otro. No obstante, ellos poseen Preguntas más Frecuentes ("Frequently Asked Questions" [FAQs]) o instrucciones que explican cómo el sitio específico funciona. Tomando algo de tiempo para leer estas FAQs antes de comenzar la búsqueda te ahorrará mucho tiempo.

Si no estas contento con tus resultados, trata de usar otro motor de búsqueda; coteje su gramática; o trate de usar sinónimos o términos relacionados, más amplios o más limitados. Siempre emplee algun tipo de estrategia. La mayoría de los motores de búsqueda permiten al usuario seguir técnicas avanzadas. Por ejemplo, algunos motores de búsqueda te permiten escribir preguntas (búsquedas en lenguaje natural) como si la estuvieras pensando o hablándola (e.g., **¿porqué el cielo es azul?**). Como veremos más adelante, también se pueden emplear frases, i.e., cuando el usuario desea que las palabras aparezcan juntas en un orden específico. Comunmente, los signos de comillas (" ") se emplean para agrupar palabras en frases, de manera que en la búsqueda se le considere como una unidad. Algunos motores de búsqueda emplean paréntesis, comas o guión en vez de las comillas. Por ejemplo, "**vitamina A**"; "**cama y desayuno**", entre otras. Otra estrategia de búsqueda incluye la **búsqueda de campo**. Esta función restringe las búsquedas a ciertas partes de los documentos del Web. Le permite al usuario especificar que las palabras claves de búsqueda aparezcan en el título, URL o en el primer párrafo del documento. Por ejemplo, **title: cáncer**; otra búsqueda de campo puede ser **URL: epa**. La mayoría de los motores de búsqueda no son sensitivos a las mayúsculas. Sin embargo, hay algunos que reconocen las variaciones en las letras mayúsculas y minúsculas. Por ejemplo, **AIDS**, elimina la referencia de ayudantes. Si se escribe **Baker**, la búsqueda busca este nombre y elimina las referencias de bizcochos (cake) y panaderos.

Para acceder un sitio de búsqueda (los buscadores), puedes aplicar un clic sobre el botón de "**Search**" en la barra de herramienta de tu navegador o entrar el **URL** del sitio de búsqueda dentro de la línea de **Dirección** ("**Address**") y luego presione la tecla de [**Enter**]. En la mayoría de los casos usted puede simplemente escribir la porción del nombre del sitio de búsqueda del URL y luego presionar [**Enter**]; tu navegador deberá de proveer las partes de el **http://www** y **.com** de la dirección. Cuando te conectas a un sitio de búsqueda, su página principal exhibirá las alternativas de búsqueda. La mayoría de los sitios de búsqueda contienen una caja de texto dedicada para entrar palabras claves que usted quiere emplear en la búsqueda. Las palabras claves que se escriben para conducir la búsqueda se conocen como **cadena de texto**. La cadena de texto puede ser una sola palabra, frases o puede ser una cadena compleja que incorpore operadores que modifican la búsqueda. Los operadores son palabras (**AND, NOT**) o símbolos (**+, -**) que ayudan a reducir el enfoque de una búsqueda. Una vez se haya entrado la cadena de texto, se inicia la búsqueda al presionar la tecla de [**Enter**] o al aplicar un clic sobre el botón rotulado como "**Search**". El proveedor de búsqueda habra de enviar una lista de sitios Web que contienen la cadena de texto que se entró o una lista de sitios que cubren tópicos relacionados. Para una sesión de búsqueda efectiva, se recomienda aplicar lo siguiente:

- Siembre coteje por palabras mal escritas y errores al escribir el texto.
- Emplee palabras descriptivas y frases.
- Utilice sinónimos comunes y variaciones en las palabras claves.
- Busca y siga las instrucciones que el sitio de búsqueda le sugiere para construir una búsqueda productiva.
- Elimine palabras innecesarias (el, a, una, entre otras) de la cadena de búsqueda.
- Se debe concentrar en palabras claves y frases.
- Pruebe su cadena de búsqueda en diferentes buscadores.
- Explore algunos de los sitios Web que aparecen en el resultado de la búsqueda inicial y localice términos que puedan refinar su cadena de búsqueda.

Busque Ayuda

No existen estándares que gobiernen las opciones y procedimientos de búsqueda, de manera que los propios Buscadores desarrolla los suyos. Usted obtendrá mejores resultados si se coteja la sección de Ayuda del sitio de búsqueda para ver las opciones disponibles antes de realizar la búsqueda. Por ejemplo, algunos Buscadores, tal como AltaVista y Yahoo!, te permiten entrar comillas (" ") alrededor de una cadena de texto para indicar que usted quiere localizar sitios Web que contengan esta frase exacta. Por el otro lado, algunos sitios de búsqueda reconocen el signo de más (+) como un sustituto para el operador de **and**, mientras otros requieren que se entre la palabra **AND** escrita toda en mayúsculas. Otros buscadores pueden no aceptar operadores lógicos. Por lo regular, usted podrá encontrar un enlace a la página de ayuda en un sitio de búsqueda localizado debajo de la caja de texto para la búsqueda de la página principal. Este enlace puede ser identificado como "**Search Tips**", "**Search Options**" o "**Help**". Muchos buscadores también proveen recomendaciones/instrucciones para efectuar búsquedas más complejas, comunmente llamadas como "**Advanced Search**" o "**Power Search**".

