

MANUAL DEL PARTICIPANTE

# **Curso básico de Exploradores 2.0**

---

2011 © Universidad de Colima  
Dirección General de Tecnologías para el Conocimiento  
Subdirección de Tecnologías para el Conocimiento  
Av. Gonzalo de Sandoval 444, Colima, México  
Teléfono/ Fax + (312) 3161121





## 3. Estrategias de búsqueda para recuperación de información

*Al finalizar el módulo, el participante conocerá qué son las estrategias de búsqueda para la recuperación de información y cómo se utilizan*

Este módulo es uno de los más importantes, ya que con una buena estrategia, las búsquedas que realicemos seguramente serán exitosas.

¿Cuántas veces no hemos perdido horas buscando información y al final obtenemos resultados decepcionantes?

¿Cuántas veces nos hemos preguntado “cómo busco ese tema”?

¿Cuántas veces la cantidad de información que hemos encontrado es tanta que hasta nos da flojera revisarla?

En la actualidad existe una sobreabundancia de información que dificulta que podamos extraer algo útil, y el no poder manejar tanta información comienza a darnos angustia. Cornella A., (2000) designó como **infoxicación** a la “saturación de información y ruido informativo por una persona, que impide a la mayoría de profesionales definir adecuadamente sus necesidades”, entonces podemos decir que con estas cantidades exageradas de información la infoxicación ha alcanzado su mayor expresión.

Algo muy importante que se debe tener en cuenta es que cantidad de información no implica calidad de información, a no ser que ésta sea tratada adecuadamente.

El exceso de información nos puede llevar al desorden y a situaciones de incertidumbre por la imposibilidad de analizar y valorar la información para poder decidir de un modo adecuado qué es lo que realmente necesitamos, qué información nos es útil. Por este motivo, se han hecho imprescindibles tanto la habilidad para localizar información de diversas fuentes de las que

Internet puede ser la más potente pero no la única, como la de analizar y distinguir lo que es relevante para nosotros dependiendo de lo que estamos buscando y en el detalle en que lo necesitamos.

Lo primero que se debe tener en claro es QUÉ es lo que vamos a buscar. Independientemente de la herramienta de búsqueda que vayamos a utilizar, la habilidad que desarrollemos para encontrar información relevante sobre una materia nos será bastante útil a lo largo de nuestros estudios o investigaciones.

Para adquirir esta habilidad es sumamente necesario seguir estas cuatro acciones:

1. Definir la necesidad de información ---> qué estoy buscando
2. Establecer y aplicar una buena estrategia de búsqueda ---> cómo voy a buscarlo
3. Localizar las fuentes de información adecuadas ---> en donde voy a buscarlo
4. Evaluar los resultados obtenidos ---> ¿encontré lo que buscaba?

Tipos de búsquedas que podemos realizar:

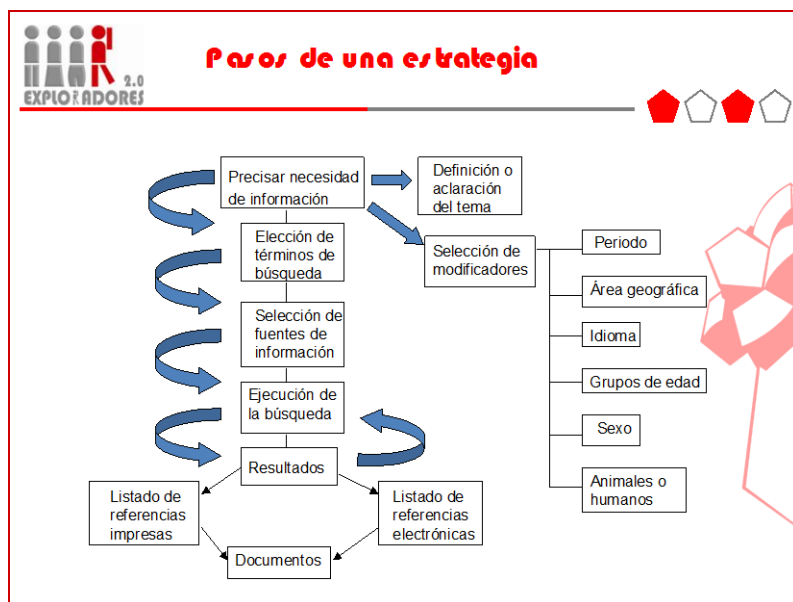
- a) Documental: localizar referencias bibliográficas y materiales de un tema específico en diversos soportes, p.e. revistas impresas, revistas electrónicas, libros, tesis, videos, etc.
- b) Factográfica: documentación de hechos concretos, p.e. participación de la mujer en la revolución mexicana
- c) Actualización: información de avances en un campo de especialización específico, p.e. nuevos fármacos para el tratamiento de la diabetes mellitus tipo II
- d) Restrospectiva: información anterior sobre un tema en particular, p.e. daños en viviendas de la ciudad de Colima por terremotos ocurridos entre 1900 y 2000
- e) Temática: investigación sobre un tema en particular, p.e. determinar que efectos produce en los trabajadores de las gasolineras la exposición prolongada al plomo.

