



enero 2009

www.bibliopos.es

Consulta y explotación de Bases de Datos científicos

1. INTRODUCCIÓN. BASE DE DATOS DOCUMENTAL BIBLIOGRÁFICA.

- Conjunto de información estructurada en registros y estos en campos.

1.1. ESTRUCTURA DE LA INFORMACIÓN. REGISTROS Y CAMPOS.

1.2. CLASIFICACIÓN DE LAS BASES DE DATOS SEGÚN EL CONTENIDO DE LOS REGISTROS.

- Bases de Datos factuales.
- Directorios.
- Bases de Datos documentales.

1.3. BASES DE DATOS DOCUMENTALES. TIPOS:

- A texto completo.
- De imágenes.
- De datos referenciales.

2. BASES DE DATOS CIENTÍFICAS. TIPOS.

- SEGÚN EL ORGANISMO PRODUCTOR.
 - Universidades y organismos e instituciones públicas:
 - Acceso público.
 - Acceso interno.
 - Empresas privadas:
 - De uso interno.
 - Comerciales.
- SEGÚN EL MODO DE ACCESO.
 - De acceso local.
 - En CD-ROM.
 - En línea.
- SEGÚN SU COBERTURA TEMÁTICA Y EL DESTINATARIO.
 - Multidisciplinares.
 - Especializadas.
- SEGÚN SU COBERTURA DOCUMENTAL.
 - Centradas en un único tipo de documento.

- Reúnen varios tipos de documentos.
- SEÚN SU ANÁLISIS DOCUMENTAL.
 - De sumarios o sin análisis de contenido.
 - Catálogos de bibliotecas.
 - Con análisis documental más completo:
 - NTIS. National Technical Information Service.
 - Scisearch. Institute for Scientific Information (ISI).
 - CA SEARCH. Chemical Abstracts.
 - PASCAL.

3. ACCESO A LA INFORMACIÓN EN BD'S DOCUMENTALES.

- Búsqueda directa.
 - Interrogación en texto libre.
 - Interrogación en campos individuales.
- Búsqueda a través de índices.
- Búsqueda jerarquizada.
- Búsqueda a través de códigos.

4. ESTRATEGIAS DE BÚSQUEDA EN BASES DE DATOS.

5. ACCESO AL DOCUMENTO ORIGINAL.

1. INTRODUCCIÓN. BASE DE DATOS DOCUMENTAL BIBLIOGRÁFICA.

Bajo la denominación genérica de bases de datos documentales se engloban diferentes modelos que pueden diferenciarse tanto por el contenido, o tipos de documentos que describen, como por la estructura de las mismas.

Pero existen otras diferencia marcadas, fundamentalmente, por los diferentes programas informáticos con los que se realiza el acceso a las mismas.

Una base de datos es un conjunto de información estructurada en **registros** y almacenada en un soporte electrónico legible desde un ordenador. Cada registro constituye una unidad autónoma de información que, a su vez, puede estar estructurada en diferentes **campos** o tipos de datos que se recogen en dicha base de datos. Por ejemplo, en un directorio de miembros de una asociación, un registro será la ficha completa de cada uno de los socios de la misma. Es decir, en cada registro se recogerán determinados datos (nombre, profesión, dirección, etc., en función de la temática u objeto de la base de datos), cada uno de los cuales constituye un campo.

1.1. ESTRUCTURA DE LA INFORMACIÓN, REGISTROS Y CAMPOS.

Las bases de datos se componen de diferentes registros con su correspondiente número de identificación. En una base de datos bibliográfica cada registro se corresponde con una referencia de un documento original, que puede ser una publicación independiente (monografía, tesis, informe, etc.) o una parte con autonomía de contenido dentro de otra

publicación (artículo incluido en una revista, una serie o una compilación; ponencia o comunicación en las actas de una reunión científica, etc.).