Búsquedas empleando un Lenguaje Natural o Común

Algunos motores de búsqueda emplean la tecnología de búsqueda de lenguaje natural. Esto permite al investigador entrar preguntas comunes en su idioma nativo (e.g., Inglés o Español), de manera que se puede reducir aún más los resultados de la búsqueda. Esto elimina el uso de palabras claves y operadores lógicos. Muchos usuarios encuentran esta función de mayor utilidad que entrar las posibles palabras claves, las cuales podrán o no resultar en pareos eficientes y precisos. Por ejemplo, el buscador "Ask Jeeves" permite el usos de lenguaje natural dentro de su cadena de texto de búsqueda.

Sensitividad a las Mayúsculas

La mayoría de los motores de búsqueda no son sensitivos a las mayúsculas. Esto significa que si usted

entra una palabra comenzando con una letra minúscula, tal como **pasteur**, la mayoría de los motores de búsqueda buscará automáticamente por la versión de la palabra en mayúscula (**Pasteur**) en adición de la versión en minúscula de la palabra. Sin embargo, si usted escribe la primera letra de la palabra en mayúscula (**Pasteur**), la mayoría de los motores de búsqueda solamente buscarán por la versión en mayúscula.

Búsqueda con Palabras Claves

Una **búsqueda con palabras claves** utiliza una cadena de texto, comunmente uno o dos términos claves, para buscar un pareo en el catálogo del motor de búsqueda. Cuando se inicia la búsqueda con palabras claves, el buscador investiga su catálogo para la presencia de la cadena de text y luego exhibe los resultados en orden de mayor a menor relevancia. Pueden haber miles y hasta millones de pareos que contienen la cadena de texto. Los pareos exhiben una página a la vez. Se puede ver la próxima página al aplicar un clic al enlace rotulado "**next page**" que provee el buscador al final de cada página de los resultados de búsqueda. La información exhibida en la página de resultados habra de variar, pero la mayoría de los sitios Web de Busqueda muestran lo siguiente:

- Un **título** para la página Web. El título se formatea como un hipervínculo en forma subrayada y texto coloreado, sobre el cual el usuario puede aplicar un clic para acceder el sitio de enlace.
- El **URL** del sitio Web.
- Una **descripción**. Esta puede ser las primeras líneas de la página Web, el título de la página Web, un repaso, resumen o palabras claves.
- El **número total de hits**. Algunos sitios indican el número total de sitios que la búsqueda a producido. Si los resultados son muy amplios (e.g., 6, 134, 200 resultado mostrado en una búsqueda en Google) esto puede ser un indicador que se deberá redicir los términos de búsqueda.

Es posible rastrear los resultados exhibidos para poder ver si un sitio Web contiene la información deseada. Si se observa un listado que se parece a lo que se esta buscando, aplique un clic sobre éste para conectarte a dicho sitio Web. Se puede siempre regresar a sitio de búsqueda al aplicar un clic sobre el botón de "**Back**" localizado en la barra de herramientas de tu navegador; luego seleccione un sitio diferente para visitar o realice otra búsqueda.

Evaluación de la Información Obtenida

Es de vital importancia que los estudiantes o público en general que se encuantra navegando la internet o buscando información posean la capacidad de evaluar efectivamente la información que se accesa. Esto se basa en el hecho de que cualquier persona que posea los recursos neccsarios puede publicar páginas en el Web. A continuación se ofrece una guía simple para evaluar páginas en la Internet/Web:

Tipo del Sitio Web

Determine el tipo de página/sitio que estas evaluando. ¿Es académico el documento, una página personal, un anuncio, de diversión, etc?. ¿Cuales son las metas del autor al crear esta página? La información puede ser influenciada/alterada en cualquier manera.

Calidad

Hágase las siguinetes preguntas:

- ¿El autor de la página/sitio posee credibilidad?
- Qué tipo de evidencia de competencia existe n el área que se discute?
- Trabajan para una organización de alta reputación?
- Cómo comparan los datos y conclusiones presentados aqui con otras fuentes que usted ha consultado?
- Hay disponibles referencias (i.e., citasiones) que apoyen las conclusiones planteadas?

- Este sitio/página del Web ha sido evaluada por individuos reconocidos, de manera que se asegure los estándares de calidad?
- Este sitio Web ha sido evaluado por uno de los servicios evaluativos de la Internet (e.g., Magellan)?

Vigencia

Para poder determinar si las páginas obtenidas están a la par con nuestra realidad presente, trate de contestar las siguientes preguntas:

- ¿Cuándo fue por primera vez que esta página fue creada?
- ¿Cuándo fue la última vez que fue actualizada?
- ¿Se actualiza con regularidad esta página?
- Los datos y conclusiones no están actualizados o son lo suficientemente recientes para ser válidos.
- Los documentos en un área de avance científico rápido pueden caducar su vigencia (estar obsoletos) o simplemente incorrectos dentro de un año de ser creados.
- Coteje que todos los enlaces en la página están aún activos. Se puede sospechar que la página no tiene un buen mantenimiento si un gran número de enlaces están inactivos.

