

Conceptos básicos para búsquedas en la base de datos

Por Marvin Hunn

Estos son algunos términos y conceptos importantes para las bases de datos, los buscadores, y las fuentes del contenido. Usted debe familiarizarse con este material antes de intentar las tareas de las búsquedas en su clase de Orientación y Método de Investigación (como RS101 y BS1004). Aquí encontrará los ejemplos de búsquedas escritos de esta manera: << términos de la búsqueda >>.

Documentos. Para nuestros usos, esto significa un trabajo intelectual registrado. Puede ser registrado en cualquier manera. Por ejemplo, un documento puede ser un libro impreso, un libro electrónico, un artículo en una revista, un archivo en el formato de MP3, un manuscrito escrito a mano, una fotografía, una tableta de arcilla, o una pintura al óleo.

Metadato. Metadato bibliográfico es información que describe un documento, como el autor, el título, la fecha, el tema, y el resumen. Para bases de datos bibliográficos, tenemos datos (documentos) y metadatos (información descriptiva de los documentos). Metadatos son organizados en archivos, un archivo por documento y los archivos son organizados en la base de datos.

Archivos. Un archivo en la base de datos es la unidad de información en la base de datos. Consiste de un índice, o más. Cada archivo es para un documento y cada índice trata con una parte del documento en específico. Por ejemplo, un archivo puede ser de un libro y puede incluir el índice para el autor, el título, el editor, el tema, y más. Este es un archivo simplificado con los índices más comunes.

| | |
|-------------------------|---|
| Autor: | Bibfeldt, Franz |
| Título: | John Knox y la Reforma Británica |
| Información del editor: | Dallas : Nonesuch Press, c2001. |
| Serie: | Estudios en la historia de la Reforma |
| Tema: | La Reforma – Gran Bretaña |
| Tema: | Gran Bretaña – Historial de la iglesia – Siglo 16 |
| Tema: | Knox, John, ca. 1514-1572 |

Vocabulario del base de datos controlado. Este es una base de datos que usa términos estandarizados para describir los documentos. Estandarizado es para poder garantizar que los nombres o conceptos siempre se expresen en una manera consistente. Eso disminuye los problemas causados por las diferentes maneras de escribir ciertas cosas (por ejemplo: Koran o Quran o Curan) y diferentes formas de expresar algo (por ejemplo: J Smith vs John Smith) y sinónimos (por ejemplo: ira vs enojo). Un vocabulario controlado ayuda en la consistencia y referencias de términos no estandarizados a términos estandarizados.

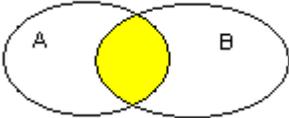
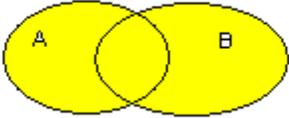
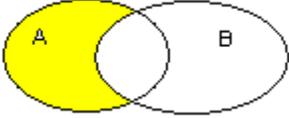
Encabezados de materia. Descriptores y encabezados de la materia son términos (palabras o frases) asignados a documentos para indicar el tema de la materia. Normalmente, estos términos so del tesoro, que lista los encabezados estandarizados.

Buscadores. Buscadores son sistemas computarizadas que intentan encontrar los artículos (archivos, documentos) cuales emparejan con las condiciones de los términos de la búsqueda. Google es un ejemplo.

Operadores de Búsquedas. Muchos operadores de búsquedas no solo requieren/permiten que usted detalle los términos de la búsqueda, sino también como la información en los términos está conectada. Los operadores le comunican a los buscadores como conectar los términos.

Frases de Búsquedas. Una frase de búsquedas es una combinación de términos, operadores y opciones que constituyen la búsqueda. Por ejemplo, << (ley OR pacto) AND (romanos OR gálatas) >> es una frase de búsqueda. Esta definición asume que la persona está usando el lenguaje con los operadores dictando lo que la persona está buscando.

Operadores Lógicas de Búsqueda. La mayoría de los buscadores usan los operadores lógicas AND, OR, NOT (Y, O, NO).

| Operador o función | Símbolo Común | Ejemplo | Explicación | |
|--------------------|---------------|-----------------------------------|--|--|
| intersección | AND (Y) | A AND B iglesia AND estado |  | 'AND' encuentra los archivos que incluyen los dos términos. |
| unión | OR (O) | A OR B clero OR pastor |  | 'OR' encuentra los archivos que incluyen cualquiera de los dos términos. |
| exclusión | NOT (NO) | A NOT B espíritu NOT santo |  | 'NOT' no incluye los archivos que contienen el segundo término. |

Agrupamientos. Muchos sistemas de búsquedas permiten que la persona use paréntesis para agrupar términos y especificar el orden en que la búsqueda será hecha. Por ejemplo, considere esta búsqueda: << café AND perro OR gato >>. Es bien claro que el perro debe de ser color café. ¿Qué del gato? La mayoría de los sistemas encontrarían gatos de cualquier color. Para especificar que el perro y el gato deberían ser del color café, usaríamos los paréntesis de esta manera: << café AND (perro OR gato) >>. Esto determina que los resultados deben de incluir perro OR (O) gato y esos resultados son filtrados por AND y la lista de resultados será de perros y gatos color café.

