

Evaluación de los principales "buscadores" desde un punto de vista documental: recogida, análisis y recuperación de recursos de información

Angeles Maldonado Martínez

CINDOC-CSIC

Elena Fernández Sánchez

CINDOC-CSIC

Resumen

El presente trabajo describe y analiza sistemáticamente las características documentales de los veinte buscadores, índices y motores de búsqueda, internacionales y en español de mayor relevancia.. En primer lugar se explica el criterio aplicado para seleccionar los buscadores. A continuación se definen las funciones básicas que debe de realizar y las características que debe reunir un buscador como bases de datos documental que es desde el punto de vista de la recogida, análisis, búsqueda y recuperación de recursos. Seguidamente se describen pormenorizadamente cada uno de los buscadores seleccionados. En último lugar se analizan los resultados obtenidos.

Palabras-clave

Internet, buscadores, motores de búsqueda, índices, características documentales, recogida de información, análisis documental, recuperación de información

1. Objetivo

El objetivo del presente estudio es describir y analizar sistemáticamente las características documentales (recogida y análisis de información, búsqueda y resultados) de los veinte buscadores, índices y motores, internacionales y en español de mayor relevancia.

2. Fases del estudio

Los pasos seguidos para la realización del presente trabajo han sido los siguientes:

- a) Definición de criterios para seleccionar los buscadores y selección de los mismos.
- b) Definición de las características documentales a cumplir por los buscadores seleccionados.
- c) Descripción de las características documentales, definidas con anterioridad, de cada uno de los buscadores seleccionados.
- d) Análisis de resultados a partir de los datos obtenidos.

3. Selección de buscadores

En la actualidad existen un gran número de buscadores, tanto índices temáticos como motores de búsqueda. Para realizar el presente estudio se han seleccionado veinte, diez de ámbito internacional y otros diez dirigidos al mundo hispano. Tanto en un caso como en otro los elegidos son aquellos que tienen un mayor número de URLs enlazadas. En un primer momento fueron preseleccionados aquellos buscadores más citados en diversos compiladores. Para realizar la selección definitiva se lanzó una búsqueda por "link" (localización de páginas enlazadas a una determinada dirección) en tres buscadores que permiten interrogar por este campo: Altavista.digital.com, Altavista.magallanes.net e Infoseek.

En el caso de los buscadores internacionales (ver gráfico 1) el ranking se estableció en función de los resultados obtenidos en Altavista.digital.com. Los buscadores seleccionados y el número de páginas enlazada a cada uno de ellos fueron los siguientes: Yahoo! (754.739), Altavista (272.761), Excite (206.979), Lycos (180.128), WebCrawler (93.904), Hotbot (93488), Magellan (60.994), Infoseek (33.214), Galaxy (26.528) y Looksmart (10.433).

Con respecto a los buscadores en español (ver gráfico 2) se primaron los datos obtenidos en Altavista.magallanes.net y los resultados fueron los siguientes: Ole (5.770), Donde? (3.673), Ozu.com (2.072), Eleano (1742), Biwe (1.415), Ozu.es (1.229), Trovator (1.155), Tarántula (1.056), Hispavista (869) y Sol (706).

4. Definición de características documentales

Los buscadores son las herramientas más importantes para localizar información en Internet. Tanto los índices temáticos como los motores de búsqueda son bases de datos de URLs. En el caso de este trabajo se han establecido dos tipos de URLs: las que representan páginas individualizadas y las que representan recursos de información completos compuestos por una o más páginas. Por norma general los índices gestionan recursos de información y los motores páginas.

Las bases de datos documentales deben realizar unas funciones básicas y cumplir una serie de características, que a los buscadores también se les debe exigir. A continuación se describen dichas funciones y características agrupadas en tres apartados: recogida y análisis de información, búsqueda y resultados.

a) Recogida y análisis

En una cualquier base de datos la captación de datos se puede hacer de forma manual o de forma automática. Lo mismo ocurre en el caso de los buscadores. El procedimiento manual, utilizado por los índices, consiste en rellenar un formulario de alta que los propios buscadores proporcionan. Los datos solicitados en estos cuestionarios varían de unos a otros. Cuanto más completa y pormenorizada sea la toma de datos, las posibilidades de búsqueda podrán ser también más completas. En este trabajo se han considerado dos tipos de datos a consignar, unos que se les ha otorgado la categoría de imprescindibles: título, URL, clasificación y descripción; y otros considerados recomendables: palabras-clave, localización, idioma, tipo de información (información académica, información comercial, directorio, publicación electrónica, etc) y tipo de propietario (organismos públicos, empresas, páginas personales, etc). En el caso de los recursos de información existentes en Internet parece excesivo utilizar las cuatro modalidades de representación de contenido: título, clasificación, palabras-clave y descripción (resumen). La clasificación viene avalada por la utilidad de las búsquedas a través del índice temático. Entre el título, la descripción y las palabras clave, se han considerado como imprescindibles las dos primeras: el título por ser la modalidad más concisa y por lo tanto la que introduce un grado de error menor en la recuperación y la descripción, porque en la totalidad de los casos su longitud está muy limitada, pudiéndose ser considerada como un conjunto de palabras clave unidas sintácticamente. Además el contenido de la descripción se puede utilizar posteriormente en la presentación de resultados. Sin embargo también se ha considerado recomendable la existencia de palabras clave, por aumentar la pertinencia en la recuperación.

El procedimiento de toma de datos automático es el que normalmente utilizan los motores de búsqueda, quienes únicamente solicitan la URL de la página a dar de alta. A continuación su robot es capaz de visitar e indizar la página dada de alta y todas aquellas otras que están enlazadas. El lenguaje HTML permite utili-

zar una serie de etiquetas, denominadas Meta, a través de las cuales se puede añadir cualquier tipo de información sobre una página, en el momento de su confección. Con ayuda de estas etiquetas se puede realizar un análisis completo de una URL desde el punto de vista documental. Principalmente se suelen utilizar para describir el contenido a través de pequeños resúmenes y palabras-clave. Hay robots que son capaces de identificar las etiquetas META y extraer la información de las mismas para ser usada en la búsqueda o en la presentación de resultados. Sin embargo, existen otros que no la detectan o no utilizan su contenido. Los datos imprescindibles y recomendables a consignar en las etiquetas Meta serían los mismos que en el caso de los formularios de alta de los índices: título, URL, clasificación, descripción, palabras-clave, localización, idioma, tipo de recurso y propietario de las páginas. Hay que resaltar que las etiquetas Meta no son obligatorias a la hora de diseñar una página, y en la actualidad hay muchas URLs que carecen de ellas.

En documentación, en términos generales, son más aconsejables los sistemas que son capaces de analizar más pormenorizadamente las características formales y de contenido de cada uno de los registros que gestionan. En el caso de los buscadores, no se deben contraponer de forma general índices y motores de búsqueda, sino más bien hay que indagar las características particulares tanto de unos como de otros.

Consecuentemente con lo anterior, si hubiera que establecer un ranking de categorías de buscadores desde el punto de vista de sistema de recogida y análisis de direcciones el orden sería el siguiente: 1) Índice temático con al menos título, URL, clasificación, descripción de cada una de las páginas. 2) Buscador con reconocimiento y análisis de etiquetas Meta. 3) Buscador sin reconocimiento y uso de las etiquetas Meta.

b) Búsqueda

Un sistema de búsqueda es mejor cuanto más flexible es y cuantas más posibilidades de recuperación ofrece. Las características que se han considerado básicas desde el punto de vista de la recuperación de la información que debe cumplir un buscador, ya sea un índice o un motor de búsqueda, son las que a continuación se citan:

- Formularios de búsqueda.

Posibilidad de elegir entre un formulario simple y otro más detallado. El simple da la posibilidad de acercarse a los buscadores a aquellas personas que no están muy familiarizadas con el uso de las bases de datos. Los formularios más completos dan la oportunidad de realizar búsquedas más complejas a los usuarios expertos.

- Herramientas de búsqueda.

Posibilidad de uso del truncado de palabras, de los operadores booleanos: AND, OR y el NOT y del paréntesis. También es recomendable la localización de términos compuestos (utilizando comillas o buscando frases). El contar con todas estas herramientas permite realizar búsquedas que definen muy ajustadamente la información que se requiere.

- Clasificación temática.

Existencia de un índice general, pues facilita mucho la localización de información a aquellos que no saben concretar su tema de búsqueda y prefieren explorar los apartados temáticos propuestos por el sistema.

- Campos de búsqueda.

La posibilidad de dirigir las búsquedas a campos determinados aumenta notablemente la pertinencia en la recuperación de información. Las búsquedas en texto libre son más exhaustivas pero menos pertinentes. La búsqueda en campos determinados está muy relacionada con la segmentación de datos en la recogida de datos. Cuanto más sectorializada esté la información extraída de un recurso de información y más se pueda concentrar la búsqueda a un tipo de dato concreto más pertinente será la recuperación. Como en el caso de la toma de datos se consideran campos imprescindibles: título, URL, descripción; y campos recomendables: palabras-clave, localización, idioma, tipo de información y tipo de propietario.

- Control del vocabulario.

Uno de los problemas fundamentales en la recuperación de información es el control de vocabulario. Eliminar las sinonimias y las polisemias es una labor a realizar en todo sistema documental que se precie, pues

disminuye considerablemente el ruido y el silencio.

- **Detección de novedades.**

En todo sistema documental es imprescindible poder detectar los nuevos registros incorporados. Existen varios sistemas a través de los que detectar las novedades: creación de ficheros aparte, uso de etiquetas especiales, acotar por fecha de alta y ordenar los recursos por fecha de alta.

c) Resultados

Es aconsejable que los usuarios puedan elegir entre diferentes formatos de presentación de resultados o incluso diseñarlos a la medida. También es recomendable que todos los documentos que componen un sistema documental vayan acompañados de una descripción de su contenido. En cuanto al orden de presentación de los resultados, es importante que se pueda seleccionar entre varios criterios de ordenación. En el caso de los motores de búsqueda, ya que gestionan URLs, es recomendable que puedan presentar las páginas recuperadas agrupadas por servidores.

5. Descripción y análisis de los buscadores seleccionados

A continuación se describen las funciones y las características documentales de cada uno de los buscadores seleccionados. También se resalta lo positivo y lo negativo de cada uno de ellos.