Confiabilidad

Si usted va a ser referencia de este documento (o sitio Web) en una manera prominente, usted podrá estar seguro que se mantendrá disponible por un tiempo razonable? Es más probable que un sitio Web del gobierno se mantenga por varios años en comparación para un sitio de una campaña pequeña. Hay disponibles sitios de espejo? En otras palabras, la página se encuentra duplicada en varios lugares para la facilidad de su acceso?

SITIOS WEB DE BÚSQUEDA (PROVEEDORES DE BÚSQUEDA)

Los sitios de búsqueda registran (en un catálogo) las páginas Web en bases de datos, de manera que el usuario pueda realizar búsquedas en éstos. Estos proveedores de búsqueda se conocen como herramientas de búsqueda o simplemente buscadores. La mayoría de los sitios de búsqueda confeccionan sus catálogos empleando aplicaciones especiales, conocidas como "**crawlers**", "**spiders**", "**robots**" o "**worms**". Estos programas investigan nuevos sitios Web y los organizan en un catálogo, comúnmente al bajar sus páginas principales o almacenando su texto. Los sitios de búsqueda se diferencian en cómo ellos ordenan/clasifican y catalogan los sitios Web y cómo ellos permiten al usuario efectuar búsquedas en sus bases de datos.

HERRAMIENTAS DE BÚSQUEDA (BUSCADORES) EN EL WEB

Las herramientas de búsqueda representan programas que permiten la búsqueda por palabras en diversos servidores WWW, dando como resultado final el lugar donde se encuentra la información. Son páginas que proveen espacios o formas para entrar palabras orientadas a localizar documentos y otras utilidades en las bases de datos que utilizan el formato HTML a través de la Internet. Los buscadores son, pues, programas de computadoras que permiten buscar y coleccionar información de lugares en el "Web". Estos programas especializados proveen un encasillado para escribir términos, llevan a cabo la búsqueda y ofrecen una lista de documentos disponibles. Es importante enfatizar que **ninguna herramienta de búsqueda es totalmente efectiva**. Básicamente, existen tres tipos o categorías de herramientas disponibles para investigar en la Internet/Web, a saber: los **motores de búsqueda** o "**robots**", **megabuscadors (meta-índices)** y los **directorios**. Una descripción de cada uno se ofrece en los siguientes párrafos.

Motores de Búsqueda o "Robots"

Los motores de búsqueda son programas de computadora que atraviesan la base de datos disponibles

(banco de datos) en los documentos del Web en busca de la estructura hipertexto de estas páginas HTML, con el fin de localizar alguna pieza de información, como puede ser un documento, una imagen o un programa de computadora. Estos tipos de buscadores localizan las diferentes Páginas WWW que tienen el término de búsqueda y proveen un índice de documentos. Utilizan múltiples palabras o términos para localizar la información solicitada. Además, aceptan frases, algunos sensitivos a la mayúsculas y emplean campos de acción. Los motores de búsqueda se enriquecen y actualizan automáticamente.

Estas herramientas de investigación son un producto comercial y están diseñadas para rastrear banco de datos en hipertexto localizados en el Web. Un gran número de estos motores de búsqueda son generales y abarcadores (banco de datos grandes). Además, la mayoría requieren el uso de palabras claves.

Los bancos de datos que incluyen los robots puede ser, a saber: 1) páginas del Web (e.g., URLs, encabezados, texto, enlaces), 2) archivos de imágenes y sonido, 3) mensajes de texto completo de grupos de noticias ("newsgroup"), 4) lista de archivos de servidores ftp, entre otros. El lugar donde decide ir el motor de búsqueda depende del "robot" que sea, puesto que todos emplean diversas estrategias. En adición, los "robots" comienzan por buscar en su propia lista de localizadores (URL's), continuando la búsqueda a través de las páginas "Whats New" y los lugares más populares en el Web.

Los métodos de clasificación de los resultados varían entre los diversos "Robots" de búsqueda. Por lo regular, los resultados se ordenan según su relevancia. Puede significar que una palabra clave aparece en la página WWW. Los resultados pueden verse afectados por diversos factores, tales como el tamaño de la base de datos, la frecuencia con que se actualiza, las capacidades de búsqueda disponibles, velocidad de respuesta, el diseño de la interfaz, la forma en que despliega los resultados, cantidad de ayudas que brinde y la composición de los meta-datos.

Cuando se realiza una búsqueda de campo en estos tipos de buscadores, se debe especificar el campo donde el término debe aparecer (e.g., Título: Salud y Medicina, URL: www.saludmed.com). Esto ayuda aumentar la calidad de los resultados. No siempre se requiere entrar palabras claves completas al emplear los "robots". Esto se conoce como **truncación**, es decir, búsquedas realizadas con la raíz de una palabra. Por ejemplo, insertando **comput**, los resultados pueden ser **computadora** o **computación**. Se recomienda evitar los plurales en cada búsqueda.