Es decir, registro es una unidad completa y autónoma de información. La información contenida en un registro bibliográfico puede estar, a su vez, estructurada en diferentes campos para facilitar su control sistemático y su recuperación individualizada. Por consiguiente, campo es una subdivisión del registro dedicada a un solo aspecto de la información contenida en el mismo. Por ejemplo: autor, título, etc.

Algunos campos reflejarán la descripción formal del documento y otros irán destinados a reflejar su contenido temático. Según el modelo de base de datos y el tipo de documento vaciado en la misma, se establece una estructura de campos determinada.

Por ejemplo, en una base de datos que incorpora referencias de artículos de revistas se encuentran campos como los siguientes:

- Autor o autores del documento original.
- Título del artículo.
- Título de la revista en la que aparece.
- Año de publicación.
- Datos fuente: volumen, número y páginas.
- Idioma en el que está escrito el documento.
- Clasificación: que describe el tema central tratado en el documento.
- Descriptores: términos que describen conceptos teóricos tratados de forma especialmente relevante en el artículo.
- Resumen: texto breve que sintetiza el contenido del documento.

También pueden existir otros campos para la descripción formal del documento:

- Afiliación institucional de los autores.
- ISSN (International Serial Standard Number) de la revista.
- País de edición de la revista.
- Editorial de publicación de la misma.
- Biblioteca o fondo documental donde se puede localizar el documento.

Un valor añadido de las bases de datos bibliográficas puede ser el enlace directo con la copia electrónica del documento original, que el usuario puede decidir imprimir o realizar una copia en su ordenador personal.

1.2. CLASIFICACIÓN DE LAS BASES DE DATOS SEGÚN EL CONTENIDO DE LOS REGISTROS.

Según el contenido de los registros, las bases de datos pueden ser:

- **Bases de datos factuales:** cuando recogen informaciones muy concretas, generalmente de carácter numérico: datos estadísticos, resultados de encuestas, etc.

- **Directorios:** que recogen datos sobre personas o instituciones especializadas en una actividad o materia concreta. Hay directorios de investigadores, de centros de investigación, de empresas, de editoriales, etc.
- **Bases de datos documentales.** En este último grupo, cada registro se corresponde con un documento, sea este del tipo que sea: una publicación impresa; un documento audiovisual, gráfico o sonoro; un documento de archivo; un documento electrónico, etc.

1.3. BASES DE DATOS DOCUMENTALES. TIPOS

Una base de datos se crea y mantiene de forma continuada con el objetivo de resolver necesidades de información concretas de un colectivo, una empresa o el conjunto de la sociedad. Dichos recursos electrónicos pueden ser consultados directamente en formato electrónico o ser utilizados para elaborar productos impresos, como bibliografías, directorios, informes, etc.

Los registros de las bases de datos documentales pueden incluir o no el contenido completo de los documentos que contienen, según lo cual se distinguen tres categorías:

- **Bases de datos de texto completo:** son las que están constituidas por los propios documentos en formato electrónico, mediante un vaciado completo de su texto. Además, pueden incorporar campos que contengan información fundamental que facilite su descripción y recuperación.
- **Bases de datos de imágenes:** están constituidas por referencias que permiten un enlace directo con la imagen del documento original, sea este un documento iconográfico (fotografías, imágenes de televisión, etc.) o un documento impreso digitalizado en formato de imagen.
- **Bases de datos referenciales:** sus registros no contienen el texto original, sino tan sólo la información fundamental para describir y permitir la localización de documentos impresos, sonoros, iconográficos, audiovisuales o electrónicos. En estos sistemas de información sólo se pueden obtener referencias sobre documentos que habrá que localizar posteriormente en otro servicio (archivo, biblioteca, fonoteca, fototeca, etc.) o solicitar a un servicio de suministro de documentos. Sin embargo, una base de datos referencial puede incluir campos que faciliten la localización del documento (bibliotecas, signaturas, direcciones de Internet, etc.) o incluso enlaces directos para obtener directamente el original a través de otro programa (tratamiento de texto, navegador en Internet, etc.).