YAHOO! (Yahoo! Inc.)

<http://www.yahoo.com/>

Contenido: cobertura internacional con nodos locales que sirven a sus respectivos ámbitos geográficos en: Alemania, Asia, Australia, Canadá, Corea, China, Dinamarca, España, Estados Unidos, Francia, Italia, Japón, Noruega, Reino Unido y Suecia.

Recogida de Páginas: la mayoría de los 'sites' son sugeridos por los propios usuarios, y más tarde son visitados y evaluados por un equipo de Yahoo!. Existe un formulario de sugerencias donde se solicita la URL, el título y la descripción del recurso a incluir. Yahoo! está organizado en categorías temáticas, por lo que el solicitante además debe indicar la(s) categoría(s) en las que piensa debe enmarcarse el recurso sugerido.

Búsqueda: se puede buscar siguiendo dos procedimientos: o bien a través de la clasificación temática o bien utilizando los formularios de búsqueda. En el primer caso se puede descender en la jerarquía hasta donde se considere oportuno y visualizar los recursos de interés. En el segundo, se teclean los términos en una ventana debiéndose elegir entre visualizar los epígrafes clasificatorios o los recursos que contengan dichos términos. Si el resultado no es satisfactorio, se tiene la posibilidad de lanzar la pregunta en las páginas indizadas por el motor de búsqueda Inktomi o conectar con otro buscador (Altavista, Hotbot, Infoseek, etc ...). El sistema admite trincar, buscar por términos compuestos y el uso de los operadores booleanos AND (+), NOT (-) y OR. La opción 'Search Options' permite hacer una acotación en función del tiempo transcurrido desde el alta del recurso, preguntar por una URL determinada o buscar los términos en el título del recurso.

Resultados: no se puede formatear el resultado de una búsqueda. De cada 'site' se visualiza el título con el enlace, una pequeña descripción y el epígrafe clasificatorio al que pertenece, en el caso de que la búsqueda se haya realizado en la base de Yahoo!. Si la búsqueda se ha lanzado en Inktomi la descripción se corresponde con las primeras líneas de las páginas. Alcanzarán un mayor rango en la presentación aquellos documentos que contengan más palabras clave, que estas estén en determinadas partes del documento (se prima el título) o que el recurso pertenezca a una categoría que esté situada en una posición elevada, es decir más general.

Ayudas: las ayudas detallan la forma de realizar las búsquedas con ejemplos, los operadores a utilizar y los criterios de relevancia que se aplican en la presentación de resultados.

Otros: a las nuevas incorporaciones de les añade la etiqueta de 'nuevo' y se las posiciona al principio de cada epígrafe durante la primera semana.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicita la URL, el título y

la descripción del mismo. 2) Existe una clasificación temática. 3) Posibilidades de búsqueda satisfactorias ya que se pueden utilizar los operadores AND, OR y NOT, se puede truncar y localizar términos compuestos. 4) Se pueden acotar las búsquedas a la URL y al título. Las búsquedas en texto libre se realizan sobre la descripción del recurso. 5) Se recuperan unidades de información y no páginas. 6) las ayudas son satisfactorias. 7) Se pueden detectar las nuevas incorporaciones al acotar por fecha de alta del recurso. 8) Tiene versión en español.

Negativo: 1) Posibilidades de búsqueda medianamente satisfactorias, ya que aunque el sistema permite truncar, localizar términos compuestos, y utilizar los operadores AND, OR y NOT, no es posible utilizar paréntesis. 2) No se controla el vocabulario. 3) No se pueden diseñar los formatos de presentación. 4) No se pueden mantener abiertas paralelamente la lista de resultados de la búsqueda y una de las direcciones recuperadas.

ALTAVISTA (Digital Equipment Corporation)

<http://altavista.digital.com/>

Contenido: cobertura internacional con nodos locales que sirven a sus respectivos ámbitos geográficos y lingüísticos en los siguientes países: Australia, Canadá, España, Estados Unidos, Japón y Suecia.

Recogida de Páginas: las páginas son recogidas por un robot que es capaz de indizarlas y de saltar a aquellas otras enlazadas con ellas. Las altas se realizan a partir de un formulario donde únicamente se solicita la dirección de la página a dar de alta. El sistema es capaz de reconocer las etiquetas META de título, descripción y palabras-clave y extraer la información que contienen.

Búsqueda: presenta dos formularios de búsqueda con casi las mismas prestaciones, ya que se puede acotar por idioma, buscar en campos concretos (dominios, servidores, URLs, links, imágenes, título e idioma), utilizar operadores booleanos (AND, OR, NOT), de proximidad y paréntesis, localizar términos compuestos y truncar. Los dos también contemplan la posibilidad de 'afinar' la búsqueda, es decir hacer una segunda selección entre las páginas recuperadas originalmente en función de la elección de nuevos términos propuestos por el sistema por el grado de coincidencia con los demandados en un principio. La diferencia entre los dos formularios reside en que en el simple las páginas seleccionadas son presentadas por un orden de relevancia establecido por el propio sistema y en el avanzado es el usuario el que decide el orden de presentación haciendo uso de la opción 'Ranking', indicando el término que deben contener las páginas presentadas en primer lugar. En el formulario para búsquedas avanzadas se puede seleccionar documentos que hayan sido publicados o modificados por última vez en un periodo de tiempo establecido.

Resultados: no se puede formatear el resultado de una búsqueda. De cada página se visualiza: el título con el enlace, la URL y las primeras palabras en el caso de que no contenga etiqueta META de descripción, si es así aparece el contenido de la misma. Además, aparece el tamaño de la página (bits), el idioma y la fecha de la última modificación. Los criterios de relevancia para la ordenación de las páginas de una búsqueda simple son: si las palabras o frases demandadas se encuentran en las primeras líneas de una página; la frecuencia de ocurrencia; si las palabras o frases se una consulta múltiple se encuentran cerca una de otra en el documento. Se puede mantener abiertas simultáneamente la página de resultados y uno de los documentos presentados.

Ayudas: las ayudas detallan con ejemplos la forma de realizar las búsquedas, los operadores a utilizar y los criterios de relevancia que se aplican en la presentación de resultados.

Otros: Altavista conecta desde su página principal con 'Looksmart', índice de páginas clasificados en categorías.

Positivo: 1) Reconoce las etiquetas META y utiliza su información para indizar las páginas y como criterio de relevancia a la hora de presentación de resultados. 2) Las posibilidades de búsqueda son bastante completas: truncado, términos completos, operadores booleanos (AND, OR, NOT) y paréntesis. 3) Posibilidad de buscar por campos concretos: dominios, servidores, URLs, links, imágenes, título e idioma. 4) Posibilidad de "afinar" la búsqueda utilizando términos propuestos por el sistema. 5) Posibilidad de mantener una página abierta a la vez que se visualiza el listado de resultados. 6) Las ayudas son bastante completas. 7) Se puede recuperar páginas publicadas o revisadas en un periodo de tiempo dado. 8) Versión en español.

Negativo: 1) Carece de clasificación temática. 2) No se puede recuperar información utilizando los campos de descripción y palabras clave. 3) No se controla el vocabulario. 4) Se recuperan páginas y no recursos

de información que además no pueden ser agrupadas por servidores. 5) No se pueden diseñar los formatos de presentación.

EXCITE (Excite Inc.)

<http://www.excite.com/>

Contenido: cobertura internacional con versiones locales en Alemania, Francia, Holanda, Suecia, Reino Unido, Japón, China y Australia.

Recogida de páginas: las páginas que contiene su base de datos son localizadas e indizadas por un robot, aunque también se pueden dar de alta usando un formulario existente para tal fin. El robot desestima toda la información proveniente de las etiquetas META.

Búsqueda: existen dos formularios de búsqueda: uno con una única ventana y otro más completo. Las búsquedas pueden realizarse a través de palabras o de términos compuestos que pueden ser combinados entre sí con operadores booleanos (AND, OR y NOT). Se pueden utilizar paréntesis y truncar. No se puede limitar por campos. En la búsqueda aparecen dos funciones interesantes, por una parte la sugerencia de palabras asociadas a los términos solicitados (Search Wizard) y por otra la posibilidad de pedir documentos con las mismas características de aquel o aquellos que más se ajustan a las preferencias del solicitante (More Documents like this one). En la presentación y ayudas se indica que Excite realiza las búsquedas por conceptos, es decir que ante una búsqueda, el sistema identifica los conceptos representados por las palabras utilizadas y es capaz de localizar esos mismos conceptos. Lo que no explica las ayudas es el procedimiento exacto para realizar esta búsqueda inteligente.

Resultados: los resultados se presentan según un orden de relevancia establecido por el sistema. De cada una de las páginas seleccionadas se puede elegir entre visualizar únicamente el título y el enlace o además la descripción (primeras palabras de la página), la URL y el % de relevancia. Las páginas pueden ser agrupadas por servidores (List by Web Site).

Ayudas: en las ayudas se detalla convenientemente como realizar una búsqueda y los criterios de presentación de recursos.

Otros: también presenta una selección de páginas clasificadas por temas.

Positivo: 1) Se pueden utilizar los operadores AND, OR y NOT, utilizar los paréntesis, truncar y localizar términos compuestos. 2) Posibilidad de 'afinar' la búsqueda utilizando términos asociados propuestos por el propio sistema. 3) Se pueden pedir documentos con las mismas características a uno anteriormente seleccionado. 4) Realiza un cierto control del vocabulario en los términos de búsqueda. 5) A la hora de presentación de resultados se puede elegir entre un formato reducido (sólo título) y uno con mayor información. 6) Las páginas recuperadas pueden ser agrupadas por servidores. 7) Las ayudas proporcionan suficiente información.

Negativo: 1) Las páginas clasificadas temáticamente representan una pequeña parte del total. 2) No reconoce las etiquetas META. 3) No se puede buscar por campos, salvo por idioma. 4) La descripción de las páginas se corresponde con las primeras palabras de la página. 5) Se recuperan páginas y no recursos, aunque estas puedan ser agrupadas por servidor. 6) No se pueden mantener varias páginas abiertas a la vez que se visualiza la lista de resultados. 7) No hay forma de detectar las páginas nuevas. 8) No existe versión en español.

LYCOS (Lycos Inc. Carnegie Mellon University)

<http://www.lycos.com/>

Contenido: cobertura internacional con nodos locales en: Alemania, Bélgica, España, Estados Unidos, Francia, Holanda, Italia, Reino Unido, Suecia y Suiza.