Operadores Lógicos (Comandos) Booleanos

No todos los tópicos de búsqueda pueden ser descritos por una sola y precisa palabra clave. Con frecuencia es necesario emplear dos o más palabras claves separadas por operadores que le indican al buscador la manera de trabajar con dichas palabras claves. Por consiguiente, al emplear una variedad de operadores y otros comandos avanzados de búsqueda, el usuario podrá llevar a cabo búsquedas más precisas y rápidas. Los **operadores** representan palabras o símbolos que modifican la cadena de búsqueda. Los **comandos booleanos** crean relaciones entre términos/palabras de búsqueda en la base de datos localizadas en el Web. Te permiten buscar el resultado de: la intersección de dos términos/palabras de búsqueda, la combinación de dos palabras o la exclusión de una palabra de la búsqueda. Dependiendo del tipo de Buscador, los comandos booleanos puede ser representados por palabras o símbolos. Los operadores booleados incluyen **conectores booleanos**, **booleanos sustitutos**, **comandos de proximidad**, y **comandos de agrupamiento**.

Conectores Booleanos. Los conectores booleanos son **and**, **or** y **not**. Al emplear el conector **and**, implica que todas las palabras deben de estar englobadas en las páginas Web solicitadas. Especifica que **ambas** palabras en cualquiera de los lados del operador deben estar contenidas en los documentos de la base de datos del Web en la cual se está buscando. Por ejemplo, **Internet and Education** busca únicamente aquellas bases de datos en las cuales se encuentran las palabras Internet y la palabra Educación.

Combinando palabras con el conector **or**, significa que cualquiera de las palabras deben estar incluidas en

los documentos. Especifica que **una o la otra o ambas** de las palabras en cualquiera de los lados del operador deben estar contenidas en los documentos de la base de datos del Web que se esta buscando. Por ejemplo, **Internet or Education** busca únicamente aquellas bases de datos en las cuales se encuentran la palabra Internet o la palabra Educación o ambas.

Por el otro lado, al utilizar **not**, se remueven los documentos que contienen una palabra dada. Este conector especifica que la palabra antes del operador debe estar incluida en los documentos de la base de datos del Web, pero la palabra después del operador **no** debe de estar contenida en los resultados. Por ejemplo, **Internet not Web** busca todas aquellas bases de datos en las cuales se encuentran la palabra Internet con excepción de aquellos registros en la cual aparece la palabra Web.

Boleanos Sustitutos. Son utilizados para sustituir los conectores booleanos **and**, **or** y **not**. Emplean los símbolos:

+ -. Por ejemplo: **+ computadora + disco duro** es equivalente a **computadora and disco duro**.

Comandos de Proximidad (Aproximación). Estos tipos de operadores lógicos son empleados entre dos palabras de búsqueda para indicar que la palabra debe estar contenida en los documentos de la base de datos dentro de una distancia específica de cada una para el pareo en dicho banco de datos. Se caracterizan por seguridad cerca de la encriptación y ayuda con la relevancia. Además, asocia una palabra con la otra. Los comandos de aproximación puede determinar cuan cerca (e.g., distancia entre dos palabras claves), dentro de tres palabras, cuatro palabras, etc.

Los operadores de aproximación se componen de una **palabra o letra** que indica la dirección y un **número** que indica la distancia en palabras. El comando de aproximación **near** se usa para acercar o aproximar dos palabras dadas. Por ejemplo, **university NEAR3 college** parea cualquier banco de datos en la cual la palabra university y college se encuentran contenidas entre tres o menos palabras de cada una en cualquier dirección.

Comandos de Agrupamiento. Se encargan de unir palabras y frases para que sean tratadas como una sola unidad o para determinar el orden en el cual los operadores Booleanos son aplicados. Los más comunes son las **comillas dobles**: “ ” y el **paréntesis**: (). Al emplear las comillas dobles (“ ”), los documentos encontrados por la búsqueda deben de contener las palabras dentro de las comillas dobles exactamente como se entraron. Por ejemplo, **“World Wide Web”** busca aquellas páginas que contienen la frase World Wide Web y no las palabras individuales separadas por otras palabras o las mismas palabras en minúsculas. Por el otro lado, el paréntesis () agrupa palabras y operadores en una cadena de búsqueda para definir el orden en el cual los operadores Booleanos deben aparecer. Su uso es indicado cuando se emplean varios operadores simultáneamente. Por ejemplo, si desea buscar información con referente a un visualizador Web (navegador), la cadena de busqueda que se debe entrar es: **browser and (INTERNET or Web)**; esta cadena encontrará páginas que contienen las palabras Internet y browser o las palabras Web y browser.

Comodines(Wildcards)

Representan caracteres como asteriscos (*) or signos de interrogación (?) que representan a otros caracteres y que aparecen en el mismo sitio, de manera que pueden ser empleado para sustituir cualquier caracter o línea de caracteres. Por ejemplo, **teen*** busca teenage, teenagers o teens. Empleando el signo de pregunta, **Herz?** busca a Herzegovina; **wom?n** hace una búsqueda por woman o women.