Operador de Truncamiento. Este operador permite resultados de partes de palabras. Se marca usando el * (asterisco). Por ejemplo, en algunos sistemas, bapt* resulta en cualquier palabra que empieza con las letras "bapt". En algunos sistemas, uno debe de usar el asterisco (*) para resultados en el posesivo. Es bien fácil incluir resultados que uno no está buscando cuando uno trunca los términos de la búsqueda. Los sistemas no dejan que uno use el asterisco al comienzo de la palabra. Algunos sistemas tienen

operadores que uno puede usar entre medio de la palabra (por ejemplo, muj#er resultaría en mujer o mujeres).

Operadores Predeterminados. El buscador usara los operadores predeterminados si la frase de la búsqueda no indica como filtrar varios términos. Por ejemplo, si uno busca dos palabras como << palabra1 palabra2 >>, el buscador usara AND (y) entre medio de las palabras, o buscara la frase exacta o algo similar.

Códigos de Índice. Los códigos de índice son los índices que son examinados. Por ejemplo, << au:calvin AND ti:institutos >> puede resultar en una búsqueda por el autor 'calvin' y el título que incluye 'institutos'. Los códigos usualmente son abreviaciones de dos letras. Muchos de los sistemas incluyen una manera de escoger los índices que uno desea buscar y uno no tiene que memorizar los códigos, pero no permite que uno busque con frases complicadas.

URL. Uniform Resource Locator (Localizador Uniforme de Recursos). El enlace. Por ejemplo:
<http://www.dts.edu>

Palabra Clave. La palabra clave es cualquier palabra en la búsqueda. No significa que es la palabra más importante, es solo una palabra que se encuentra en cualquier parte del archivo o el documento.

Pare-palabras. Estas son palabras específicas que el sistema no usa en la búsqueda. Para ahorrar tiempo, el sistema ignora ciertas palabras comunes que no tienen mucho significado, como conjunciones, preposiciones, y artículos. En ciertas búsquedas, palabras como "el" o "la" o "un" no son necesarios. Algunos sistemas incluirán estos tipos de palabras en la búsqueda si están entre comillas.

Mala coordinación es una relación semántica no desea entre palabras juntadas en una búsqueda de varias palabras. Uno ejemplo es alguien buscando "car polish" (esmalte de carro) usando "car AND polish" y encuentra información para "Polish car" (carro polaco). Esto es común cuando uno usa AND para conectar palabras en una búsqueda. Es menos común cuando uno usa frases exactas en la búsqueda.

Un **Resultado Falso** es un archivo/documento en la lista de resultados que no es relevante a la búsqueda.

Relevancia de los resultados de la búsqueda. La relevancia de los resultados depende del contexto y cambia, dependiendo del tiempo y la perspectiva de la persona. Un documento es relevante si trata del tema correcto – la relevancia depende del tema. Esto no es completamente correcto. In otro sentido, el documento es relevante si provee la información que la persona está buscando. También hay la relevancia que depende si el documento es de ayuda a la persona. Documentos pueden ser relevantes si tratan de cierto tema o ayudan en aclarar un problema que no necesariamente tiene que ver con el tema original.

Precisión en la búsqueda y los resultados. No espere que sus resultas sean una lista perfecta. El éxito de la búsqueda es vista en la precisión de los resultados y la lista de los resultados. Lo más preciso los resultados, la lista será más corta y vice versa.

Bases de datos de imágenes versus Base de datos textuales. Muchas bases de datos incluyen metadato y enlaces a documentos completos (para leer los libros o artículos en línea). Estos documentos pueden estar

disponibles textualmente o como imágenes (o los dos). Una base de datos de imágenes provee imágenes escaneados para tener acceso en línea.

Muchas veces, “textuales” significa “documento completo”. Uno podría ver un archivo de sonido en el formato MP3 (por ejemplo, sermones) aunque no es un documento textual. Hay ciertas bases de datos textuales que no tienen texto.

Sugerencia automática. Este programa automáticamente muestra una lista de términos (frases) que son similares a lo que uno está escribiendo. Por ejemplo, si uno empieza a ingresar << deducción de impuestos >> las siguientes sugerencias aparecerán antes de que uno termine de escribir la frase:

- Límites de deducción de impuestos
- Calculadora de deducción de impuestos

Navegando versus Búsqueda. Ciertas bases de datos tienen dos formas principales de encontrar recursos. 1) buscando palabras o 2) navegando una lista de títulos. La búsqueda encuentra archivos basados en la combinación de palabras en cualquier parte del documento. Esto es lo que uno usa, usualmente. Navegando es completamente diferente. Es un proceso de dos pasos. Primero, ingrese una palabra o frase para encontrar ciertos índices. Después, uno escoge de una lista de títulos más específicos en la lista producida. Por ejemplo, empiece con << Jesucristo >>. El sistema responderá con una lista que empieza así:

- Jesucristo – apariencias
- Jesucristo – ascensión
- Jesucristo – autoridad
- Jesucristo – bautismo

Usted escoge una de las opciones. Esto es similar a las sugerencias automáticas.

Textos Marcados. Un documento esta “marcado” si cierta información y vocabulario específico ha sido ingresado al texto para proveer información de partes del texto. Por ejemplo, si el documento incluye dialogo, puede ser marcado con información de quien está hablando. Esto hace posible que uno busque las palabras dichas por esa persona, específicamente. LOGOS tiene muchos documentos marcados.

Orden de Rango. Algunos sistemas ordenan los documentos en una manera que intentan mostrar los mejores documentos primero. El orden de rango depende de una combinación de factores como relevancia, popularidad, y otras cosas.

Para más definiciones, visite: <http://library.dts.edu/Pages/RM/Helps/glossary.shtml>