Recogida de Páginas: las páginas son localizadas por un robot a partir de las URL's dadas de alta en el formulario que existe para tal fin, donde sólo se solicita la URL y el e-mail de la persona que realiza el alta. Ignora las etiquetas META de descripción y palabras clave.

Búsqueda: cuenta con un formulario simple y otro más complejo (Lycos Pro). En ambos se puede acotar por idioma y utilizar: operadores booleanos (AND, OR y NOT), operadores de proximidad y paréntesis. Sin embargo, no se admite el truncado. Además en el formulario complejo se puede elegir entre buscar los tér-

minos de búsqueda en todo el documento, el título o la URL. También se pueden hacer búsquedas en un web determinado existente en Lycos, indicando su dominio en el formulario de búsqueda. La opción "más resultados" permite visualizar las páginas de iguales características que una seleccionada. Se pueden localizar imágenes y sonidos.

Resultados: se puede elegir entre visualizar las páginas seleccionadas agrupadas por dominios o hacerlo individualizadamente. De cada página se visualiza: el título con el enlace, la URL, las primeras palabras a modo de resumen, el tamaño en bits de la página y el grado de pertinencia. El usuario puede elegir entre varios el criterio de relevancia a utilizar para el orden de presentación de las páginas.

Ayudas: las ayudas detallan la forma de realizar las búsquedas, los operadores a utilizar y los criterios de relevancia a aplicar en la presentación de resultados.

Otros: además de la base de datos confeccionada a partir del motor de búsqueda, Lycos contiene lo que denomina Webguide, que no es más que un índice de recursos seleccionados clasificados por temas. No tiene motor interno de búsqueda.

Positivo: 1) Se puede buscar por los siguientes campos: idioma, título y URL. 2) Se pueden hacer búsquedas dentro de un web determinado existente en Lycos. 3) Visualización de páginas con igual características que una seleccionada. 4) Se pueden localizar imágenes y sonidos. 5) Se recuperan páginas, pero pueden ser agrupadas por dominios. 6) El usuario puede elegir los criterios de relevancia en el orden de presentación de los resultados. 7) Las ayudas son bastante completas. 8) Existe versión en español.

Negativo: 1) No reconoce las etiquetas META de palabras claves ni de descripción, por lo tanto su contenido no puede ser utilizado ni en la indización ni en la presentación de resultados. 2) Sólo están clasificadas parte de las páginas. 3) Las posibilidades de búsqueda son incompletas, pues no admite el truncado. 4) No se controla el vocabulario. 5) En los resultados la descripción de la página se realiza a través de las primeras líneas del texto de la misma. 6) No se pueden detectar las novedades.

WEBCRAWLER (Excite Inc.)

<http://www.webcrawler.com/>

Contenido: cobertura internacional.

Recogida de páginas: los usuarios pueden sugerir nuevas páginas a partir de un formulario, advirtiéndose que de un web únicamente se consigne las URLs de la página principal y aquellas otras de las secciones relevantes. Webcrawler indiza todas las palabras de una página hasta 1 MB de texto. No utiliza la información de las etiquetas META. Expertos de Webcrawler confeccionan un directorio ("The WebCrawler Guide") donde una selección de "sites" se encuentra organizada en categorías.

Búsqueda: las búsquedas se realizan a través de un formulario con una única ventana donde deben consignarse los términos a localizar. Las búsquedas pueden realizarse a través de un término o de términos compuestos que pueden ser combinados entre sí con operadores booleanos (AND, OR, NOT). Se pueden utilizar paréntesis. No se puede limitar por campos. Al igual que en Excite y Magellan se pueden pedir documentos con las mismas características de aquel o aquellos que más se ajustan a las preferencias del solicitante ("Similar Pages"). También se puede consultar el directorio y seleccionar páginas descendiendo en la jerarquía hasta encontrar el epígrafe adecuado.

Resultados: existen dos formatos de presentación: uno exclusivamente con los títulos de las páginas y otro, además de con el título, con la URL, las primeras palabras a modo de descripción y el grado de relevancia. En ambos casos los resultados se ordenan según su grado de relevancia estimado a partir del número de palabras solicitadas encontradas en el documento. En el caso del directorio, en primer lugar aparecen las páginas recomendadas por WebCrawler y a continuación el resto ordenadas alfabéticamente con una breve descripción del contenido de las mismas.

Ayudas: en las ayudas se detalla como funciona el sistema y como realizar una búsqueda convenientemente. Se presentan ejemplos.

Positivo: 1) Las posibilidades de búsqueda son totalmente satisfactorias, pues se pueden utilizar los operadores AND, OR y NOT y los paréntesis, se puede truncar y localizar términos compuestos. 2) se pueden localizar páginas con las mismas características que una predeterminada. 3) Se puede elegir entre dos formatos de presentación. 4) los resultados se ordenan por criterios de relevancia. 5) Las ayudas son bastante completas.

Negativo: 1) No utiliza la información de las etiquetas META. 2) Las páginas están parcialmente clasificadas. 3) No se puede limitar por campos. 4) Se recuperan páginas que no pueden ser agrupadas por servidores. 5) La descripción de la página viene dada por las primeras líneas de la misma. 6) No se pueden detectar las novedades. 7) No existe versión en español.

HOTBOT (Wired Digital Inc.)

<http://www.hotbot.com/>

Contenido: cobertura internacional.

Recogida de Páginas: las páginas son localizadas por un robot a partir de las URLs dadas de alta en un formulario, donde debe indicarse la URL y el e-mail de la persona que la realiza. El robot extrae información de las etiquetas META de título, descripción y palabras clave.

Búsqueda: cuenta con dos formularios de persiana: uno simple y otro más complejo. En el primero se puede acotar por continente y fecha de alta. En el más complejo, además de por los conceptos anteriores, se añaden más elementos. En ambos casos se pueden utilizar las siguientes Meta words (etiquetas de campo): dominio, título, links, fechas de creación y modificación de las páginas. También se pueden localizar páginas dentro del servidor designado. Reconoce términos compuestos y se pueden utilizar: operadores booleanos (AND, OR y NOT), operadores de proximidad, paréntesis y truncado.

Resultados: se puede elegir entre tres formatos de visualización. En el formato completo de cada página se visualiza: el título con el enlace, la URL y las primeras palabras a modo de resumen o en el caso de contar con etiqueta META de descripción del contenido existente en la misma. Además aparece el tamaño de la página (bits), la fecha de alta en el sistema y el grado de relevancia. Las páginas se presentan por orden de relevancia donde se tiene en cuenta la posición de los términos solicitados (título y etiqueta Meta de palabras-clave), la frecuencia en la página y la profundidad de la misma. Hotbot es capaz de mantener simultáneamente la conexión con el listado de resultados y con una de las páginas recuperadas.

Ayudas: las ayudas detallan con ejemplos la forma de realizar las búsquedas, los operadores a utilizar y los criterios de relevancia que se aplican en la presentación de resultados.

Otros: Hotbot enlaza con el índice de páginas clasificados "Looksmart" a partir de la presentación de sus categorías.

Positivo: 1) Reconoce las etiquetas META de título, descripción y palabras clave. 2) Las posibilidades de búsqueda son muy completas, ya que se pueden utilizar todos los operadores booleanos y los paréntesis, se puede truncar y localizar palabras compuestas. 3) Se puede acotar por dominio, título, continente, país, fecha de alta o modificación de las páginas, link y por numerosos campos que tienen que ver con las características informáticas de las páginas. 4) Se puede elegir entre tres formatos de visualización. 5) Para la descripción de las páginas se utiliza el contenido de esta etiqueta Meta en el caso de que exista. 6) El orden de presentación de resultados obedece a unos criterios de relevancia. 7) Se puede mantener conexión simultáneamente con el listado de resultados y con una de las páginas recuperadas. 8) Las novedades pueden ser detectadas utilizando el campo de fecha de alta y modificación. 9) Las ayudas son bastante completas.

Negativo: 1) Las páginas no están organizadas por clasificación temática. 2) No se puede acotar ni por descripción, ni por palabra clave. 3) No se controla el vocabulario. 4) Se recuperan palabras y no recursos de información y las páginas no pueden ser agrupadas por recursos. 5) No existe versión en español.

MAGELLAN (The Mckinley Group, Inc., a subsidiary of Excite Inc)

<http://www.mckinley.com/>

Contenido: cobertura internacional.

Recogida de páginas: los usuarios pueden sugerir nuevas páginas a partir de un formulario en el que únicamente deben consignar la URL de la página a incluir. Las páginas son visitadas y dadas de alta definitivamente por un equipo de expertos de Magellan.

Búsqueda: Magellan está organizado en categorías en las que se encuentran enmarcadas las páginas visitadas y evaluadas por los expertos. Existen dos procedimientos para realizar las búsquedas: o bien descendiendo por las categorías de su clasificación hasta llegar al epígrafe adecuado o bien tecleando los términos deseados sobre un formulario de búsqueda. En este formulario se debe especificar el conjunto de documen-

to sobre el que se quiere buscar: documentos revisados por Magellan, documentos con luz verde o en la base de datos de Excite (The Entire Web). No funciona el truncado. Las búsquedas pueden realizarse a través de unitérminos o de términos compuestos que pueden ser combinados entre sí con operadores booleanos (AND, OR y NOT). Se pueden utilizar paréntesis. No se puede limitar por campos. Al igual que en Excite se pueden pedir documentos con las mismas características de aquel o aquellos que más se ajustan a las preferencias del solicitante (Find Similar). También como en Excite en la presentación se indica que el sistema realiza las búsquedas por conceptos, es decir que ante una pregunta el sistema identifica los conceptos representados por las palabras utilizadas y es capaz de localizar estos mismos conceptos en otras páginas. Lo que no se explica es el procedimiento exacto para realizar esta búsqueda inteligente.

Resultados: los resultados se presentan según un orden de relevancia establecido por el sistema, que no se explica en las ayudas. De cada una de las páginas seleccionadas aparece el título con el enlace, un pequeño resumen que sirve como descripción de la página, la URL y el % de relevancia.

Ayudas: en las ayudas se detalla convenientemente como realizar una búsqueda presentándose ejemplos.

