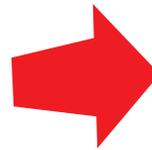




Búsqueda de información mediante términos de la lengua natural en fuentes electrónicas y evaluación de las mismas



Búsqueda de información empleando términos de la lengua natural:



El usuario elige y combina palabras o frases que podrían aparecer en un texto cualquiera que tratara el tema elegido.

Algunas características de este modo de búsqueda:

- Ampliamente utilizado
- De aparente simplicidad.
- Los bibliotecarios suelen utilizarlo como un complemento de otros recursos.
- Los usuarios finales obtienen por lo general resultados útiles, pero no exhaustivos.
- Frecuentes resultados de baja relevancia.
- La poca normalización y los frecuentes cambios de las interfaces y los lenguajes confunden a los usuarios, que suelen utilizar en un sistema tácticas adaptadas a las características de otro.
- Algunos sistemas admiten consultas redactadas como preguntas, pero los resultados no siempre son satisfactorios.

Estrategias comunes en los usuarios:

- Apuntar a temas amplios.
- Emplear algunos términos que parecen relevantes, para seleccionar luego ítems pertinentes dentro de los resultados parciales con estrategias heurísticas.

Causas de dificultades:

- Mal diseño de las interfaces, en particular de las ayudas.
- Inadecuado manejo de las técnicas de búsqueda.
- El usuario intenta describir algo que no conoce.

Soluciones:

Las dificultades pueden enfrentarse mejor si el público dispone de ayudas adecuadas o participa de actividades de capacitación y formación. Para ser utilizado en las mismas, se generó un modelo descriptivo. No todas las acciones enunciadas pueden llevarse a cabo en cualquier sistema. El modelo tiene aplicación en el desarrollo de un criterio para evaluar las fuentes con el fin de estudiar la conveniencia de incorporarlas a las colecciones reales o virtuales de las bibliotecas, los tipos de usuarios a los que pueden ser recomendadas, etc.

Operadores usuales:

Operador	Acción general	Notas	Ejemplos de simbología utilizada
Intersección de conjuntos de términos	Selecciona documentos que contienen ambos términos que se encuentran a los lados del operador.	Ver además INCLUSIÓN y PROXIMIDAD DE TÉRMINOS.	and & *
Suma de conjuntos de términos	Selecciona documentos que contienen al menos uno de los términos que se encuentran a los lados del operador.		or +
Diferencia de conjuntos de términos	Excluye los documentos que contienen el término que se encuentra a continuación del operador.	Ver además EXCLUSIÓN	not and not ! ^
Proximidad de términos	Selecciona documentos donde los términos que se encuentran a ambos lados del operador se encuentran cercanos uno del otro (adyacentes, con hasta un determinado número de palabras en el medio, etc.)	Es posible lograr resultados parecidos con frases exactas; ambos operadores ubican términos en un contexto. Son formas más exigentes de la intersección de conjuntos	near ~ adj \$. (f) (g) [con ligeras variantes en su modo de acción]
Frase exacta	Selecciona documentos en los que está presente la expresión compuesta, en el orden indicado.	Ver además PROXIMIDAD DE TÉRMINOS.	**
Paréntesis	Indica niveles de precedencia en las operaciones.		()
Inclusión - exclusión	Indica que el término que sigue al operador debe estar presente (+) en los documentos seleccionados, o ausente (-) en los mismos.	La inclusión puede usarse para producir un efecto semejante a la intersección de conjuntos; la exclusión, a la diferencia de conjuntos	+ -
Comodin	Se intercala en un término de búsqueda para indicar que en esa posición puede haber cualquier carácter (o cadena de caracteres, según el caso)	Los comodines y el truncamiento sustituyen la enumeración de términos emparentados ortográficamente	? * % [con ligeras variantes en su modo de acción]
Truncamiento	Se ubica a continuación de un fragmento de una palabra, para indicar que se buscan documentos que contengan términos que comiencen con la raíz expuesta, sin importar cual sea la terminación.	Ver además COMODIN	? * \$

Acciones usuales durante procesos de búsqueda en lengua natural:

Búsquedas de algunos documentos o fragmentos acerca de un tema (no exhaustivas)
 Seleccionar alguna palabra importante que represente, aunque sea de manera parcial, el tema; en lo posible debe ser poco frecuente fuera de ese contexto.
 Si el sistema lo admite, puede intentarse:
 Ingresar varias palabras sin indicar operadores específicos.
 Utilizar preguntas en lengua natural.

