

LA ECUACIÓN DE BÚSQUEDA

Contenido

LA ECUACIÓN DE BÚSQUEDA	1
1. OPERADORES	1
A. Operadores Lógicos o Booleanos.....	1
B. Operadores Sintácticos o de Proximidad	4
C. Operadores de Comparación	7
2. BÚSQUEDA POR RANGOS DE VALORES	7
3. LIMITADORES DE BÚSQUEDA (EN, IN)	8
4. TRUNCAMIENTOS Y CARACTERES COMODÍN	8
5. PARÉNTESIS	9
6. ALGUNAS CONSIDERACIONES SOBRE LOS OPERADORES Y LOS PARÉNTESIS ...	11

La **Ecuación de Búsqueda** es la formulación estructurada de las expresiones de búsqueda.

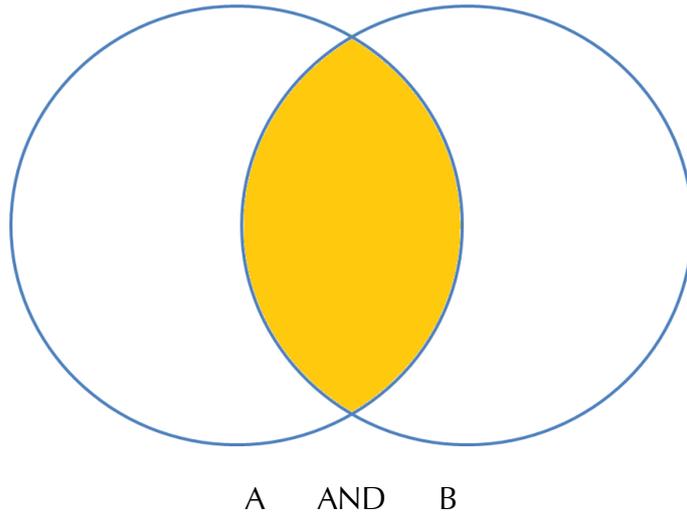
En la ecuación de búsqueda se tienen que considerar: diversos operadores, los rangos de valores, los limitadores, los truncamientos y caracteres comodín y los paréntesis.

1. OPERADORES

A. Operadores Lógicos o Booleanos

- Intersección (Y, AND, ET)

Es un operador de intersección, producto lógico o conjunción. Este operador se utiliza para realizar búsquedas de registros que contengan todos los términos que representan los principales conceptos.



Ejemplo:

Queremos localizar en una base de datos aquellos trabajos en los que conste la **contaminación** en **Almería**.

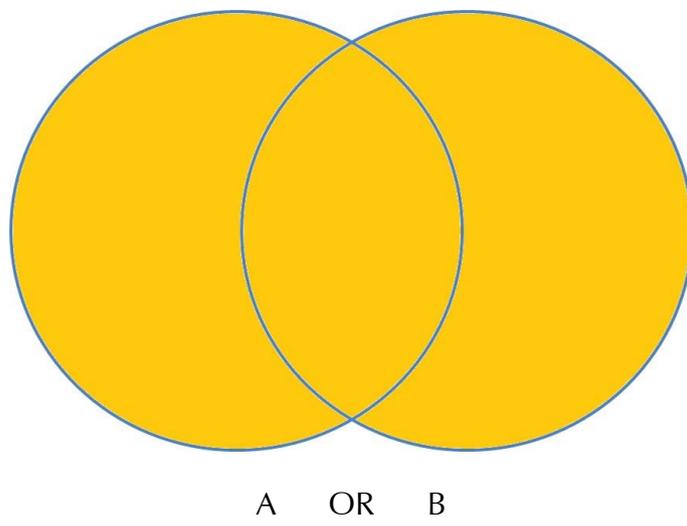
La **ecuación de búsqueda** que realizaríamos sería:

Contaminación AND Almería

- **Reunión o Suma Lógica (O, OR, OU)**

Es un operador de unión lógica o disyunción. Este operador se utiliza para realizar búsquedas que contengan todos los descriptores o palabras claves que describen un mismo concepto.

La suma lógica no tiene por qué coincidir con la aritmética, ya que si un mismo documento contiene los dos términos de una suma lógica, el programa de recuperación no los cuenta dos veces sino una; es decir, que en la recuperación de información, este operador elimina los documentos duplicados.