Positivo: 1) La selección de páginas la hace un equipo de expertos del propio Magellan. 2) Todas los recursos de información existentes en la base están clasificados temáticamente. 3) Existe cierto control del vocabulario. 4) Se pueden pedir documentos de iguales características a uno designado. 5) Se recuperan unidades de información y no páginas. 6) La descripción del recurso se hace a través de un pequeño resumen. 7) Los recursos se ordenan utilizando criterios de relevancia. 8) Las ayudas a la búsqueda son completas, pero no se explican ni los criterios de indización ni los de relevancia para establecer el orden de presentación de recursos.

Negativo: 1) Las posibilidades de búsqueda no son completas, pues no se puede truncar. 2) No se puede buscar por campos concretos. 3) No se pueden seleccionar formatos. 4) No hay forma de detectar las novedades. 5) No existe versión en español.

INFOSEEK (Infoseek Corp.)

<http://guide.infoseek.com/>

Contenido: cobertura internacional con versiones locales en España, Alemania, Francia, Italia, Dinamarca, Suecia, Reino Unido y Japón.

Recogida de Páginas: las páginas que contiene su base de datos son localizadas e indizadas por un robot, también se pueden dar de alta usando un cuestionario que el propio sistema presenta donde exclusivamente se solicita la URL de la página. Las páginas se dan de alta individualizadamente. El robot es capaz de identificar las etiquetas META de palabras clave ("keywords") y de descripción existentes en las páginas utilizando la primera para la indización de la página y la segunda como resumen del contenido en la visualización de resultados.

Búsqueda: la búsqueda se puede realizar utilizando cualquiera de los dos formularios existentes: uno para búsquedas simples y otro para las más complejas. El sistema permite localizar palabras y frases que pueden ser combinadas con los operadores AND, OR y NOT, pero no se puede utilizar paréntesis. Se pueden truncar términos. Las búsquedas pueden hacerse en texto libre o bien limitarse a los siguientes campos: URL, "hyperlinks", título y lugar geográfico.

Resultados: los resultados son presentados siguiendo un orden de relevancia primándose aquellas páginas en las que los términos requeridos aparecen en el título o cerca del inicio, contienen un mayor número de ocurrencias o los términos solicitados son considerados relevantes (palabras poco comunes en una base de datos). De cada página seleccionada aparece el título con el enlace, la URL, la descripción (la que aparece en la etiqueta META "description", en el caso de que exista, o las primeras líneas que aparecen en la página propiamente dicha), el tamaño de la página (bits) y la fecha de alta en el sistema.

Ayudas: en las ayudas se detalla convenientemente como realizar una búsqueda y los criterios de presentación de recursos.

Otros: también presenta una sección de páginas clasificadas directamente por un equipo de personas de Infoseek.

Positivo: 1) Reconoce las etiquetas META de palabras-clave y de descripción. 2) Posibilidades de búsqueda bastante completas: truncado, búsqueda de términos compuestos y operadores booleanos (AND, OR,

NOT). 3) Posibilidad de buscar por los siguientes campos: URL, "hiperlink", título y lugar geográfico. 4) Las páginas recuperadas se presentan siguiendo un orden de relevancia. 5) Se utiliza el texto de la etiqueta META de descripción para resumir el contenido de cada página recuperada. 6) Detección de novedades por la fecha de alta en el sistema. 7) Las ayudas son bastante completas. 8) Versión en español.

Negativo: 1) El alta debe hacerse página a página. 2) Las páginas se encuentran parcialmente clasificadas. 3) No permite el uso del paréntesis. 4) No se puede buscar ni en el campo de palabras-clave ni en el de descripción. 5) No se controla el vocabulario. 6) Existe un único formato de presentación. 7) Se recuperan páginas y no recursos que además no pueden ser agrupadas por servidores. 8) Se conoce la fecha de alta de las páginas, pero no se puede acotar por este criterio.

GALAXY (CyberGuard Corp.)

<http://www.einet.net>

Contenido: cobertura internacional.

Recogida de Páginas: las altas se producen a partir de un formulario donde se solicita título, URL, idioma, tipo de información (institución académica, organización comercial, directorio, revista, etc), localización geográfica y descripción.

Búsqueda: existen dos procedimientos de búsqueda: utilizar uno de los dos formularios que existen o descender por la clasificación temática. El formulario simple tan sólo es una ventana donde se teclean los términos de búsqueda. El formulario destinado a las búsquedas avanzadas está compuesto de tres ventanas: una para los términos, otra para restringir a un epígrafe de la clasificación y un tercero para elegir el formato de salida. También se puede elegir entre buscar exclusivamente en el título y la URL o hacerlo en el registro completo. Se puede truncar y localizar términos compuestos y utilizar los operadores lógicos AND, OR y NOT.

Resultados: se puede elegir el formato de visualización entre tres opciones. El formato más completo contiene la siguiente información: título con el enlace, grado de pertinencia, descripción, palabras más frecuente dentro del recurso, tipo de recurso, clasificación, URL, fecha de alta en el sistema y fecha de indización. El orden de presentación obedece a ciertos criterios de relevancia que no son explicados.

Ayudas: se explica convenientemente como realizar las búsquedas y como rellenar el formulario de alta de nuevos recursos, pero no se da información sobre como funciona el sistema ni cuáles son los criterios de relevancia en el orden de presentación de recursos.

Positivo: 1) Las altas se realizan a partir de un formulario donde se solicitan: título, URL, idioma, tipo de información (institución académica, organización comercial, directorio, revista, etc), localización geográfica y descripción. 2) Los recursos están todos clasificados temáticamente. 3) Las posibilidades de búsqueda son casi completas pues se pueden utilizar todos los operadores lógicos (AND, OR y NOT), se puede truncar y buscar términos compuestos. No se pueden utilizar paréntesis. 4) Se recuperan unidades de información y no páginas. 5) Se puede seleccionar entre tres formatos de presentación. 6) La información aportada para cada registro recuperado es muy amplia, pues en el formato completo se reproducen los siguientes datos: título con el enlace, grado de pertinencia, descripción, palabras más frecuente dentro del recurso, tipo de recurso, clasificación, URL, fecha de alta en el sistema y fecha de indización. 7) Los recursos recuperados se ordenan en función de ciertos criterios de relevancia. 8) Se pueden detectar las novedades, pues existe un fichero especial donde se almacenan las nuevas incorporaciones.

Negativo: 1) A pesar de que la ficha de toma de datos es muy completa tan sólo se puede acotar por título, URL y clasificación. 2) No se controla el vocabulario. 3) Las ayudas no son todo lo completas que deberían ser, pues no explican la significación de todos los datos que aparece en el formato de visualización completo ni los criterios de relevancia utilizados para ordenar los resultados de la búsqueda.

LOOKSMART (Looksmart International Limited)

<http://www.looksmart.com/>

Contenido: cobertura internacional con versiones locales en Australia y el Reino Unido.

Recogida de Páginas: existe un formulario de alta donde los interesados pueden sugerir nuevos recursos. En dicho formulario se demanda título, URL y descripción como datos obligatorios, y palabras-clave y

fecha en la que el recurso fue puesto en la red como datos opcionales. Los datos deben de ser rellenados en el epígrafe clasificatorio en el que se piensa debe de estar enclavado el recurso. Esta información es revisada por un equipo de expertos de Looksmart quienes dan el alta definitiva.

Búsqueda: las búsquedas pueden realizarse o bien descendiendo por la clasificación temática o utilizando un formulario. El formulario es muy simple, pues tan sólo cuenta con una ventana donde han de teclearse los términos. En el formulario de búsqueda las palabras tecleadas son por defecto unidas con el operador AND. No se pueden utilizar ni operadores de búsqueda ni truncar ni localizar términos compuestos. No se puede acotar por campos. Existe la posibilidad de enlazar con el formulario para búsquedas avanzadas de Hotbot y lanzar la interrogación en este sistema.

Resultados: cuando se desciende por la clasificación lo que se obtiene son los recursos dados de alta en Looksmart. Si se utiliza el formulario de búsqueda, en primer lugar aparecerán los recursos existentes en Looksmart y a continuación las páginas obtenidas en Altavista. En el primer caso se obtiene el título con el enlace, la descripción y la categoría de la clasificación temática a la que pertenece con un enlace activo. En el segundo los datos proporcionados por Altavista. En el caso de Looksmart no se aplican criterios de relevancia en el orden de presentación de los recursos.

Ayudas: las ayudas son escasas en cuanto a la forma de realizar las búsquedas.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicita título, URL, clasificación, descripción, palabras clave y fecha de puesta en marcha del recurso. 2) Todos los recursos están clasificados siguiendo su índice temático. 3) Se recuperan recursos y no páginas. 4) Los recursos van acompañados de un resumen descriptivo del contenido. 4) Las ayudas son aceptables.

Negativo: 1) No se puede utilizar ningún operador de búsqueda ni truncar ni localizar términos compuestos. 2) No se puede acotar por campos. 3) No se controla el vocabulario. 4) No se puede seleccionar formatos. 5) No se aplica ningún criterio de relevancia en el orden de presentación de recursos. 6) No se pueden detectar las novedades.

OLE (Ole)

<http://www.ole.es/>

Contenido: webs en español.

Recogida de Páginas: el método utilizado para localizar e incluir recursos es el alta por parte de los responsables a partir de un formulario. Antes de introducir los datos es necesario posicionarse en aquella categoría de la clasificación general en la que se piensa debe ubicarse el nuevo recursos. En el formulario se solicitan los siguientes datos: título, URL, país, idioma, ámbito, zona de influencia, tipo de servicio, descripción, palabras-clave, e-mail de contacto técnico y de contacto de contenido.

Búsqueda: se puede acceder a la información de dos modos: o descendiendo por el árbol de sectores o realizando una búsqueda desde cualquier punto de este mismo árbol utilizando un formulario desde donde se definen las búsquedas debiendo decidir si se truncan las palabras, si se enlazan con AND o con OR y en que campo se desea sean localizados los términos: título, URL, descripción, palabras clave, zona de influencia, sector, idioma, ámbito, tipo de servicio y país de origen. No se puede utilizar el operador NOT ni buscar términos compuestos ni usar paréntesis. En su opción "Temas" ofrece la posibilidad de visualizar por orden alfabético todas las categorías de su clasificación y en la de "Arbol" de desplegar la clasificación por apartados.