### **3.1. Estrategias de búsqueda de información**

Podemos definir estrategia de búsqueda como:

- Modelos, planes o procedimientos que se emplean para la recuperación de información, descartando aquella no relevante o pertinente.
- Operaciones lógicas que se deben llevar a cabo en un proceso de búsqueda y el orden en que éstas deben realizarse para el logro de los mejores resultados.

En todos los casos, la estrategia que utilicemos irá de acuerdo con nuestra necesidad de información.



Los pasos de una estrategia de búsqueda son los siguientes:

### 1) Precisar la necesidad de información.

Lo primero que tenemos que definir es para qué necesitamos la información, que sería nuestro objetivo de búsqueda. No es lo mismo buscar información para una investigación o para una tesis, que para hacer un trabajo, una monografía, un ensayo, o resolver una duda. Es pertinente que tengamos claro de cuánto tiempo disponemos y qué tanto sabemos del tema, cual es nuestro grado de conocimiento y como se relaciona con otras temáticas.

En este punto debemos seleccionar algunos atributos conocidos como delimitadores o modificadores: qué es lo que nos interesa incluir y excluir en la investigación, vamos a trabajar sobre un periodo determinado, sobre un área geográfica determinada, con grupos étnicos determinados, con un sexo en particular, de una edad en particular, el estudio será con personas, plantas o animales, en qué otros idiomas voy a buscar información, etc.

### 2) Elección de los términos de búsqueda.

Una vez que identificamos nuestro tema, podemos escribir frases cortas de los conceptos más significativos, sus sinónimos y algunas palabras relacionadas con el objetivo de nuestra investigación.

Podemos traducir los conceptos con un diccionario especializado, o bien recuperarlos en otros idiomas utilizando tesauros, índices, listas de términos o listas de encabezamientos. Aquí **el uso de descriptores y palabras clave son fundamentales**.

### 3) Selección de fuentes de información.

Dependiendo del tipo de documentos que deseemos recuperar será el sitio a donde nos vamos a ir a buscar. La selección de las fuentes de información va de acuerdo con las que existen para el propósito de nuestra investigación y aquellas a las que podemos tener acceso: bibliotecas, hemerotecas, archivos, videotecas, fonotecas, etc.; se toma en cuenta además la gran diversidad de fuentes electrónicas que existen: bases de datos, catálogos en línea, revistas en línea, tesis en formato digital, etc.

### 4) Ejecución de la búsqueda.

Una vez seleccionadas las palabras clave o descriptores que vayamos a utilizar y también las fuentes de información en donde vamos a buscar, procederemos a hacer la búsqueda tomando en consideración las herramientas e instrucciones de uso del sitio.

Como regularmente siempre se utilizan varias palabras clave o descriptores, los **operadores lógicos** o **“booleanos”** nos van a ayudar a formar una “sentencia” o “secuencia” de búsqueda.

### 5) Resultados.

Debemos observar que tanto es lo que recuperamos, "mucho o poco" es decir, el resultado contiene una gran cantidad de documentos o solo un número muy reducido. A este fenómeno se le conoce como silencio o ruido documental.