Ejemplo:

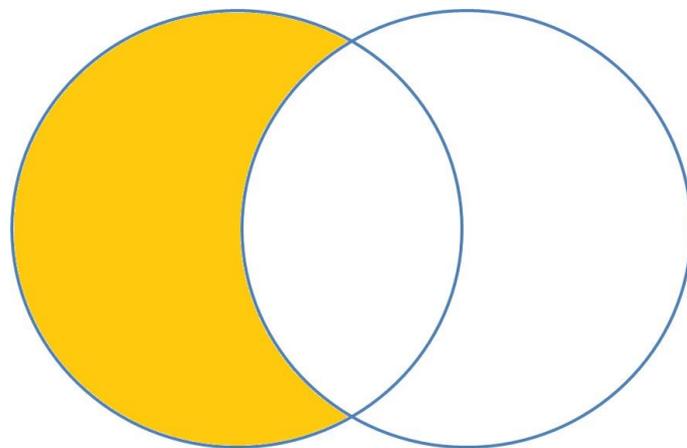
Queremos localizar en una base de datos aquellos artículos que contengan los términos **Contaminación** o **Polución**

La **ecuación de búsqueda** que realizaríamos sería:

Contaminación **OR** Polución

- **Negación (NO, NOT, AND NOT, BUT NOT, NON, SAUF)**

Es un operador de negación lógica o exclusión. Este operador se utiliza para realizar búsquedas de registros que contengan el primer término y no el otro.



A AND NOT B

Ejemplo:

Queremos localizar en una base de datos aquellos documentos que contengan el término **Polución** y no contengan el término **Automóvil**.

La **ecuación de búsqueda** que realizaríamos sería:

Polución NOT Automóvil

B. Operadores Sintácticos o de Proximidad

Los operadores sintácticos se intercalan entre los descriptores para establecer entre ellos una relación de proximidad, es decir, permiten concretar la posición de dos o más palabras en la ecuación de búsqueda y por tanto en el registro. Son similares al operador de intersección pero impone más condiciones respecto a la localización de los términos en el registro.

- **Adyacencia (ADJ, ADY)**

Este operador permite recuperar registros en los que los términos que figuran en la pregunta aparecen uno junto a otro dentro del documento, sin ninguna palabra que los separe.

Ejemplo:

Queremos localizar en una base de datos documentos que contengan la expresión **química orgánica**.

La **ecuación de búsqueda** que realizaríamos sería:

Química ADJ orgánica

- **Presencia de términos a una distancia de cómo máximo "N" palabras (W)**

Este operador permite recuperar registros en los que los términos de la pregunta aparecen en el documento separados uno de otro por un número de palabras igual o menor a la que se indica.

Ejemplo:

En la **ecuación de búsqueda** ponemos:

Química 6w orgánica

Recuperaríamos documentos en los que se dijera:

- La **química orgánica** ha avanzado mucho en los últimos años
- La **química inorgánica** avanza más que la **orgánica**.

No recuperaríamos documentos en los que se dijera:

- La **química inorgánica** resulta más complicada a los estudiantes que la **orgánica**

- **Presencia dentro de una misma frase (NEAR, SENTENCE)**

Este operador se utiliza para localizar aquellos documentos que contienen ambos términos en la misma frase.

Ejemplo:

Queremos localizar en una base de datos aquellos artículos que contengan en la misma frase **tridimensional** y **espacio**.

La **búsqueda** que realizaríamos sería:

tridimensional NEAR espacio

Nota: Este operador permite especificar el número de palabras que pueden existir entre los dos términos.

Ejemplo:

Queremos localizar aquellos artículos en los que conste en la misma frase **tridimensional** y **espacio**, como máximo separado por dos palabras.

En este caso, la **estrategia de búsqueda** sería:

tridimensional NEAR2 espacio

- **Uso de Comillas (" ")**

Busca los términos tal y como aparecen entre las comillas, la frase exacta.

Ejemplo:

Queremos recuperar todos los documentos que contienen la expresión **administración y dirección de empresas** exactamente expresada así.

"administración y dirección de empresas"

- **Dos términos en el mismo campo sin importar el orden en que se encuentren (F, WITH, W)**

Este operador permite recuperar aquellos registros en los que se encuentran los dos términos buscados en el mismo campo.

Ejemplo: Queremos localizar en una base de datos los documentos en los que figuran las palabras **química** y **orgánica** en el título o en el resumen.

La **búsqueda** que realizaríamos sería:

química W orgánica

C. Operadores de Comparación

Permiten poner condiciones a los campos con formato con longitud fija y en concreto a los campos con datos numéricos (fechas, números de documentos, números de registro, datos numéricos, etc...)

Los signos utilizados son:

>	Mayor que
<	Menor que
=	Igual
>=	Mayor o igual que
<=	Menor o igual que

2. BÚSQUEDA POR RANGOS DE VALORES

Se utiliza también con campos numéricos para especificar, entre qué límites ha de estar un valor. Se trata, por lo tanto, de un “mayor que y menor que”

Ejemplo:

Queremos limitar una búsqueda a documentos que hayan sido publicados entre 1990 y 1995. Se pondría así:

1990 to 1995

1990..1995

3. LIMITADORES DE BÚSQUEDA (EN, IN)

El operador **IN** se utiliza para realizar búsquedas de registros que contengan un término en un campo específico.

Ejemplo:

Queremos localizar en una base de datos aquellos trabajos en los que figure el término **tridimensional** en el campo título.