Resultados: no existe posibilidad de elección de formato de salida, apareciendo los recursos seleccionados agrupados por categorías y dentro de cada una de ellas por orden alfabético. De cada recurso se presenta el nombre con el enlace y la descripción. Se destacan las nuevas incorporaciones colocándolas por delante una "N".

Ayudas: en la ayuda se orienta muy someramente sin ejemplos sobre como se debe realizar una búsqueda.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicitan los siguientes datos: título, URL, país, idioma, ámbito, zona de influencia, tipo de servicio, descripción, palabras-clave, e-mails de contacto técnico y de contacto de contenido. El alta debé de hacerse desde el epígrafe clasificatorio adecuado. 2) Los recursos están organizados según una clasificación temática. 3) Se puede limitar por los siguientes campos: título, URL, descripción, palabras clave, zona de influencia, sector, idioma, ámbito, tipo

de servicio y país de origen. 4) Se recuperan recursos de información y no páginas. 5) En la visualización de resultados se reproduce la descripción proporcionada por el responsable del recurso. 6) Las novedades pueden ser detectadas, pues las nuevas incorporaciones se destacan del resto durante 15 días a partir de su incorporación.

Negativo: 1) Las posibilidades de búsqueda son limitadas pues se puede truncar y utilizar los operadores AND y OR, pero no se puede utilizar el operador NOT ni el paréntesis ni localizar términos compuestos. 3) No se controla el vocabulario. 4) No se puede elegir formato de salida. 5) No se utilizan criterios de relevancia en la presentación de recursos. 6) Las ayudas no son muy completas y no presentan ejemplos de búsquedas.

DONDE? (Grupo Nuevas Tecnologías de la Información - Universidad Jaime I)

<http://donde.uji.es/>

Contenido: webs españoles. Cerca de 35.000 URLs.

Recogida de Páginas: los recursos se recogen a partir de un formulario de alta que demanda los siguientes datos: título, URL, tipo de recurso, datos institucionales (nombre, unidad o departamento y naturaleza de la organización responsable), descripción del contenido, localización geográfica (ciudad y comunidad autónoma), persona y e-mail del responsable del recurso.

Búsqueda: los recursos se encuentran agrupados por comunidades autónomas. Las búsquedas se pueden realizar en la totalidad de la base o en una comunidad autónoma determinada. Cuando se decide ir directamente a una comunidad autónoma es necesario visualizar por orden alfabético los recursos agrupados bajo esa categoría. En el caso de que se decida hacer una búsqueda en la totalidad de la base, se puede utilizar o un formulario de una única ventana o bien otro más completo donde se puede acotar por los siguientes campos: título, URL, descripción, organización responsable, localidad y comunidad autónoma. Se permite el uso del asterisco para truncar, de las comillas para la localización de términos compuestos, y de los operadores lógicos AND y OR. No se puede utilizar ni el operador lógico NOT ni los paréntesis.

Resultados: se puede elegir entre un de formato de salida reducido y otro más completo donde aparece título, organización responsable, descripción, ciudad, comunidad autónoma y grado de relevancia. La visualización está limitada a 100 registros ordenados según su nivel de relevancia. En la presentación de páginas se utiliza un orden de relevancia. En los recursos recuperados aparece un resumen con el contenido del mismo. Los recursos más recientes tienen una etiqueta que indica su novedad. También existe un apartado denominado "Lo + nuevo", donde se agrupan las últimas entradas en orden cronológico inverso.

Ayudas: en las ayudas se informa sobre el funcionamiento en general del sistema y de la manera de realizar las búsquedas (aunque con muy pocos ejemplos). No se menciona los criterios de relevancia a la hora de la presentación de recursos.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicitan los siguientes datos: título, URL, tipo de recurso, datos institucionales, descripción del contenido, localización geográfica, persona y e-mail del responsable del recurso. 2) Se puede acotar por los siguientes campos: título, URL, descripción, tipo de organización responsable, localidad y comunidad autónoma. 3) Se recuperan unidades de información y no páginas. 4) Se puede elegir entre dos formatos de salida. 5) El formato de presentación completo presenta abundante información sobre el recurso en cuestión: título, institución responsable, ciudad, comunidad autónoma, resumen de contenido y grado de relevancia. 6) Las páginas recuperadas se presentan siguiendo un orden de relevancia 7) Los recursos más nuevos presentan una etiqueta que indica su novedad.

Negativo: 1) Los recursos no están organizados atendiendo a una clasificación temática. 2) Las posibilidades de búsqueda no son completas, pues aunque se puede truncar, buscar términos compuestos y los operadores AND y OR, no se pueden usar ni el operador NOT ni los paréntesis. 3) No se controla el vocabulario. 4) Las ayudas no son muy completas, pues apenas se presentan ejemplos de búsquedas y no se mencionan los criterios de relevancia a la hora de presentación de recursos.

OZU.com (OzuCom S.L.)

<http://www.ozu.com/>

Contenido: webs españoles.

Recogida de Recursos: a través de un formulario donde debe consignarse la URL, el título y la descripción de la página junto con el e-mail en el caso de que la inscripción haya sido hecha por el creador o el administrador de la misma. Los recursos en OZU.com están organizados siguiendo una clasificación temática. El alta debe hacerse desde el epígrafe donde se piensa pueda estar incluido un recurso dado.

Búsqueda: la búsqueda puede realizarse o bien descendiendo por la clasificación temática hasta encontrar el epígrafe deseado o bien utilizando uno de los dos formularios de búsqueda que existen. Uno de ellos tan sólo tiene una única ventana donde deben escribirse los términos. El otro formulario es más completo y permite acotar por campos (URL, título y descripción), seleccionar un epígrafe determinado de la clasificación general, elegir entre utilizar los operadores AND y OR y truncar las palabras si así se desea. No se puede utilizar el operador NOT, ni los paréntesis ni localizar términos compuestos. Cuando los términos de busca son muy generales el sistema presenta los epígrafes de la clasificación y obliga a restringir a uno de ellos.

Resultados: no se puede elegir el formato de salida. Cuando se lanza una pregunta, primero se obtienen las secciones de la clasificación que concuerdan con las palabras buscadas, a continuación aparecen los recursos (título, descripción y URL) ordenados por secciones de la clasificación. Dentro de cada epígrafe clasificatorio los recursos se presentan por orden alfabético. A las páginas que han sido añadidas recientemente se les añade una etiqueta que así lo indica. Existe un apartado donde se presentan agrupados por epígrafes los recursos dados de alta recientemente.

Ayuda: se da información sobre el funcionamiento de OZU.com, y se detallan las posibilidades de búsqueda sin incluir ejemplos.

Otros: OZUEXPRESS es un servicio de búsqueda a la carta que envía periódicamente nuevas páginas relacionadas con el tema de interés expresado en el formulario existente en esta sección. Se puede elegir entre la periodicidad del envío de las novedades (mensual, semanal o diario).

Positivo: 1) El alta de los recursos se hace a través de un formulario donde se solicita la URL, el título y la descripción de la página, una vez situado en el epígrafe clasificatorio adecuado. 2) Los recursos están organizados atendiendo a una clasificación temática. 3) Se puede acotar por los siguientes campos: URL, título y descripción. 4) Se recuperan recursos y no páginas. 5) De los recursos recuperado se presenta las descripciones de los mismos. 6) Las nuevas incorporaciones se destacan con una etiqueta. 7) Las ayudas son bastante completas, aunque no presentan ejemplos de búsqueda. 8) Existe un servicio de búsqueda a la carta con envío periódico de resultados.

Negativo: 1) Las posibilidades de búsqueda son deficientes, pues no se pueden localizar términos compuestos ni utilizar el operador NOT ni los paréntesis. 2) No se puede elegir formato de salida. 3) No existen criterios de relevancia en la presentación.

ELCANO

<http://www.elcano.es/>

Contenido: webs principalmente del ámbito hispano.

Recogida de Páginas: las páginas se recogen a partir de un formulario donde se solicitan los siguientes datos: URL, título, tipo de recurso (lista de correo, publicación electrónica, etc.), apartado del índice temático donde debe insertarse, naturaleza del propietario (empresa comercial, institución académica, etc), idioma, descripción y e-mail del responsable. Todos los registros son revisados antes de ser dados definitivamente de alta en la base de datos.

Búsqueda: las búsquedas pueden realizarse o bien entrando por la clasificación temática que está dividida en dos partes diferenciadas, "Áreas de Interés" y "Profesionales y Colectivos", hasta encontrar el epígrafe adecuado, o bien a través de un formulario donde se deben escribir los términos de la búsqueda y elegir el idioma y naturaleza del propietario de las páginas. No se pueden utilizar ni operadores ni paréntesis. Las palabras son interpretadas como cadenas de caracteres y cuando se teclea más de una como frase.

Resultados: No se puede elegir el formato de salida. Los resultados aparecen ordenados por fecha de alta en el sistema, en primer lugar los más recientes. De cada recurso se presenta el título con el enlace, la des-

cripción, el idioma, el tipo de recurso, la fecha de alta, el e-mail del responsable y el nombre de la persona que ha proporcionado la información. A todos los recursos se les otorga una calificación. Ayudas: No existe ningún tipo de ayuda.

Otros: Existe un apartado de novedades donde aparecen todos los recursos dados de alta más recientemente, ordenados por fecha de alta.

Positivo: 1) El alta se realiza a partir de un formulario donde se solicitan los siguientes datos: URL, título, tipo de recurso, epígrafe clasificatorio, tipo de propietario, idioma, descripción y e-mail del responsable. 2) Existencia de una clasificación temática. 3) Se recuperan recursos de información y no páginas. 4) De cada recurso recuperado se presenta el título con el enlace, la descripción, el idioma, el tipo de recurso, la fecha de alta, el e-mail del responsable y el nombre de la persona que ha proporcionado la información. 5) Todos los recursos están valorados. 6) Los recursos se presentan por fecha de alta, lo que en todo momento permite conocer los más recientes.

Negativo: 1) Las posibilidades de búsqueda son bastante deficientes, ya que no se pueden utilizar ni operadores lógicos ni paréntesis y las palabras siempre son interpretadas como cadenas de caracteres. Se pueden localizar términos compuestos. 2) Exclusivamente se puede acotar por idioma y por tipo de propietario. 3) No se controla el vocabulario. 4) No se puede elegir el formato de salida. 5) No se aplican criterios de relevancia en la presentación de recursos. 6) No existe ningún tipo de ayuda.