Motores de Búsqueda disponibles en el Web

Según la NEC Research Institute (1998), la cobertura de los Robots en el Web indican que **HotBot** posee un 34 por ciento, **Altavista** un 28 por ciento, **Northem Light** con 20 por ciento, **Excite** 14 por ciento, **Infoseek**

10 por ciento y **Lycos** 3 por ciento. Algunos de los Robots de búsqueda que existen en el Web se describen en la Tabla 1.

Tabla 1: Motores de Búsqueda.

Motores de Búsqueda con Interfaz en Inglés				
Los más Comunes y Usados				
MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)	BOOLEANOS EMPLEADOS	SENSITIVIDAD A MAYÚSCULAS	ALCANCE
Google	www.google.com			
Alta Vista	www.altavista.com	AND, OR, NOT Usa booleanos sustitutos	sí	Web, News
HotBot	www.hotbot.com			Web, News
FAST Search	www.alltheweb.com			
Nothern Light	www.northernlight.com			
Excite	www.excite.com	AND, OR, NOT Usa booleanos sustitutos	sí	Web, News
Lycos	www.lycos.com	Usa booleanos sustitutos		Web, gopher, ftp
WebCrawler	www.webcrawler.com	AND, OR, NOT	no	Web
InfoSeek	www.infoseek.com			Web, News
Open Text	www.opentext.com		no	
LookSmart	www.looksmart.com			
GoTo	www.goto.com			
Otros Buscadores Robots				
MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)			
DejaNews	http://www.deja.com/			
InfoHiway	http://www.infohiway.com/			
ThunderStone	http://www.thunderstone.com/			
ScrubTheWeb	http://www.scrubtheweb.com/			
Netscape	http://www.search.netscape.com/			
About.com	http://search.about.com/fullsearch.htm			
PlanetSearch	http://www.planetsearch.com/			
Teoma	http://www.teoma.com/index.asp			
WiseNut	http://www.wisenut.com/			
FindWhat	http://www.findwhat.com/			
iLOR,	http://www.ilor.com/			
Oingo	http://www.oingo.com/			
Search Engines Worldwide	http://www.twics.com/~takakuwa/search/			
Subjex	http://subjex.com/			
NBCi.com	http://nbc.msnbc.com/nbc.asp			
Daypop	http://www.daypop.com/			

Rocket News	http://www.rocketnews.com/
Voila	http://www.world.voila.com/
Findinfo	http://www.findinfo.com/

Motores de Búsqueda con Interfaz en Español

Los más Comunes y Usados

MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)
Google en Español	http://www.google.com/intl/es/
HispaVista	http://www.buscar.hispavista.com/
Biwe	http://www.biwe.cesat.es/
Encuentrelo	http://www.encuentrelo.com/
Telepolis	http://www.telepolis.com/
MexSearch	http://www.yellow.com.mx/
Señas	http://www.combios.net/senyas/
MexWeb	http://www.mty.itesm.mx/Mexweb/
LatinWorld	http://www.latinworld.com/
Mundivia	http://www.mundivia/es/mundi2.html/
Ubbi	http://www.ubbi.com/
Yupi	http://www.yupimsn.com/index.asp

NOTA. Adpatado de: *Características principales de una selección de herramientas de búsqueda del WWW.* Edwin R. Reyes (enero de 1997).

Motores de Búsqueda Especializados en el Web

En el Web existen una variedad de Robots dedicados a buscar información en bases de datos particulares, tales como materias académicas (e.g., educación, ciencias, salud/medicina, entre otros), direcciones de personas, entre otros. Algunos de los Robots de búsqueda especializados más comunes disponibles en el Web se describen en la Tabla 2 y en el listado que sigue:

Tabla 2: Motores de Búsqueda Especializados.

Educación	
General	
MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)
SearchEdu.com	http://www.searchedu.com/
Education World	http://www.education-world.com/
Federal Resources for Educational Excellence	http://www.ed.gov/free/subject.html
Enciclopedias	
MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)
Encarta	http://encarta.msn.com/default.asp
Columbia Encyclopedia, Sixth Edition	http://www.bartleby.com/65/
Investigación General	
MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)
iTools	http://www.itools.com/
Find Articles	http://www.findarticles.com/
Ciencia	

General	
---------	--

MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)
-------------------	---------------------------

Scirus (motor de búsqueda de ciencia): (Información Científica)	http://scirus.com/ http://www.scirus.com/
--	--

Salud y Medicina	
------------------	--

MOTOR DE BÚSQUEDA	DIRECCIÓN EN EL WEB (URL)
-------------------	---------------------------

Medical World Search	http://www.mwsearch.com/
MedHunt:	http://www.hon.ch/cgi-bin/find?1
Medscout:	http://www.medscout.com/
MedBot:	http://www-med.stanford.edu/medworld/medbot/