- Silencio documental se refiere a los documentos almacenados en la base de datos pero que no han sido recuperados debido a que la estrategia de búsqueda ha sido demasiado específica o que las palabras clave utilizadas no son las adecuadas para definir la búsqueda.
- Ruido documental se refiere a los documentos recuperados por el sistema pero que no son relevantes. Esto suele ocurrir cuando la estrategia de búsqueda se ha definido demasiado genérica.


En ambos casos es necesario replantear la estrategia de búsqueda y los descriptores y operadores lógicos que estamos utilizando.

## 3.2. Uso de tesauros, descriptores y palabras clave

La definición o elección de los términos de búsqueda incluye el uso de palabras clave, descriptores y términos en uno o varios idiomas, que nos serán de utilidad en la consulta para la correcta recuperación de la información.

El uso de tesauros, listas de encabezamientos, vocabularios e índices es muy recomendable, ya que al manejar un lenguaje normalizado presentan un vocabulario previamente elaborado, y admiten un limitado número de modificaciones al momento de utilizarlas. Controlan además las ambigüedades propias del lenguaje, tanto para la representación como para la búsqueda y recuperación de la información (sinonimia, polisemia, homonimia), y las relaciones semánticas (equivalencia, jerárquica, asociativa).

El lenguaje o vocabulario controlado le proporciona al usuario un punto de búsqueda, reduce la posibilidad de que la búsqueda sea incompleta, permite el uso de términos técnicos adecuados para un área del conocimiento determinada, el uso de términos válidos, eliminar ambigüedades semánticas, seleccionar sinónimos, utilizar los términos correctos en otros idiomas, entre otros.



**Tesauros**

TESAURO *THESAURUS* → lista controlada y estructurada de términos

Vocabulario controlado y dinámico, compuesto por términos que tienen entre ellos relaciones semánticas y genéricas y que se aplica a un dominio particular del conocimiento (Norma ISO-2788), lo fundamental del tesauro es, por una parte, ser un lenguaje que utiliza herramientas conceptuales de control para eliminar las ambigüedades semánticas y por otra, ser aplicable a un área específica del conocimiento e incluso hacerlo desde un cierto ángulo en particular

Instituto de Estudios Documentales sobre Ciencia y Tecnología (IEDCyT) España <a href="http://thes.cindoc.csic.es/index_esp.php">http://thes.cindoc.csic.es/index_esp.php</a>	Biblioteca Dag Hammarskjöld UN <a href="http://unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/">http://unhq-appspub-01.un.org/LIB/DHLUNBISThesaurus.nsf/</a>
UNESCO Inglés, Español, Francés y Ruso <a href="http://databases.unesco.org/thesaurus/">http://databases.unesco.org/thesaurus/</a>	OCDE Macrotesauro Inglés, Francés, Español y Alemán <a href="http://168.96.200.17/ar/oced-macroth/">http://168.96.200.17/ar/oced-macroth/</a>
HASSET (Humanities and Social Science Electronic Thesaurus) <a href="http://www.data-archive.ac.uk/find/hasset-thesaurus">http://www.data-archive.ac.uk/find/hasset-thesaurus</a>	Tesauro Agrícola de la Biblioteca Nacional

NOTAS

Un **tesauro** (del latín *thesaurus* y del griego *thesauros* que significan “tesoro”) es una lista estructurada de conceptos, destinados a representar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental determinado, y a ayudar al usuario en la indización y en la recuperación de documentos.

La forma y la estructura de un tesauro se desarrollan con las disposiciones de las normas internacionales ISO 2788 (para tesauros monolingües) e ISO 5964 (para tesauros multilingües) (Moreiro G., J.A. et al. 2003; Marzal, M.A. et al. 2006).

Algunos tesauros:

- **Biblioteca Dag Hammarskjöld (Naciones Unidas)**  
<http://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf>
- **Australian Pictorial Thesaurus**  
<http://www.picturethesaurus.gov.au/about.html>
- **Tesauro de Arte y Arquitectura**  
<http://www.aatespanol.cl/taa/publico/portada.htm>
- **Instituto de Estudios Documentales sobre Ciencia y Tecnología** (varios tesauros)  
[http://thes.cindoc.csic.es/index\\_esp.php](http://thes.cindoc.csic.es/index_esp.php)

- **UNESCO**  
<http://databases.unesco.org/thessp/>
- **Tesoro Agrícola de la Biblioteca Nacional**  
<http://agclass.nal.usda.gov/>

**ACTIVIDAD 3.1. Localización de tesauros en su área de interés**

*Tiempo: 15 minutos*

1. Por medio del Google, busque 2 tesauros en su área de interés.
2. Anote como se llaman, su dirección electrónica (URL), contenido temático y en qué idiomas están los descriptores.