OZU.es (AdverNet, S.L.)

<http://www.ozu.es/>

Contenido: webs españoles.

Recogida de Recursos: a través de un formulario donde debe consignarse la URL, el título, la descripción de la página y la sección de la clasificación temática de OZU donde se piensa debe estar ubicada. También se solicita el nombre y el e-mail de la persona que realiza el alta.

Búsqueda: la búsqueda puede realizarse o bien descendiendo por la clasificación temática hasta encontrar el epígrafe deseado o bien utilizando un formulario donde se elige entre localizar los términos de búsqueda en el documento completo o en alguno de los siguientes campos: título, descripción, URL o e-mail. También se puede realizar la búsqueda en el interior de cada una de las categorías del índice, en una comunidad autónoma o en la totalidad de OZU.

El operador utilizado por defecto en el formulario de búsqueda es el OR. Se puede utilizar el operador AND, pero no así el NOT. Tampoco se puede hacer uso de los paréntesis. Las palabras son interpretadas por el sistema como cadenas de caracteres y no como palabras delimitadas. Se permite el uso de comillas para la localización de términos compuestos.

Resultados: no se indica el número total de respuestas. No se puede elegir el formato de salida. Los resultados se presentan agrupados por epígrafes clasificatorios y dentro de cada uno de ellos por orden alfabético. De cada recurso se presenta el título con el enlace, la descripción y el e-mail del responsable del mismo.

Ayuda: no detalla los operadores que pueden ser utilizados.

Positivo: 1) El alta se realiza a través de un cuestionario donde se solicitan la URL, el título, la descripción de la página y la sección de la clasificación temática. 2) Existe una clasificación temática. 3) Se puede acotar por los siguientes campos: URL, título, descripción, y comunidad autónoma. 4) Se recuperan recursos de información y no páginas. 5) Los recursos seleccionados aparecen acompañados de una descripción.

Negativo: 1) Las posibilidades de búsqueda son limitadas pues no se puede utilizar ni el operador NOT ni usar los paréntesis. Además el sistema interpreta lo escrito como cadenas de caracteres y no como palabra completa. 2) No se controla el vocabulario. 3) No se indica el resultado total de las búsquedas. 4) No se puede elegir formatos de salida. 5) No existe ningún orden de relevancia para establecer el orden de presentación de los resultados. 6) No se detectan las novedades. 7) Las ayudas son muy exiguas.

BIWE (Cesat)

<http://biwe.cesat.es/>

Contenido: webs españoles.

Recogida de Páginas: a través de los propios usuarios y responsables de los webs, quienes una vez situados en la categoría pertinente del índice temático rellenan un formulario con los siguientes datos del web a dar de alta: título, URL, e-mail, breve descripción, localización, categoría donde se da de alta y palabras-clave. La información introducida por este sistema es procesada y verificada por el sistema informático.

Búsqueda: la búsqueda se realiza a través de dos procedimientos: o bien descendiendo por las diferentes categorías en las que están distribuidos los recursos o bien utilizando uno de los dos formulario de búsqueda que presenta. El más sencillo consiste en una única ventana donde se escriben los criterios de búsqueda realizándose por defecto la unión entre ellos. En el formulario más completo se pueden utilizar exclusivamente los operadores AND y OR; deben limitarse a los campos de búsqueda existentes: URL, título, descripción, palabras-clave; y no se permite el uso de términos compuestos ni paréntesis ni se puede trunca. Una vez obtenido el resultado de una búsqueda se puede acotar más utilizando nuevos términos.

Resultados: no existe posibilidad de elección de formato de salida, ya que siempre aparecen el título y la descripción, pero si se puede optar entre presentar los recursos agrupados por categorías o no y elegir el orden de presentación entre las siguientes opciones: "los más accedidos", "los más nuevos" o "por orden alfabético".

Ayudas: en las ayudas se aconseja sobre como realizar una búsqueda utilizando ejemplos. También se dan explicaciones sobre la presentación de los recursos.

Otros: Biwexpres.

Positivo: 1) El alta se hace a través de un formulario donde se solicita: título, URL, e-mail, descripción, localización, categoría de la clasificación y palabras-clave. 2) Todos los recursos se encuentran ordenados siguiendo una clasificación temática. 3) Se puede acotar por los siguientes campos: URL, descripción y palabras-clave. 4) Se recuperan unidades de información y no páginas. 5) Los recursos recuperados vienen acompañados por una descripción. 6) Se pueden detectar las novedades al elegir la opción "los más nuevos" entre las diferentes posibilidades de presentación de resultados. 7) Las ayudas son bastante completas.

Negativo: 1) Las posibilidades de búsqueda son muy limitadas, pues no se puede utilizar ni el operador NOT ni los paréntesis ni localizar términos compuestos ni trunca. 2) No se controla el vocabulario. 3) No se pueden seleccionar formatos de salida. 4) La ordenación de resultados no corresponde a ningún criterio de relevancia.

TROVATOR (Cambios)

<http://trovator.cambios.es/>

Contenido: webs en español: 400.000 páginas hispanas y 15.000 webs clasificados.

Recogida de Páginas: aunque Trovator cuenta con un robot capaz de moverse y localizar nuevas páginas autónomamente, existe un formulario que puede ser utilizado para dar de alta a nuevos recursos en el propio Trovator y en Señas, directorio de recursos clasificados. En este formulario se solicitan los siguientes datos de las páginas a introducir: título, URL, epígrafes de la clasificación de Señas, descripción, país de origen y e-mail.

Búsqueda: las búsquedas se realizan descendiendo por la clasificación de Señas o introduciendo los términos de búsqueda en una formulario que permite acotar por dominio. Se puede elegir entre introducir los términos directamente sin ningún vínculo entre ellos o realizar una búsqueda avanzada y utilizar todos los operadores permitidos: AND, OR, NOT, NEAR, además del asterisco para trunca. También se pueden usar paréntesis. Tan sólo se puede acotar por países.

Resultados: no se pueden elegir formatos. En el caso de Trovator aparecen las páginas ordenadas según un orden de relevancia, presentándose tan sólo la URL con el enlace y en la mayoría de los casos las primeras palabras de la página a modo de descripción. En el caso de Señas aparecen los títulos y la descripción de los recursos también sin ningún orden aparente.

Ayudas: explicación breve pero clara de cómo realizar una búsqueda. No dice nada sobre los criterios de relevancia aplicados a la presentación de las páginas.

Otros: Existe una sección de nuevos donde aparecen sin ningún orden las páginas introducidas en Trovator en la semana en curso.

Positivo: 1) Sistema mixto de recogida de nuevas páginas a través de un robot o de un formulario en el que se solicita: título, URL, epígrafes de la clasificación de Señas, descripción, país de origen y e-mail. 2) Las posibilidades de búsqueda son muy completas, pues se pueden utilizar todos los operadores booleanos y los paréntesis, se puede truncar y buscar términos compuestos. 3) Los resultados se presentan ordenados según ciertos criterios de relevancia. 4) Las ayudas son suficientes en el caso de la búsqueda, pero no explican los criterios de relevancia aplicados en la presentación de los resultados.

Negativo: 1) Los recursos están clasificados por temas sólo parcialmente. 2) Sólo se puede acotar por país. 3) No se controla el vocabulario. 4) Se recuperan principalmente páginas. 5) No se puede elegir formatos de salida. 6) No todas las páginas aparecen con una descripción propia. 7) Las páginas recuperadas no pueden ser agrupadas por servidores. 8) Es muy difícil la detección de novedades, pues aunque existe un apartado que agrupa las altas de la semana en curso, no se pueden realizar búsquedas.

TARANTULA (Intercomputers, S.A. de C.V.)

<http://www.tarantula.com.mx/>

Contenido: webs en español.

Recogida de Páginas: a través de un formulario de alta donde tan sólo es necesario escribir la URL de la página, el nombre del responsable de la misma y su e-mail. A continuación Tarántula inspeccionará las páginas automáticamente.

Búsqueda: las búsquedas se realizan a partir de un formulario con una única ventana. Las palabras tecleadas sobre él son interpretadas como cadenas de caracteres. El sistema utiliza el operador AND por defecto, no reconociendo ni el OR ni el NOT. No se puede utilizar paréntesis. No se pueden localizar términos compuestos. No se puede acotar por campos.

Resultados: no se puede elegir el formato de salida. Las páginas se presentan por orden de relevancia, apareciendo la siguiente información de las mismas: título con enlace, primeras palabras a modo de descripción, e-mail del responsable, tamaño en bits y fecha de alta. No se señalan los recursos nuevos.

Ayuda: no se da apenas información del funcionamiento del sistema en general ni de como realizar las búsquedas.

Positivo:

Negativo: 1) No existe clasificación temática. 2) Las posibilidades de búsqueda son incompletas, pues no se puede utilizar ni los operadores OR y NOT ni los paréntesis ni localizar términos compuestos. 3) No se puede acotar por campos. 4) No se controla el vocabulario. 5) Se recuperan páginas y no recursos de información que no pueden ser agrupadas por servidores. 6) No se pueden elegir formatos de salida. 7) Las páginas recuperadas se acompañan de las primeras líneas a modo de descripción. 8) No existen criterios para establecer el orden de presentación de páginas recuperadas. 9) No se distinguen las páginas nuevas. 10) Prácticamente no existen ayudas.

HISPAVISTA (Hispavista)

<http://www.hispavista.com/>

Contenido: webs en español.

Recogida de Páginas: existe un formulario para dar de alta a nuevas páginas, en el que hay que facilitar los siguientes datos: título, URL, categoría de la clasificación a elegir entre los epígrafes sugeridos, descripción y país de la página sugerida junto con el nombre y el e-mail de una persona de contacto.

Búsqueda: la búsqueda puede realizarse o bien descendiendo por la clasificación temática hasta encontrar el epígrafe deseado o bien utilizando uno de los dos formularios de búsqueda que existen. Uno de ellos tan sólo tiene una única ventana donde deben escribirse los términos. El otro formulario, difícil de encontrar, es un poco más completo y da la posibilidad de truncar y de buscar términos compuestos, aunque el sistema no es capaz de localizar términos compuestos unidos por preposiciones desde la opción palabra exacta. Las palabras por defecto se relacionan con el operador AND. El operador OR lo reconoce como AND y no funciona el NOT ni el paréntesis. No se puede acotar por campos.