- AgriSurf! - The Farmers Search Engine: <http://www.agrisurf.com/agrisurfscrips/agrisurf.asp>
- Ask A+ Locator (para expertos): <http://www.vrd.org/locator/>
- Biography Center: <http://www.biography-center.com/>
- Cinemachine: <http://www.cinemachine.com/>
- Deja.com (para grupos de discusión): <http://www.deja.com>
- Daypop (news): <http://www.daypop.com/>
- Direct Search (for databases not covered by the search engines): <http://gwis2.circ.gwu.edu/~gprice/direct.htm>
- FasTV (video searching): <http://www.fastv.com/>
- FindArticles.com (para artículos de revistas a texto completo desde 1998): <http://www.findarticles.com/PI/index.jhtml>
- Forum One (para grupos de discusión): <http://www.forumone.com>
- Internet Movie Database: <http://www.imdb.com>
- Invisible Web (para buscar información escondida en miles de bases de datos): <http://www.invisibleweb.com>
- lexisONE (información legal): http://lexisone.com/html/index_home.html
- Inomics (Economía): <http://www.inomics.com/query/search>
- Liszt (para buscar grupos de discusión por e-mail: <http://www.liszt.com>
- LitEngine.com: <http://www.litengine.com/>
- LLEK Bookmarks: <http://www.scientific-search-engines.com/>
- USgovsearch: <http://usgovsearch.northernlight.com/publibaccess/>
- Pilot-Search: A Literary Search Engine: <http://www.pilot-search.com/>
- Speechbot (para palabras habladas): <http://speechbot.research.compaq.com/>
- SOAR: Searchable Online Archive of Recipes: <http://soar.berkeley.edu/recipes/>
- StartSpot: <http://startspot.com>
- Subject Directory of Search Engines: <http://www.searchiq.com/subjects/>
- Tile.net (listservs): <http://tile.net/lists/>
- Traffick: The Guide to the Portals: <http://www.traffick.com/>
- Ultimate Band List (bands, musicians, lyrics): <http://www.ubl.com>
- WebExhibits: <http://www.webexhibits.com/>
- World Lecture Hall: <http://www.utexas.edu/world/lecture/index.html>
- xrefer (referencia): <http://www.xrefer.com>
- Search Adobe PDF: <http://searchpdf.adobe.com/>
- Search Sonoma: <http://www.searchsonoma.com/>

- **MatWeb:** <http://www.matweb.com/>
- **InfoMine:** <http://infomine.ucr.edu/>

Megabuscadores

Estos tipos de buscadores realizan búsquedas simultáneas en varios "Robots", i.e., representan programas que ejecutan búsquedas simultáneas en múltiples herramientas. Los megabuscadores no poseen una base de datos propia, eliminando los duplicados.

Megabuscadores disponibles en el Web

En la Tabla 3 se describen los principales megabuscadores que cuenta el Web.

Tabla 3: Megabuscadores.

MEGABUSCADOR	DIRECCIÓN EN EL WEB (URL)
MetaCrawler	http://www.metacrawler.com
SavvySearch	http://www.savvysearch.com
Beaucoup	http://www.beaucoup.com
Dogpile	http://www.dogpile.com
ProFusion	http://www.profusion.com
All-in-One Search	http://www.allonesearch.com http://www.unam.mx/allone
StartingPoint	http://www.stpt.com
MetaFind	http://www.metafind.com
Inference Find	http://www.infind.com
Mamma	http://www.mamma.com
Cyber411	http://www.c4.com
AskJeeves	www.ask.com
Pregunta.com	http://www.pregunta.com/
Ixquick	http://ixquick.com/
Vivisimo	http://vivisimo.com/
Turbo10.com	http://turbo10.com/
C4.com,	http://www.cyber411.com/
Infosearch2000	http://www.infosearch2000.com/
Seek123	http://seek123.com/
MetaIQ,	http://www.metaiq.com/index.html
MetaEureka	http://www.metaeureka.com/
MetaSearch	http://www.metasearch.com/
QueryServer	http://queryserver.com/
Metor	http://www.metor.com/
Pandia Search Central	http://www.pandia.com/
Search.com	http://www.search.com/

Direct Search,	http://gwis2.circ.gwu.edu/~gprice/direct.htm
SurfWax	http://www.surfwax.com/
RedeSearch	http://www.redesearch.com/
2TROM Metasearch	http://www.2trhttp://www.2trom.com/om.com/
InfoZoid	http://www.infozoid.com/
Highway61	http://www.highway61.com/
Metagopher	http://metagopher.com/
PowerSearch	http://www.powersearch.com/
SearchTurtle	http://www.searchturtle.com/
WebScout	http://www.webscout.com/
What You Seek	http://www.whatuseek.com/
SherlockHound Metasearch Engine	http://sherlockhound.com/

Directorios (Índices)

Los Directorios son creados manualmente sin la ayuda de ningún "Robot". Por consiguiente, son más precisos que los "robots", pues son personas quienes compilan y mantienen la información. Estos Índices proveen una lista de materiales organizada alfabéticamente: basados en listas o catálogos y organizados jerárquicamente por materias. Los directorios son más pequeños en contenido en comparación con los "Robots"

Los Directorios o Índices representan, entonces, un banco de datos de lugares de internet. Estos lugares son sometidos por sus creadores o evaluadores. Como fue mencionado arriba, los Directorios se organizan en categorías.