Los términos del tesoro incluyen **descriptores** (términos preferidos), utilizados para la indexación y la recuperación de información y **no descriptores** (términos no preferidos) los cuales nos brindan puntos de entrada adicionales para los descriptores por medio de una referencia de “USE” (use for UF) (Fig. 3.1.) (Lamarca Lapuente, M.J., 2007) (Universidad de Salamanca, España).

Los descriptores son palabras o expresiones del lenguaje común retenidas por el constructor del tesoro para designar los conceptos representativos del documento y las preguntas utilizadas por el indizador, suelen acompañarse de una nota de alcance o, menos usualmente, de una definición en los casos en que el mero registro del término puede provocar problemas de ambigüedad en su interpretación o cuando este tiene diferentes significados ya que es utilizado en diferentes disciplinas, **no forma parte del descriptor** aunque aparezca en el tesoro (Fig. 3.1.).

Las notas de alcance pueden ser explicativas, de definición, de aplicación o históricas.

<p><b>DESARROLLO ECONOMICO</b>                  02.04.00 – DESARROLLO  <b>Nota de alcance:</b> Para el incremento en la producción de bienes y servicios solamente  <b>Úsese</b> CRECIMIENTO ECONOMICO</p>	<p><b>DEVELOPPEMENT ECONOMIQUE</b>                  02.04.00 – DEVELOPPEMENT  <b>Note d'application:</b> Pour la simple croissance de la production  <b>Employer</b> CROISSANCE ECONOMIQUE</p>	<p><b>ECONOMIC DEVELOPMENT</b>                  02.04.00 – DEVELOPMENT  <b>Scope note:</b> For increased output of goods and services alone  <b>Use</b> ECONOMIC GROWTH</p>
--	--	---

**Figura 3.1** Ejemplo de descriptores tomados del tesoro de Naciones Unidas (UNBISNET)

A los descriptores se les conoce también como términos de indexación ya que son los términos con los que se indizará y recuperará la información en el tesoro.

Los no descriptores son sinónimos o cuasi-sinónimos de los descriptores o términos que designan en el lenguaje de uso conceptos afines a los que cubren los descriptores. Los no descriptores no pueden ser utilizados para la indexación del documento o de las preguntas, pero cada uno de ellos reenvía a uno o dos descriptores para representar los conceptos correspondientes.



En el siguiente ejemplo, “evasión de capitales” sería el descriptor y “fuga de capitales” sería el no descriptor:

**Evasión de capitales**  
SN Fuga de dinero hacia el exterior (nota de alcance)  
UF Fuga de capitales (útese para)

**Fuga de capitales** (no admitido)  
USE Evasión de capitales (admitido)

Las **palabras clave** o **keywords** son términos significativos y representativos del documento que las contiene, que el indizador extrae del texto, del título o del resumen de dicho documento, para ser usado como término. A diferencia de los descriptores, las palabras clave no se toman de un tesoro, sino que son términos designados por los autores del artículo para que su trabajo pueda ser recuperado (Figura 3.2).

**Palabras clave:** Hipermedia, Patrimonio cultural, Digitalización, Cultura digital, Sociedad del Conocimiento.  
**Key words:** Hypermedia, Cultural Heritage, Digitalization, Digital Culture, Knowledge Society.

Figura 3.2 Ejemplo de palabras clave tomadas de la revista Zer (ISSN 1137-1102)

### ACTIVIDAD 3.2. Uso de tesauros para búsqueda de descriptores

*Tiempo: 15 minutos*

1. De los tesauros que localizó y de los que se vieron en la sesión, utilice 2
2. Busque descriptores de su tema de interés tanto en español como en inglés y anótelos.

Estos descriptores serán utilizados posteriormente para hacer las búsquedas en el Google, en el Google Académico y en las fuentes de información.