Resultados: no se puede elegir el formato de salida. Cuando se lanza una pregunta, las páginas seleccionadas se presentan bajo su epígrafe clasificatorio, reproduciéndose el título y la descripción. Dentro de cada epígrafe clasificatorio los recursos se listan por orden alfabético. Existe un apartado donde se presentan agrupados por epígrafes los recursos dados de alta recientemente.

Ayuda: se da información muy general sobre el funcionamiento del sistema no detallándose adecuadamente las posibilidades de búsqueda. No se incluyen ejemplos.

Positivo: 1) Las altas se producen a través de un formulario donde se solicitan los siguientes datos: título, URL, categoría de la clasificación, descripción y país con el nombre y el e-mail de una persona de contacto. 2) Todos los recursos están ordenados siguiendo una clasificación temática. 3) Se recuperan recursos de información y no páginas aisladas. 4) Los recursos recuperados van acompañados con una descripción de los mismos. 5) Se pueden detectar las novedades en un apartado especial.

Negativo: 1) Las posibilidades de búsqueda son incompletas, pues no funcionan los operadores OR ni NOT y no se pueden utilizar paréntesis. 2) No se puede acotar por campos. 3) No existe control de vocabulario. 4) No se pueden elegir formatos de salida. 5) No existen criterios de presentación. 6) Las ayudas son muy exigüas.

SOL (Sol)

<http://www.sol.es/>

Contenido: webs españoles.

Recogida de Páginas: las páginas son recogidas por un robot que es capaz de indizarlas y de seguir los demás links. Las altas se realizan a partir de un formulario donde únicamente se solicita la dirección de la página a dar de alta.

Búsqueda: presenta un único formulario de búsqueda. Se pueden utilizar los operadores booleanos (AND, OR, NOT), las comillas para localizar términos compuestos y el asterisco para truncar palabras. No se puede utilizar paréntesis ni acotar por campos.

Resultados: se puede elegir entre visualizar el formato completo y otro simplificado. En el completo aparece el título de la página, las primeras palabras de la misma a modo de descripción, la URL, el tamaño de la página en bits y la fecha en que se han introducido en el sistema. No se destacan las páginas nuevas. Se pueden mantener abiertas simultáneamente la página de resultados y una de las recuperadas.

Ayudas: se explica cuales son los operadores que se pueden utilizar y la forma de realizar búsquedas presentando ejemplos. No se habla de los criterios de ordenación de las páginas una vez realizadas una búsqueda.

Positivo: 1) Se puede elegir entre dos formatos de visualización. 2) Se puede mantener abiertas simultáneamente la página de resultados y una de las páginas recuperadas. 3) Las ayudas son bastante completas a la hora de explicar el modo de realización de las búsquedas.

Negativo: 1) No existe clasificación temática. 2) Las posibilidades de búsqueda son incompletas, pues no se pueden utilizar los paréntesis. 3) No se puede acotar por campos. 4) No existe control de vocabulario. 5) Se recuperan páginas y no recursos de información que no pueden ser agrupadas por servidores. 6) En las páginas recuperadas aparecen las primeras líneas del texto a modo de descripción. 7) No existen criterios de relevancia en la presentación de resultados. 8) No se pueden detectar las novedades.

6. Análisis de resultados

Una vez estudiadas las características documentales de cada uno de los buscadores seleccionados y a partir de las mismas se pueden hacer las siguientes observaciones.

a) Tipos de buscadores

De los veinte buscadores seleccionados, nueve son motores de búsqueda y once índices temáticos. De los nueve motores, seis son internacionales: Altavista, Excite, Lycos, WebCrawler, Hotbot Infoseek; y tres reúnen URLs en español: Trovator, Tarántula y Sol. Al contrario, de los once índices temáticos, siete contienen

direcciones en español: Ole, Donde?, Ozu.com, Elcano, Ozu.es, Biwe e Hispavista; los cuatro restantes tienen una cobertura internacional.(Yahoo!, Magellan, Galaxy y Looksmart.

b) Recogida y análisis (ver Tabla 1)

- Buscadores Internacionales

De los seis motores de búsqueda internacionales tan sólo tres (Altavista, Infoseek y Hotbot) reconocen y utilizan las etiquetas Meta de las páginas recogidas. Los otros tres (Excite, Lycos y WebCrawler) no las reconocen o las ignoran.

Los cuatro buscadores restantes son índices que recogen la información de los recursos a través de un formulario de alta, tres de ellos (Yahoo!, Galaxy y Looksmart) solicitan todos los datos considerados imprescindibles (Título, URL, clasificación y descripción). Para los datos considerados recomendables estos son los resultados: palabras-clave (1/4), localización (1/4), idioma (1/4), tipo de recurso (1/4) y propietario de las páginas (0/4). En Magellan tan sólo se solicita la URL del recurso a dar de alta y un equipo interno extrae la información, aunque no se ha podido averiguar los datos extraídos. El que más información recoge es Galaxy, que de los nueve datos solicitados, recoge información de siete.

- Buscadores en español

De los siete índices que recogen recursos en español, excepto Donde?, que no clasifica sus direcciones, el resto recogen de cada uno de los recursos los datos considerados imprescindibles: título, URL, clasificación y descripción. Pero ninguno de ellos recoge todos los datos recomendables (palabras-clave, localización, idioma, tipo de información y tipo de propietario): Ole tres, Elcano y Donde? y Biwe dos e Hispavista uno. De los tres motores de búsqueda (Trovaor, Tarántula y Sol) no se ha podido averiguar si reconocen y analizan el contenido de las etiquetas Meta.

c) búsqueda (ver Tabla 2)

- Buscadores internacionales

- Formularios de búsqueda. Excepto WebCrawler, Magellan y Looksmart que no presentan nada más que un formulario simple para realizar las búsquedas, el resto de los buscadores dan la posibilidad de elegir entre uno simple con una única ventana y otro más completo donde plasmar estrategias de búsqueda más complejas. Sólo en el caso de Altavista las búsquedas avanzadas se realizan en un formulario de una única ventana.

- Herramientas de búsqueda. Casi todos los buscadores permiten utilizar la totalidad de los operadores de booleanos (AND, OR y NOT) y los paréntesis, truncar y buscar términos compuestos. Las excepciones son las siguientes: Yahoo, Looksmart y Galaxy no permiten utilizar el paréntesis; Infoseek, Lycos, WebCrawler, Magellan y Looksmart no permiten truncar. Además en Looksmart no se puede ni utilizar los operadores OR y NOT ni localizar términos compuestos.

- Clasificación temática. Tanto los índices como los motores de búsqueda presentan una clasificación temática a través de la cual se pueden localizar direcciones. La diferencia entre los índices y los motores es que en los primeros todas las direcciones están clasificadas, mientras que en los segundos solamente están clasificados una parte de sus recursos.

- Campos de búsqueda. De los índices se puede decir que es Yahoo! el único que permite acotar por los tres campos considerados imprescindibles (título, URL y descripción), además de por idioma. Galaxy admite que se acote por título y URL, y Magellan y Looksmart no dan la oportunidad de localizar direcciones utilizando campos determinados, a pesar de que en sus formularios de toma de datos solicitan los dos: título, URL y descripción de las direcciones a dar de alta. Galaxy pide además la localización, el idioma y el tipo de información y Looksmart las palabras clave.

- De los motores que son capaces de reconocer y extraer la información de las etiquetas Meta: Altavista permite seleccionar páginas por título, URL, localización e idioma; Infoseek por título, URL y localización; y Hotbot exclusivamente por título y localización. Del resto de los motores que afirman no utilizar las etiquetas Meta: Excite permite acotar exclusivamente por idioma, Lycos por título, URL e idioma; y WebCrawler no permite acotar.

- Control del vocabulario: ninguno de los buscadores controla el vocabulario, salvo en cierta forma Excite y Magellan. El sistema que utilizan ambos para llevar a cabo dicho control, no está lo suficientemente explicado.

- Detección de novedades: de los diez buscadores estudiados en tan sólo cuatro se pueden detectar las nuevas incorporaciones: Yahoo! (acotación por fecha de alta del recurso), Infoseek (registro de salida con fecha de alta en el sistema), Hotbot (fecha de creación y modificación de páginas) y Galaxy (fichero aparte).

- Buscadores en español

- Formularios de búsqueda. De los diez buscadores estudiados, todos presentan una pantalla de búsqueda simple, y seis de ellos (Ole, Donde?, Ozu.com, Ozu.es, Biwe e Hispavista) añaden otra más completa para realizar estrategias de búsqueda más complejas.

- Herramientas de búsqueda. Menos Elcano, el resto de buscadores pueden interseccionar los términos utilizando el operador booleano AND. Sin embargo, el grupo que da la oportunidad de utilizar el OR es menor, pues Elcano, Tarántula e Hispavista no permiten su uso. El NOT, y los paréntesis tan sólo pueden ser usados en Trovator y Sol. En todos los buscadores, salvo en Biwe, se puede truncar. Donde?, Ozú.es, Trovator y Sol dan la oportunidad de localizar términos compuestos.

- Clasificación temática. De los diez buscadores estudiados, siete de ellos, es decir la totalidad de los índices, dan la posibilidad de buscar a través de la clasificación temática utilizada para la ordenación de sus recursos. En el caso de los motores de búsqueda tan sólo Trovator clasifica parcialmente sus recursos y da la oportunidad de consulta a través de un índice.

- Campos de búsqueda: de los tres motores de búsqueda hay que señalar que tan sólo Trovator permite la interrogación por localización como único campo. En los índices las posibilidades de acotar por campos son mayores. Ole es el más completo en este aspecto, pues permite preguntar por título, URL, descripción, palabras-clave, localización e idioma. A continuación le sigue Donde? que ofrece la posibilidad de interrogar por los campos considerados como imprescindibles (título, URL y descripción), además de por localización y tipo de propietario. En Ozú.com y Ozú.es se puede preguntar igualmente por los campos considerados imprescindibles, añadiendo Ozú.es la posibilidad de acotar por localización. En Biwe se pueden utilizar los campos URL, descripción y palabras-clave. Finalmente los índices más pobres en este apartado son Elcano (idioma y tipo de propietario) e Hispavista (ninguno).

- Control de vocabulario: ninguno de los buscadores controla el vocabulario en la interrogación.