Pueden incluir un robot de búsqueda para buscar en su banco de datos. Estos buscadores son similares a catálogos o páginas amarillas. Los sitios en el web cubren el mismo tópico agrupados en uno. Una ventaja de los Índices o directorios es que son buenos para curiosear ("browsing").

Los Directorios o Índices se conoce también como bibliotecas virtuales. Algunas de estas bibliotecas más comunes se describen en la siguiente Tabla:

Tabla 3: Bibliotecas Virtuales.

BIBLIOTECA VIRTUAL	DIRECCIÓN EN EL WEB (URL)
Electric Library	http://ask.elibrary.com/
Internet Public Library	http://www.ipl.org/
The Library of Congress	http://www.loc.gov/
The WWW Virtual Library	http://vlib.org/
Librarians' Index to the Internet	www.lii.org
The Virtual Library of Virginia	http://www.viva.lib.va.us/viva/search/search.html

Directorios (Índices) disponibles en el Web

En la Tabla 3 se describen los principales directorios que cuenta el Web.

Tabla 3: Directorios.

LOS MÁS COMUNES

General

ÍNDICE O DIRECTORIO	DIRECCIÓN EN EL WEB (URL)
Yahoo!	http://www.yahoo.com
Magellan	http://magellan.mckinley.com
Snap!	http://www.snap.com
Web Help	http://www.webhelp.com
Argus Clearinhouse	http://www.clearinghouse.net
World Wide Web Virtual Library	http://www.vlib.org
Britannica.com	http://www.britannica.com
Librarians' Index to the Internet	www.lii.org
DMOZ Open Directory Project	http://www.dmoz.org/
Direct Search	http://gwis2.circ.gwu.edu/~gprice/direct.htm
Invisible Web Directory	http://invisible-web.net/
Search IQ	http://www.zdnet.com/searchiq/subjects/
Google Directory,	http://directory.google.com/
Multnomah Homework Center,	http://www.multcolib.org/homework/
High School Hub	http://highschoolhub.org/hub/hub.cfm
Pinakes	http://www.hw.ac.uk/libWWW/irn/pinakes/pinakes.html
LibrarySpot	http://www.libraryspot.com/
About.com	http://about.com/
IPL Pathfinders	http://www.ipl.org/ref/QUE/PF/
Academic Info	http://www.academicinfo.net/index.html
BUBL LINK	http://bubl.ac.uk/link/
Infomine	http://infomine.ucr.edu/
Suite101.com	http://trackstar.hprtec.org/
DeskRef	http://www.rcls.org/deskref/
Dewey Browse	http://www.deweybrowse.org/
Ready Reference Using the Internet	http://www.winsor.edu/library/rref.htm
CyberDewey	http://www.anthus.com/CyberDewey/CyberDewey.html
Digital Librarian	http://www.digital-librarian.com/
Awesome Library	http://www.awesomelibrary.org/
LibraryHQ.com	http://www.libraryhq.com/

Best Information on the Internet	http://library.sau.edu/bestinfo/Default.htm
Galaxy	http://www.galaxy.com/
HotSheet	http://www.hotsheet.com/
Invisible Web,	http://www.invisibleweb.com/
Kathy Schrock's Guide for Educators	http://school.discovery.com/schrockguide/
KidsClick!	http://sunsite.berkeley.edu/KidsClick!/
Kid Info	http://www.kidinfo.com/SchoolSubjects.html
LookSmart,	http://www.looksmart.com/
Open Directory Project	http://dmoz.org/
Yahooligans	http://www.yahooligans.com/
StudyWeb	http://info.studyweb.com/
Ringworld: The Webring Directory	http://www.webring.org/rw
Internet Scout	http://scout.cs.wisc.edu/index.html
Cybrary for Middle School and Beyond	http://www.geocities.com/Athens/Academy/6617/
100Hot.com	http://www.100hot.com/index.html
TrackStar	http://trackstar.hprtec.org/
Where to Do Research	http://www.wheretodoresearch.com/
ingenta	http://www.ingenta.com/
Terra	http://www.terra.es/

De Universidades

General

ÍNDICE O DIRECTORIO

DIRECCIÓN EN EL WEB (URL)

Indiana University

<http://www.indiana.edu/~librcsd/internet/>

El Web También cuenta con unos Índices o Catálogos especializados, los cuales se describen la ta Tabla 5:

Tabla 5: Índices o Catálogos Especializados.

Ciencia	
General	
ÍNDICE O CATÁLOGO	DIRECCIÓN EN EL WEB (URL)
Biología	
ÍNDICE O CATÁLOGO	DIRECCIÓN EN EL WEB (URL)
Biology Online	http://www.bioscience.org/urllists/biology.htm
Salud/Medicina	
ÍNDICE O CATÁLOGO	DIRECCIÓN EN EL WEB (URL)
Best Medical and Health Resources	http://www.infopeople.org/Wksp/Past/1999/Medical/medbkmrks.htm
Hardin Medical Directory:	http://www.lib.uiowa.edu/hardin/md/
Indexes to Resources in Medicine:	http://www.smh.toronto.on.ca/hslibrary/medindex.htm
Achoo	http://www.achoo.com/main.asp