Las bases de datos bibliográficas son, generalmente, bases de datos documentales referenciales, cuyos registros contienen referencias de documentos impresos o de texto.

2. BASES DE DATOS CIENTÍFICAS. TIPOS

Se pueden establecer diferentes tipologías para caracterizar los diferentes modelos de bases de datos documentales:

- Según el organismo productor.
- Según el modo de acceso.
- Según la cobertura temática y el destinatario.
- Según la cobertura documental.
- Según el modelo de tratamiento documental.

SEGÚN EL ORGANISMO PRODUCTOR

Las bibliotecas y centros de documentación de **universidades y organismos e instituciones públicas** de investigación elaboran gran cantidad de recursos de información:

- **Bases de datos de acceso público**, sean gratuitas o no.
- **Bases de datos de acceso interno**, con información de acceso restringido.

Los centros de documentación, bibliotecas y archivos de **empresas privadas** pueden elaborar distintos tipos de sistemas de información:

- **Bases de datos de uso interno**, para facilitar la circulación de información dentro de la empresa. Ocasionalmente, ofrecen dicho servicio al exterior: usuarios particulares o instituciones.
- **Bases de datos comerciales**, diseñadas específicamente para ser utilizadas por usuarios externos.
- **Bases de datos realizadas por cooperación en red**: se trata de sistemas de información cuya elaboración es compartida por diversas instituciones. Algunas bases de datos internacionales se elaboran a través de este sistema de trabajo, con diversos centros nacionales responsables de la información perteneciente a cada país.

SEGÚN EL MODO DE ACCESO

- **Bases de datos de acceso local**. Para consultarlas es necesario acudir al organismo productor, a su biblioteca o centro de documentación. Pueden ser consultables en monpuestos o en varios puestos de una red local.
- **Bases de datos en CD-ROM**. Pueden adquirirse mediante compra o suscripción, bien directamente por un particular o por una biblioteca o centro de documentación que permita su consulta a sus usuarios. En algunas instituciones se instalan diferentes CD-ROM en una red local para permitir su consulta desde cualquier ordenador conectado a la misma.
- **Bases de datos en línea**. Pueden ser consultables desde cualquier ordenador conectado a través de un servidor de Internet. La consulta puede ser libre (gratuita) o exigir la firma de un contrato.

Una misma base de datos puede tener acceso local y, además, una edición en CD-ROM y un sistema de acceso en línea. Sin embargo, puede haber diferencias en el contenido presente en

cada uno de los formatos o en el grado de actualización de la información. Por ejemplo, el productor de una base de datos puede ofrecer la conexión en línea a la base de datos completa con actualización diaria y, en cambio, editar un CD-ROM que tan sólo contenga los últimos cinco años de información y se actualice semestralmente.

SEGÚN SU COBERTURA TEMÁTICA Y EL DESTINATARIO

Aunque las bases de datos documentales pueden contener información sobre cualquier tema, las que nos interesan en este tema son las “Bases de datos científico-tecnológicos”, que contienen información destinada a los investigadores de cualquier ámbito científico o técnico. A su vez, este grupo puede dividirse en:

- **Bases de datos multidisciplinares**, que abarcan varias disciplinas científicas o técnicas.
- **Bases de datos especializadas**, que recopilan y analizan documentos referentes a una disciplina concreta: investigación química, biomédica, energética, medioambiental, agroalimentaria, etc.

SEGÚN SU COBERTURA DOCUMENTAL

- **Bases de datos centradas en un único tipo de documento.** Su objetivo es recopilar y permitir la localización de un tipo documental muy concreto. Hay bases especializadas en patentes, tesis doctorales, informes, artículos de revista, etc.