### TALLER EXTRA CLASE 3.1. Localización de tesauros para búsqueda de descriptores

*Tiempo: 30 minutos*

1. Por medio del Google, busque 5 tesauros de su área de interés, aparte de los 2 que encontró en clase. Anote como se llaman, su dirección electrónica (URL), contenido temático y en qué idiomas están los descriptores.
2. Busque descriptores de su área de interés y escríbalos en español e inglés.
3. Los resultados se socializarán con los demás participantes en la sesión presencial y se entregará la lista con la descripción de los tesauros y los descriptores localizados.

Ejemplo:

**EuroVoc** <http://eurovoc.europa.eu/>

Es un tesoro multilingüe y multidisciplinario de la Unión Europea. Abarca términos de los ámbitos de actividad de la Unión Europea, especialmente los relacionados con las labores parlamentarias. Se puede consultar en 22 idiomas.

Descriptores:

derecho criminal – criminal law  
expulsión, deportación – deportation

terrorismo – terrorism  
protección a testigos – witness protection

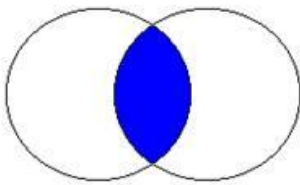
### 3.3. Operadores Booleanos

Los operadores booleanos se conocen también como operadores lógicos, son instrumentos del modelo booleano (modelo clásico de recuperación de la información basado en la teoría de los conjuntos).

Los más importantes son AND que tiene la función de intersección, OR que tiene la función de unión y NOT que tiene la función de exclusión.

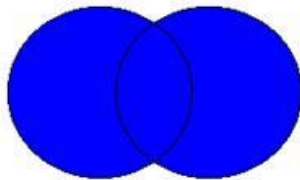
En los buscadores de Internet, cada herramienta de búsqueda tiene distintos operadores, aunque existen unos cuantos que tratan de ser universales, como son los tres anteriores.

#### Operador AND



Es la intersección de los dos conjuntos de búsqueda. Apareciendo en el resultado únicamente los elementos que aparecen en los dos conjuntos. Es un operador de reducción.

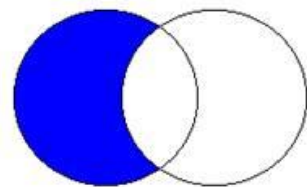
#### Operador OR



Es un operador para la unión de conjuntos. Se utiliza para ampliar el enfoque de la búsqueda e incrementa por lo general, el número de documentos a recuperar.

Al utilizar OR se indica al buscador que se desea recuperar documentos donde aparezca uno, otro o al menos uno de los argumentos indicados.. El operador OR es especialmente útil para indicar asociaciones de palabras o sinónimos en la estrategia de búsqueda .

#### Operador NOT



Es el operador de exclusión de conjuntos. El resultado de este operador son los registros que contienen los elementos del primer conjunto y que no son los del segundo. AND NOT o NOT es muy útil para minimizar los problemas ocasionados por la polisemia.

Generalmente, se utiliza después de explorar de forma preliminar los resultados devueltos por el buscador, y así refinar de la búsqueda los resultados irrelevantes.

Hay que tener precaución con su uso, ya que puede representar la pérdida de documentos relevantes; y revisar la ayuda del buscador para comprobar si utiliza la expresión AND NOT o la palabra NOT o sus equivalentes en español. (y no, no).

### ACTIVIDAD 3.3. Uso de los operadores booleanos para búsqueda de información

*Tiempo: 15 minutos*

1. Abra el Google y el Google Académico
2. Con los descriptores de su tema, haga búsquedas de información utilizando los operadores AND, OR y NOT.

## 3.4 Selección de fuentes de información

La selección de las fuentes de información va de acuerdo con las que existen para el propósito de nuestra investigación y aquellas a las que podemos tener acceso: bibliotecas, hemerotecas, archivos, videotecas, fonotecas, etc.; se toma en cuenta además la gran diversidad de fuentes electrónicas que existen: buscadores, bases de datos, catálogos en línea, revistas en línea, tesis en formato digital, bibliotecas digitales, bibliotecas virtuales y repositorios.

Para propósito de este módulo solo se utilizarán el Google y el Google Académico. Las demás fuentes de información se ven a detalle en el módulo 4.