- Detección de novedades: de los tres motores de búsqueda (Trovator, Tarántula y Sol) en ninguno se puede detectar las nuevas incorporaciones. Sin embargo, la situación es la contraria para los índices, pues, salvo en Ozú.es, en el resto se pueden detectar a través de varios procedimientos: con etiquetas especiales (Ole, Donde?, Ozu.com), poniendo la fecha de alta en los registros recuperados (Elcano) y con un fichero aparte (Biwe e Hispavista)

Si se comparan de manera global los buscadores internacionales con los buscadores en español, se llega a la conclusión de que los internacionales son más completos desde el punto de vista de la búsqueda, pues alcanzan un mayor grado en los apartados de pantallas, herramientas de búsqueda, clasificación de recursos y control de vocabulario. En el caso de la búsqueda por campos las posiciones se aproximan más. En el único apartado en que los buscadores españoles aventajan a los internacionales es en la detección de novedades.

d) Resultados

- Buscadores internacionales

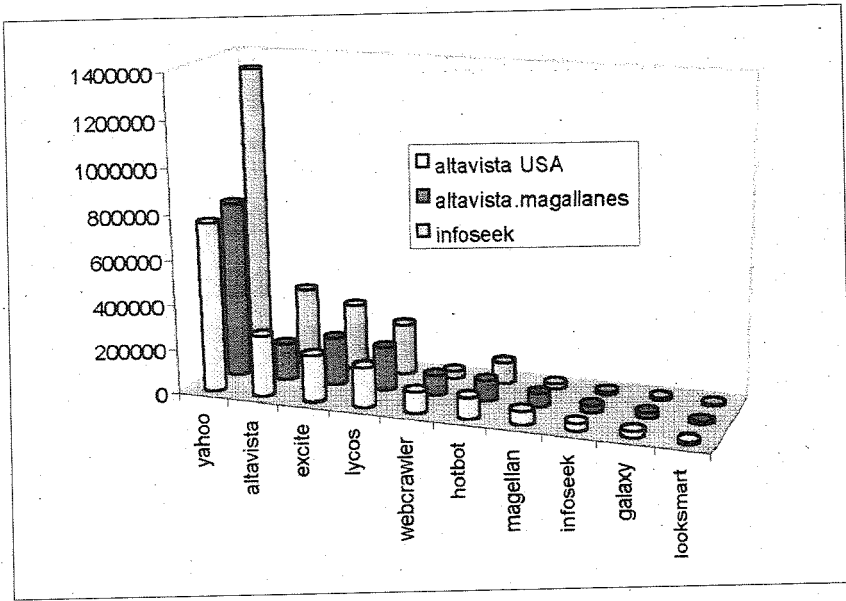
Tan sólo pueden elegirse formatos de salida en cuatro de los diez buscadores estudiados: Excite, Web-Crawler, Hotbot y Galaxy. Sin embargo, en los diez las páginas y los recursos de información vienen acompañados con una descripción de su contenido, en el caso de los motores de búsqueda representado por las primeras palabras aparecidas en las páginas o por el contenido de las etiquetas Meta en el caso de que existan. Todos los buscadores, tanto índices como motores, ordenan los registros de salida en función de unos criterios de relevancia. El punto más negativo de este apartado tiene que ver con la posibilidad de agrupación de las páginas recuperadas en los seis motores, pues tan sólo Excite y Lycos cuentan con esta opción.

- Buscadores en español

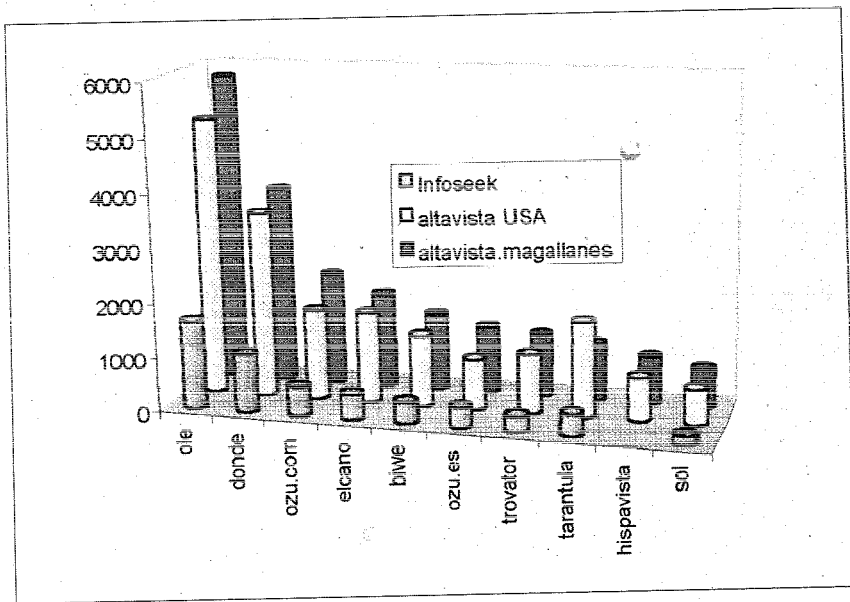
En el caso de los buscadores en español, se puede decir que en tan sólo en dos (Donde? y Sol) se puede elegir entre formatos de salida; que todos presentan una descripción de contenido de las páginas y los recursos gestionados; que únicamente en tres (Donde?, Trovator y Tarántula) se utilizan criterios de relevancia en el orden de presentación de resultados; y por último que ninguno de los motores de búsqueda permite agrupar las páginas recuperadas por servidores.

Bibliografía

- CODINA, Lluís. "Cómo funcionan los servicios de búsqueda en Internet: un informe especial para navegantes y creadores de información (Parte I)". *Information World en Español*, 1997, vol. 6, n° 5, p. 22-27.
- CODINA, Lluís. "Cómo funcionan los servicios de búsqueda en Internet: un informe especial para navegantes y creadores de información (Parte II)". *Information World en Español*, 1997, vol. 6, n° 6, p. 18-26.
- DÍEZ FERREIRA, Miguel Ángel. "Los 20 mejores buscadores". *World 9: Suplemento de Pcworld*, 1997, n° 137, pp. 34-44.
- MARCOS MORA, Mari Carmen. "Motores de recuperación de información: un análisis comparativo (parte I)". *Information World en Español*, 1998, vol. 7, n° 1-2, pp. 18-22.
- MARCOS MORA, Mari Carmen. "Motores de recuperación de información: un análisis comparativo (parte II)". *Information World en Español*, 1998, vol. 7, n° 3, pp. 13-20.
- SENSO, J. A. "Herramientas para realizar búsquedas en Internet: una revisión". *Information World en Español*, 1998, vol. 7, n° 1-2, pp. 24-25.



GRAFICA 1. BUCSADORES INTERNACIONALES



GRAFICA 2. BUCSADORES EN ESPAÑOL

Buscadores En Español (1)	TIPO DE DIRECCIONES		ROBOT				Datos Imprescindibles				INSCRIPCION					TOTAL					
	Recursos	Paginas	Meta	No Meta	?	Titulo	URL	Clas.	Descrip.	P. clave	Loc.	Idioma	T. Inform.	Prop							
Ole	X					X	X	X	X	X										7	
Donde	X					X	X	X	X											X	5
Ozu.com	X					X	X	X	X												4
Elcano	X					X	X	X	X											X	6
Oz.es	X					X	X	X	X												4
Bive	X					X	X	X	X												6
Trovator (mb)		X																			
Taramula (mb)		X																			
Hispavisia	X					X	X	X	X												5
Sol (mb)		X				X															
Total (1)	7		3	0	0	3	7	7	6		7	2	4	2	0		2				
Buscadores (2) Internacionales																					
Yahoo		X					X	X	X		X										4
Altavista (mb)		X				X															
Excite (mb)		X				X															
Lycos (mb)		X				X															
Webcrawler(mb)		X				X															
Holbot (mb)		X				X															
Magellan		X									X										1
Infoseek (mb)		X				X															
Galaxy	X						X	X	X		X		X							X	7
Looksmart	X						X	X	X		X		X								5
Total (2)	4		6	3	3	3	3	3	9		3	1	1	1	1	0					
Total (1+2)	11		9	3	3	3	10	10	3		10	3	5	3	1	2					

(mb) = motor de búsqueda

TABLA 1. RECOGIDA Y ANALISIS DE DIRECCIONES

Buscadores en Español (1)	Herramientas de Búsqueda							Clasif Tems	Pantallas			Búsqueda por Campos							Control Vocabulario	Novedades	e-mail	
	OR	AND	NOT	(")	O	*	simple		avanza	Recomendables												
										Título	URL	Descrip.	P. clave	Loc.	Idioma	Tipo Informa.	Prop.					
Ole	X	X				X	X	X	X	X	X								X			
Donde	X	X		X		X	X	X	X	X	X								X			
Ozu.com	X					X	X	X	X	X									X			
Elcano						X	X	X	X	X									X			
Oznes	X	X		X		X	X	X	X	X									X			
Bive	X	X				X	X	X	X	X									X			
Trovator (mb)	X			X		X	X	X	X	X												
Taramula (mb)						X	X	X	X	X												
Hispavisca	X	X				X	X	X	X	X									X			
Sol (mb)	X	X		X		X	X	X	X	X												
Total (1)	7	9		2		4	1	1	9	7									0		6	4
Buscadores (2)																						
Internacionales																						
Yahoo	X	X		X		X	X	X	X	X									X			X
Altavista (mb)	X	X		X		X	X	X	X	X									X			X
Excite (mb)	X	X		X		X	X	X	X	X									X			X
Lycos (mb)	X	X		X		X	X	X	X	X									X			X
Webcrawler (mb)	X	X		X		X	X	X	X	X									X			X
Hobot (mb)	X	X		X		X	X	X	X	X									X			X
Magellan	X	X		X		X	X	X	X	X									X			X
Infoseek (mb)	X	X		X		X	X	X	X	X									X			X
Galaxy	X	X		X		X	X	X	X	X									X			X
Looksmart		X																				
Total (2)	9	10		9		9	7	5	10	10									0		4	5
Total (1+2)	16	19		11		13	8	14	17	20									2		10	11

(mb) = motores de búsqueda

* Índice parcial

TABLA 2. CARACTERÍSTICAS DE LA BÚSQUEDA