Como ejemplo de este modelo pueden citarse:

- La base de datos NTIS (National Technical Information Service), especializada en documentos técnicos oficiales de EEUU.
- Las bases de datos especializadas en tesis doctorales (Dissertation Abstracts), de carácter internacional.
- **Bases de datos que reúnen varios tipos de documentos.** Su objetivo es dar información sobre una disciplina, incorporando para ello diferentes tipologías documentales.

Como ejemplo de este modelo puede citarse la base de datos “Energy Science and Technology”, especializada en energía, que recoge referencias bibliográficas de artículos de revista, abstracts, congresos, monografías, informes, tesis, etc.

SEGÚN SUS ANÁLISIS DOCUMENTAL

- **Bases de datos de sumarios o sin análisis de contenido.** Se componen de referencias bibliográficas sencillas, en las cuales el productor se limita a grabar los datos de la propia fuente y no realiza ningún análisis del contenido. Incorporan solamente los datos descriptivos, fundamentales para localizar el documento, presentes en el sumario: autor, título y datos de la fuente. En ocasiones, puede contener el resumen editado con el propio documento original, generalmente elaborado por el propio autor.

Un ejemplo de este modelo es el “Current Contents” del Institute for Scientific Information.

También podrían incluirse dentro de este modelo los servicios de suscripción a revistas electrónicas establecidas por empresas de distribución de publicaciones periódicas.

- **Catálogos de bibliotecas.** Son bases de datos que responden estrictamente a los fondos contenidos en una biblioteca o red de bibliotecas (catálogos colectivos). Tienen una alta homogeneidad, gracias a la aplicación de normas internacionales de catalogación.
- **Bases de datos con análisis documental más completo.** Son sistemas de información que incorporan un mayor número de puntos de acceso para facilitar su localización. Cada registro bibliográfico incluye un resumen del contenido del documento original y/o un conjunto de términos representativos de los términos tratados en el mismo. Dichos términos constituyen entradas en un índice de materias (palabras clave o descriptores) y al proceso de asignación de los mismos a un documento se le denomina “indización”.

Dentro de este modelo se incluyen numerosas bases de datos especializadas en distintas disciplinas científicas y de gran importancia y prestigio internacional como:

- NTIS. National Technical Information Service. Cobertura informativa desde 1964 y actualización semanal. Consiste en resúmenes de actividades de investigación, desarrollo e ingeniería patrocinadas por el Gobierno de EE.UU. Contiene, además, análisis elaborados por agencias federales o entidades contratadas o subvencionadas y es el medio por el que se puede acceder a informes de agencias como la NASA, DOE, etc., hasta un número aproximado de 240 agencias que están disponibles para distribución pública. NTIS también da acceso a los resultados de I+D de países distintos de los EE.UU. Así, entre las organizaciones que contribuyen a NTIS están el Ministerio de Industria y Comercio Internacional de Japón, laboratorios administrados por el Departamento de Industria del Reino Unido, el Ministerio Federal de Ciencia y Tecnología o el CNRS francés.
- Scisearch. Institute for Scientific Information (ISI): es un índice internacional y multidisciplinario, de acceso a la literatura de ciencia, tecnología, biomedicina, y otras disciplinas relacionadas. Contiene todos los registros publicados en el Science Citation Index además de registros adicionales en ingeniería, ciencias físicas, agricultura, biología, ciencias ambientales, medicina clínica, y las ciencias de la vida. SciSearch indiza todos los artículos significativos (artículos, revisiones, resúmenes de reuniones, cartas, editoriales, reseñas de libros, correcciones, etc.) de más de 6,100 revistas internacionales científicas y técnicas.
- CA SEARCH. Chemical Abstracts: La base de datos CA SEARCH®: Chemical Abstracts® incluye más de 17 millones de citas de la literatura mundial sobre química y sus aplicaciones desde 1967 en adelante, con actualización semanal. CA SEARCH se corresponde con la información bibliográfica y los índices completos que se encuentran la versión impresa de Chemical Abstracts.