Realimentación de la búsqueda (si no se halló algún registro, documento o fragmento que corresponda a lo esperado):
 Intentar con otra palabra (sinónimo, o término de significado más general o más específico, según las necesidades)
 Revisar los resultados para detectar nuevas pistas, términos a emplear, etc.
 Revisar la ortografía de las palabras empleadas.
 Si luego de varios ensayos no se encuentra información, intentar con algunas de las acciones empleadas para búsquedas exhaustivas.

Búsquedas temáticas que intentan ser exhaustivas:

Determinar expresiones que describan el tema.
 Analizar las palabras presentes y sus relaciones de proximidad mutuas.
 Buscar sinónimos o cuasi-sinónimos (recordando incluir las variantes ortográficas), así como términos relacionados (aunque no se utilicen para nominar al tema en sí; por ejemplo, pueden elegirse nombres que correspondan a aspectos más específicos del mismo o asuntos muy frecuentemente relacionados). Repetir los dos primeros pasos.
 Seleccionar un conjunto mínimo de términos que representen todas las variantes consideradas. No es necesario incluir todas las palabras que conforman cada enunciación: se prefieren aquéllas que puedan estar menos representadas en registros que correspondan a temas diferentes, o sean menos ambiguas.
 Descartar las palabras que aportan poco sentido en el contexto (por ejemplo, la palabra Instituto en una base de datos sobre Institutos).
 Seleccionar raíces que representen varios términos.
 Descartar (si se estima que disminuirán demasiado la precisión) las raíces que pueden referir a expresiones no relacionadas con el tema. Si es necesario, se las reemplaza por una enumeración de términos completos.
 Descartar raíces demasiado breves, ya que pueden provocar demoras y otros inconvenientes como la inclusión de términos no deseados.
 Conectar las palabras que adquieren un sentido particular estando cerca de otras, por medio de operadores de proximidad o bien integrarlas en una frase.
 Efectuar una suma de conjuntos con cada grupo de sinónimos, cuasi-sinónimos o términos relacionados, conformando clases.
 Asociar las diversas clases que deben estar presentes mediante intersección de conjuntos, para restringir el resultado.
 Si hay relaciones de precedencia, utilizar los paréntesis que correspondan.
 Si es posible, reducir la longitud de las expresiones usando paréntesis, mediante la propiedad distributiva de los operadores.
 Si es necesario, restringir el resultado mediante resta de conjuntos; esta opción debe evitarse cuanto sea posible, ya que puede llevar a la pérdida de información relevante (quizás los términos restados aparezcan en contextos que no justifican excluir los registros que los contienen; por ejemplo enumeraciones cualesquiera, comparaciones o incluso frases donde se aclara que los conceptos representados por ellos están fuera del enfoque del documento).
 Si se desea restringir mediante intersección de conjuntos, pero no hay certeza de que los términos que conforman esa clase estén siempre presentes cuando se trata el tema, es conveniente no hacerlo salvo que el resultado sea demasiado extenso. En general es preferible revisar algunos registros más y no perder información relevante.

Realimentación de la búsqueda:

Estudiar si hay palabras que produjeron problemas (demasiado generales, demasiado específicas, polisémicas, etc.)
 Si hay frases, revisar el espaciado, la puntuación, etc.
 Revisar la ortografía
 Revisar los resultados para detectar nuevas pistas, términos a emplear, etc.

Si hay exceso de recuperación:

Estudiar si es factible reemplazar términos de alcance general por otros más específicos, o bien suprimirlos.
 Estudiar si se puede acotar el resultado agregando relaciones de proximidad u otras restricciones.

Si hay escasa recuperación:

Agregar sinónimos, términos más específicos o relacionados.
 Suprimir restricciones (operadores de proximidad, intersecciones o diferencias de conjuntos)
 Buscar palabras que representen temas más generales.

Búsquedas de documentos o fragmentos específicos, de los cuales se conoce parte del texto:

Emplear una expresión que figure en el texto. Puede ser una sola palabra poco frecuente, sobre todo si la base de datos es pequeña; también una expresión compuesta mediante la intersección de conjuntos, operadores de proximidad o una frase exacta.