NOAH	http://www.noah-health.org/
WWW Medical Indexes	http://www.intmed.mcw.edu/MedIndex.html

Programas/Utilidades de Búsqueda

Representan aplicaciones de utilidad empleadas para hacer megabúsquedas en la Internet. Estos programas se puedan bajar del Web. Los más comunes se describen en la siguiente Tabla:

Tabla 6: Utilidades de Búsqueda

General	
PROGRAMA	DIRECCIÓN EN EL WEB (URL)
FerretSoft	http://www.zdnet.com/ferret/
Copernic	http://www.copernic.com/
Infoseek Express	http://express.infoseek.com/
BullsEye	http://www.intelliseek.com/
SearchWolf	http://www.trellian.net/search/
SearchPad	http://www.searchpad.com/
Inforian Quest	http://www.inforian.com/

ENLACES

Tipos de Motores de Búsqueda

- [Los Principales Motores de Búsqueda](#)
- [Motores de Búsqueda Globales](#)
- [Nuevos Motores de Búsqueda](#)
- [Motores de Búsqueda Regionales: America Latina](#)
- [Metabuscadores: Principales](#)
- [Metabuscadores: Otros](#)
- [Motores de Búsqueda Pagados](#)
- [Motores de Búsqueda Especializados: Salud/Medicina](#)
- [Guía para Motores de Búsqueda](#)
- [Utilidades de Búsqueda: Búsquedas Meta](#)
- [Utilidades de Búsqueda: Otros](#)

Operadores/Comandos de Búsqueda

- [Características de los Motores de Búsqueda](#)
- [Matemática de los Motores de Búsqueda](#)

Artículos

- [Metabuscadores](#)
- [Guías](#)

Operadores/Comandos de Búsqueda

- [Caraterísticas de los Motores de Búsqueda](#)
- [Matemática de los Motores de Búsqueda](#)

REFERENCIAS

Libros

Katsaropoulos, C., & Stevenson, N. (2001). *Learning the Internet for Business* (2da. ed., pp. 84-110). New York: DDC Publishing, Inc.

Web

Tutoriales

Ross Tyner, R., & Slany, W. (2002). *Sink or Swim: Internet Search Tools & Techniques* (Version 5.0). (<http://www.ouc.bc.ca/libr/connect96/search.htm>).

Gray, T. A. *How to Search the Web: A Guide To Search Tools* (<http://daphne.palomar.edu/TGSEARCH/>).

OWL at Purdue University and Purdue University (2002). *General Tips for Searching the Web* (<http://owl.english.purdue.edu/internet/search/general.htm>).

TUG Libraries (2000). *Internet Search Tools* (<http://www.tug-libraries.on.ca/internetsearch/>).

Search EngineWatch (<http://www.searchenginewatch.com/>).

University of Sydney Library (2002). *Searching the Internet* (<http://www.library.usyd.edu.au/skills/isearch.html>).

Listado de Buscadores y otras Herramientas de Investigación

Los Angeles County. Office of Education . *Internet Search Tools* (<http://teams.lacoe.edu/documentation/internet/search.html>) (NOTA: Incluye artículos y cómo evaluar la información)

Rice University (2001). *Internet Navigating Tools* (<http://www.rice.edu/Internet/>).

Library of Congress (2001). *Internet Search Tools: Explore the Internet* (<http://www.loc.gov/global/search.html>)

Berkely Digital Library SunSITE (2000). *Internet Search Tool Details* (<http://sunsite.berkeley.edu/Help/searchdetails.html>)

Southern Oregon University (2002). *Internet Searching Tools* (<http://www.sou.edu/library/searchtools/>).

The University of Queensland, Brisbane Australia (2002). *Cybrary Internet Searching* (<http://www.library.uq.edu.au/internet/schhints.html>).

The Snark Hunter's Page (2002). *Internet Search: Tools for Finders and for Wanderers* (<http://www.islandnet.com/~pjhughes/search.htm>)

UBC Library (2002). *Internet Search Tools* (<http://www.library.ubc.ca/home/isearch.html>).

Ramapo Catskill Library System (2002). *Search Tools* (<http://www.rcls.org/search.htm>).

NISS (2202). *WWW Search Engines* (<http://www.niss.ac.uk/lis/search-engines.html>).

Ensor, P. (2002). *Tool Kit for the Expert Web Searcher* (<http://www.lita.org/committe/toptech/toolkit.htm>).

Department of History, Tennessee Technological University (2000). *Internet Search Tools* (<http://www.tntech.edu/www/acad/hist/search.html>).

Jean & Charles Schulz Information Center (2002). *Internet Search Tools* (<http://libweb.sonoma.edu/web/searchtools.html>)

Comparación de las Capacidades de Investigación entre los Diferentes Buscadores

University of Pennsylvania Library (2002). *Search Tools* (<http://www.library.upenn.edu/resources/search/search.html>).

[▲ Regresar Arriba](#)

[▲ Regresar a: Navegación y Búsqueda](#)

[▲ Regresar a: Informática y Telecomunicaciones](#)

[▲ Regresar a la: Página Principal](#)

Copyright © 2002 Edgar Lopategui Corsino