- **PASCAL**: producida por el Institut de l'Information Scientifique et la Técnica (INIST) del Consejo Nacional de Investigación francés (CNRS) y proporciona acceso a la literatura científica y técnica mundial e incluye aproximadamente 450,000 nuevas citas por año. Disponible en la forma legible por máquina desde 1973, PASCAL se corresponde con la publicación impresa de la Bibliographie Internationale (antes Boletín Signaletique).

4. ACCESO A LA INFORMACIÓN EN BASES DE DATOS DOCUMENTALES.

Un mismo sistema de información puede estar accesible desde distintos soportes: acceso vía telnet, acceso desde una página Web y edición en CD-ROM. Cada uno de estos formatos puede tener su propio sistema de interrogación y una forma de presentación o interfaz de usuario diferente. En el acceso tradicional en línea (telnet), la consulta puede realizarse a través de un formato de menús, en los que el usuario va seleccionando sucesivas opciones, o a través de complejos lenguajes de recuperación formados por comandos y operadores. Por el contrario, en la consulta en páginas Web y en la mayoría de los CD-ROM, el usuario debe completar formularios diseñados en entorno Windows.

En general, todos los sistemas de recuperación de información se presentan en forma de comandos, de menús o formularios y permiten realizar diferentes modalidades de búsqueda:

- **Búsqueda directa**: se teclean directamente una o varias palabras en el espacio reservado para ello por el sistema de interrogación en la base de datos. Dentro de esta modalidad pueden distinguirse, a su vez, dos tipos:
 - **Interrogación en texto libre**: el usuario realiza su consulta sin tener en cuenta la estructura de campos de los registros de la base de datos. Generalmente, el sistema busca el término o los términos solicitados en la totalidad de los campos de la base. En algunas bases de datos esta búsqueda se realiza solamente en determinados campos, aquellos que el administrador de la misma ha establecido.
 - **Interrogación en campos individuales**: el usuario realiza su consulta directamente sobre el campo o campos que ha seleccionado previamente (autor, revista, descriptores, etc.)
- **Búsqueda a través de índices**: en vez de teclear un término, el usuario visualiza un diccionario o índice alfabético de las entradas de todos los campos o de uno en concreto (índice de autores, de títulos, descriptores, etc.) y selecciona los más adecuados a su búsqueda. El proceso puede parecer más lento pero es más seguro, ya que la búsqueda se adapta mejor al contenido real de los registros de una base de datos.
- **Búsqueda jerarquizada**: la interrogación se realiza a través de una estructura jerárquica. Es decir, a partir de un concepto genérico se pueden localizar, no sólo los registros en los que aparece dicho término sino, además, todos aquellos en los que figure algún concepto más específico de su campo semántico. Esta opción de búsqueda sólo será posible en aquellos sistemas de información que cuenten con

campos codificados o que utilicen lenguajes estructurados jerárquicamente en la clasificación o en los descriptores.

- **Búsqueda a través de códigos:** en determinados campos de algunas bases de datos la interrogación no se realiza a través de un texto sino a través de códigos numéricos o alfanuméricos. Pueden ser campos codificados: la clasificación, el idioma, la tipología documental, etc.

5. ESTRATEGIAS DE BÚSQUEDA EN BASES DE DATOS

Para realizar una consulta eficaz en una base de datos se debe intentar encontrar la estrategia más adecuada en cada sistema de información. Para ello se recomienda seguir los siguientes pasos:

1. Seleccionar la base de datos más adecuada a la consulta. La elección tendrá en cuenta la cobertura temática, la homogeneidad del lenguaje utilizado y la tipología documental analizada por el sistema de información.
2. Familiarizarse con las características de la base de datos y del sistema de recuperación.
3. Seleccionar los conceptos de búsqueda que representen de forma precisa el tema de interés. Deben expresarse con todas sus posibles variantes (sinónimos, conceptos afines, etc.), aunque deben evitarse los términos innecesarios.
4. Determinar la estrategia de búsqueda: en que campos puede encontrarse de forma más adecuada la información buscada y como combinar los diferentes conceptos, con operadores booleanos o de adyacencia. Las herramientas como el truncado son muy útiles, pero siempre que se compruebe su alcance para evitar la selección de otros términos no deseados. Cuando sea necesario, se consultarán directamente los índices de los campos seleccionados y otros recursos (tesauros, listados, etc.)
5. Realizar la búsqueda y visualizar los resultados, valorando la adecuación de los mismos y la eficacia de la estrategia utilizada.
6. Si el resultado es poco satisfactorio o existen dudas sobre su idoneidad, se debe repetir la búsqueda ampliando o limitando la estrategia utilizada.
7. Seleccionar los registros adecuados a la necesidad de la información y realizar una petición de suministro de documentos o bien obtener un listado de referencias por impresora o en fichero.

Para utilizar eficazmente las posibilidades de recuperación en un sistema de información es recomendable adaptarse, lo más posible, a sus características particulares. Para ello es necesario estar informado previamente de su contenido y de cómo debe realizarse la búsqueda, es decir, debe tenerse en cuenta:

- Si el sistema permite la consulta de una única base de datos o de varias simultáneamente. En este caso, habrá que valorar si se trata o no de bases de datos de características similares.
- Cuál es su cobertura temática. En las bases de datos especializadas se pueden realizar búsquedas prescindiendo de aquellos conceptos que definen el propio ámbito temático del sistema.
- Si se trata de un catálogo o de una base de datos documental con descriptores y/o con resúmenes.
- En el caso de que haya descriptores o encabezamientos de materias, si estos proceden o no de un léxico controlado accesible.
- Cuál es la estructura de los campos y en cuales se realiza la búsqueda por defecto cuando no se selecciona uno en concreto.
- Cuando se teclean dos términos, se debe saber si el sistema interpreta que debe buscar registros en los que aparezcan ambos a la vez (intersección) o bien, cualquiera de los dos (suma booleana). A esto se le denomina operador por defecto. El resultado, obviamente, será distinto.

6. ACCESO AL DOCUMENTO ORIGINAL

En algunos sistemas de información, como en las bases de datos con información factual, las bases de datos con texto completo y los directorios, el usuario obtiene directamente los datos o los documentos que precisa. Sin embargo, las bases de datos referenciales sólo cubren un primer paso en la búsqueda de información, ya que el usuario tan sólo obtiene la noticia de la existencia de un documento.

Por ello, estos sistemas pueden tener asociado un servicio de acceso al documento original. A través de él, el usuario que ha recibido el resultado de una consulta bibliográfica puede solicitar directamente aquellos documentos que son de su interés.

El servicio de acceso al documento puede realizar las siguientes funciones:

- Préstamo interbibliotecario de monografías. Las grandes bibliotecas y las redes de bibliotecas universitarias han desarrollado los sistemas internacionales y nacionales de préstamo interbibliotecario.
- Envío de tesis, informes o documentos de archivo en microfilm o en microficha.
- Envío de artículos en fotocopia, a través del correo ordinario.
- Envío de artículos a través del correo electrónico, En muchos centros se está utilizando el programa ARIEL, para envío de copias de los artículos a través del correo electrónico. Este sistema consigue una alta calidad de las copias que supera

claramente al envío por fax y, al mismo tiempo, permite una gran rapidez en el suministro de documentos frente al correo ordinario.

En la actualidad, muchos documentos están en formato electrónico. Muchas bases de datos poseen, como valor añadido, el enlace directo con la copia electrónica del documento original, que el usuario puede decidir imprimir o realizar una copia en su propio ordenador personal.

www.bibliopos.es



Licencia [Creative Commons Reconocimiento-No comercial 3.0](https://creativecommons.org/licenses/by-nc/3.0/)