Nota: La abreviatura del campo título es **ti**

La **búsqueda** que realizaríamos sería:

tridimensional EN ti

4. TRUNCAMIENTOS Y CARACTERES COMODÍN

Los términos seleccionados para realizar la búsqueda pueden truncarse o segmentarse por el lexema común a varios términos.

En relación con los truncamientos hay que tener en consideración que:

- Es **aconsejable en término largos** e inequívocos pues se obtienen formas importantes de un mismo término (gerundios, participios, plurales, etc.) así como sustantivaciones o adjetivaciones de verbos.
- No es aconsejable hacer un truncamiento con menos de cuatro letras en la raíz.
- Los truncamientos que dan lugar a la selección de un número muy grande de términos corren el **riesgo de generar ruido** y bloquear el ordenador, pues éste necesitará mucho tiempo para realizar todas las sumas lógicas.
- Debe utilizarse con precaución para evitar que se introduzca ruido con términos impensados.
- Permite **aumentar la exhaustividad** de la recuperación.

Existen varios tipos de **truncamientos y comodines**. Los más comunes son:

* (**asterisco**). Colocado al principio o final de la palabra, sustituye un número indeterminado de caracteres.

Puede adoptar otras formas según la fuente consultada, p. ej.: ! (exclamación) o \$ (dólar).

Ejemplo:

Econom* recuperaría documentos que contengan los siguientes términos: Economía, Econometría, Económico/a, Económicos/as, Economy, Economic, Economics...

? (**interrogación**). Colocado en cualquier lugar de la palabra sustituye un sólo carácter.

Ejemplo:

Wom?n, recuperaría woman y women

5. PARÉNTESIS

Los paréntesis pueden ser usados:

- Cuando se desea introducir más de un término de búsqueda en un mismo campo utilizando cualquier operador lógico. De esta manera se evita volver a teclear el prefijo del campo y el delimitador con cada uno de los términos.
- Si se van a utilizar diferentes operadores lógicos en la misma etapa de búsqueda, aunque sea el mismo campo. Es conveniente el uso de paréntesis, ya que estos establecen también precedencias (las operaciones entre paréntesis se realizan siempre en primer lugar), evitando así posibles errores o ambigüedades.
- Cuando se va a operar con diferentes etapas de búsqueda y distintos operadores lógicos.

Ejemplos:

(**gene AND therapy**) Article Title. Recupera documentos que contengan estas dos palabras en el título del artículo.

(ácidos grasos AND (leche OR pescado OR carne)) Article Title.

Recupera documentos que contengan en el título del artículo **ácidos grasos** en la **leche** o en el **pescado** o en la **carne**.

Esta ecuación está compuesta por dos subecuaciones :

- Parte A: **Ácidos grasos** y
- Parte B: **leche, pescado** o **carne**.

La parte B está conectada con el operador OR, y por tanto hay que encerrarla entre paréntesis. La parte A y B también deben estar conectadas, pero con el operador AND y, por tanto, también deben estar encerradas entre paréntesis. De ahí que aparezcan dobles paréntesis al final de la ecuación.

(ácidos grasos AND (leche OR pescado) NOT carne)) Article Title.

Recupera documentos que contengan en el título del artículo **ácidos grasos** en la **leche** o la **carne**, pero no en el **pescado**.

En este caso tenemos una ecuación de búsqueda formada por tres bloques:

- Parte A: **Ácidos grasos**;
- Parte B: **leche** o **carne**;
- Parte C: **pescado**.

La parte B está conectada con el operador OR y hay que encerrarla entre paréntesis (primer grupo de paréntesis). La parte A y B están conectadas con el operador AND y también hay que encerrarlas entre paréntesis (segundo grupo de paréntesis) y este grupo formado por las partes A y B hay que conectarlos (en bloque) a través del operador NOT con la parte C (ya no hay paréntesis), pero toda la ecuación, desde que comienza hasta que acaba se encierra entre paréntesis.

6. ALGUNAS CONSIDERACIONES SOBRE LOS OPERADORES Y LOS PARÉNTESIS

- Existe el **orden de precedencia**. Al lanzar una estrategia de búsqueda el sistema establece un orden en la ejecución de los operadores, y que no es idéntico en todas las bases de datos. Lo más característico es:
 - De izquierda a derecha.
 - Primero el **AND**, luego el **OR** y el **NOT**.
 - El paréntesis siempre se ejecuta primero
- El operador **NOT** a veces puede ocasionar errores en la búsqueda. Por ejemplo en una búsqueda queremos localizar aquellos artículos que tratan sobre **visión tridimensional** y no sobre el **espacio tridimensional**. Cuando en el área de entrada de la búsqueda introducimos “**tridimensional NOT espacio**”, excluiríamos todos los artículos que hablen de **espacio tridimensional** pero también podríamos eliminar artículos relevantes para nuestro estudio.