## 3.5 Operadores avanzados de Google

Existe una serie de operadores o comandos avanzados de búsqueda que pueden ser utilizados en Google o en Google Académico pero que no están enlistados en la ayuda del buscador. El propósito es que se conozcan para facilitar la recuperación de información y para que el uso del buscador sea más eficiente.

A nuestro requerimiento de información (query) que pueden ser una o más palabras, una frase o números se le conoce como “términos de búsqueda” y los resultados pueden variar según como los escribamos o en el orden en que lo hagamos.

Las comillas nos sirven para buscar una frase completa, los paréntesis para hacer búsquedas con varios operadores y el asterisco nos sirve para “truncar” palabras. Podemos buscar información en un tipo determinado de sitio (gob, com, org, net) o de un tipo de documento en particular (pdf, doc, ppt).

### ACTIVIDAD 3.4. Uso de los operadores avanzados de Google

*Tiempo: 15 minutos*

1. Abra el Google
2. El instructor del curso proporcionará los operadores con los que se harán los ejercicios. Se utilizarán los descriptores de su tema.

## 3.6 Uso de operadores, descriptores y palabras clave en las búsquedas

Estos son algunos tips para la búsqueda de información:

1. Utilice palabras que probablemente estén en las páginas que desea encontrar.
2. Sea específico, haga la búsqueda con los términos adecuados y no una búsqueda general.
3. Sea breve, para mejores resultados utilice términos precisos.
4. Asegúrese de que está escribiendo correctamente las palabras.

### ACTIVIDAD 3.5. Práctica general

*Tiempo: 15 minutos*

1. Abra el Google y el Google Académico
2. El instructor del curso dará las instrucciones y los operadores con los que se harán las búsquedas combinando los descriptores, operadores lógicos y comandos avanzados.

En esta liga encontrará otros operadores avanzados de Google:

[http://www.googleguide.com/advanced\\_operators.html](http://www.googleguide.com/advanced_operators.html)

### TALLER EXTRA CLASE 3.2. Uso de operadores avanzados de Google

*Tiempo: 30 minutos*

1. Abra el Google y el Google Académico.
2. Abra la página [http://www.googleguide.com/advanced\\_operators.html](http://www.googleguide.com/advanced_operators.html)
3. De la lista de comandos de la página, escoja 3 operadores diferentes a los que se vieron en clase
4. Realice las búsquedas con las instrucciones del comando, utilizando sus descriptores o las palabras clave que desee.

Los resultados se socializarán con los demás participantes en la sesión presencial y se entregará la lista con los operadores y descriptores utilizados y el número de resultados obtenidos con cada uno de ellos.

## Referencias:

**Cornella, A. (2000).** Cómo sobrevivir a la infoxicación [En línea]. Informática Educativa, Venezuela Recuperado el 01-02-2011 en:  
<http://www.informaticaeducativa.com.ve/descargable/infoxicacion.pdf>

**Lamarca Lapuente, M. J. (2007).** Hipertexto, el nuevo concepto de documento en la cultura de la imagen. Tesis doctoral, Universidad Computense de Madrid. Consultado el 11-02-11 en  
<http://www.hipertexto.info/documentos/tesauros.htm>

**Marzal, M. A. et al. (2006).** Mapas conceptuales y presentación gráfica del tesoro: aplicación a las bibliotecas educativas. Proc. of the Second Int. Conference on Concept Mapping, A. J. Cañas, J. D. Novak, Eds., San José, Costa Rica Recuperado el 01-02-2011 en:  
<http://cmc.ihmc.us/cmc2006Papers/cmc2006-p70.pdf>

**Moreiro González, J. A. et. al. (2003).** Utilización de estructuras verbales en la identificación de relaciones y descriptores en tesauros II. Proyecto financiado por la CICYT (Comisión Interministerial de Ciencia y Tecnología) de España. Consultado el 11-02-11 en  
<http://www.bibliociencias.cu/gsd/collect/eventos/index/assoc/HASH016c/759f01df.dir/doc.pdf>

**Universidad de Salamanca.** Tesauros: concepto, elaboración y mantenimiento [documento electrónico]. Salamanca, España. Consultado el 11-02-11 en:  
<http://web.usal.es/~alar/Bibweb/Temario/Tesauro.PDF>