

TEMA 1.	Bases de Datos y Técnicas Documentales	5 horas
1.1	Introducción a la documentación automatizada	20 minutos
1.2	Bases de datos documentales	2 horas
1.2.1	Definición y contexto	
1.2.2	Bases de datos bibliográficas	
1.2.3	Bases de datos factuales	
1.2.4	Bases de datos textuales	
1.2.5	Bases de datos mixtas	
1.2.6	Bases de datos comerciales	
1.2.7	Protocolos de acceso	
1.3	Análisis documental	1 hora
1.3.1	Definición de documento	
1.3.2	Elementos de la cadena documental	
1.3.3	Indización manual	
1.3.4	Indización automática	
1.3.5	Clasificación, palabras clave y descriptores	
1.4	Lenguajes documentales	1 hora
1.4.1	Definición y contexto	40 minutos
1.4.2	Lenguajes de estructura jerárquica	
1.4.3	Lenguajes de estructura asociativa	
1.4.4	Tesaurus y su representación informática	

## CONTENIDOS

## Objetivos generales del tema

## OBJETIVOS

- Presentar al alumno la perspectiva de las bases de datos documentales, su tipología y normas de acceso
- Estudiar los esquemas de representación para cada uno de los tipos de bases de datos documentales
- Analizar los niveles de estructuración de las bases de datos documentales
- Estudiar el proceso del análisis documental para la indización de documentos
- Analizar las propuestas de indización manual y automática y su correspondencia con el nivel de estructuración de una base de datos documental
- Presentar los diferentes modos de indización, basados en clasificaciones temáticas, palabras clave y descriptores

- Introducir los lenguajes documentales como herramienta de control terminológico, y su correspondencia con los modos de indización
- Estudiar el lenguaje documental tesoro, proponiendo una representación informática de las relaciones de control entre descriptores

#### PLANTEAMIENTO Planteamiento general del tema

Este tema pretende aportar una visión genérica de las técnicas documentales, que constituya la base de la presentación de los sistemas de recuperación de información o documentación automatizada. En primer lugar, presenta la definición de una base de datos documental y el contexto de ubicación de este tipo de base de datos, comparándolo con los sistemas de gestión de bases de datos habituales. En esta comparativa, se establecen las diferencias con respecto a estructuración, lenguajes y formas de acceso a la información.

A continuación, definimos los distintos tipos de bases de datos documentales, haciendo especial hincapié en el nivel de estructuración de cada uno de ellos. Para cada tipo, presentamos un esquema de representación, donde se podrá comprobar que las bases de datos bibliográficas están altamente estructuradas, mientras que las bases de datos textuales tienen un nivel de estructuración mínimo. Es importante que el alumno sepa asociar el nivel de estructuración con la información a almacenar, decidiendo el tipo de base de datos más adecuada. Además, proporcionamos referencias de bancos de datos comerciales, así como productores y proveedores de bases de datos documentales. Esta primera parte del tema, finaliza con la presentación de los protocolos de acceso a los fondos de documentación automatiza. Dentro de estos protocolos, tenemos el acceso local vía CD-ROM, o el acceso en línea, destacando el estándar Z39.50 como protocolo de acceso a bibliotecas digitales en entornos cliente-servidor.

La segunda parte del tema se centra en el análisis documental como mecanismo de selección, almacenamiento y organización de ideas relevantes de un documento para su posterior localización. Para ello, definiremos el concepto de documento y los elementos tradicionales de la documentación, tales como catalogación, indización, clasificación y resumen. Dentro del proceso de indización, presentaremos, de forma breve, los mecanismos de indización manual y automática. En este punto del tema, el alumno debe identificar la correspondencia entre el nivel de estructuración de la base de datos documental y el mecanismo de indización a utilizar para la recuperación de documentos. Por último, proponemos tres modelos de indización clásicos que tendrán su correspondencia con los lenguajes documentales.

La última parte del tema se centra en los lenguajes documentales como herramienta de control terminológico para la formulación de una petición de información en una base de datos documental. Presentaremos los lenguajes documentales de estructura jerárquica, conocidos como los primeros mecanismos de indización, para pasar a los lenguajes asociativos, donde destaca el lenguaje

documental por excelencia, el tesoro. Para el tesoro, presentaremos, de forma breve, las relaciones que incluye, y una posible representación informática de las relaciones más importantes: relaciones semánticas, jerárquicas, similitud, parecido y definitorias.

### Bibliografía básica

BIBLIOGRAFÍA
--------------

[PBR02] R. Peña, R. Baeza-Yates and J. Rodríguez. *Gestión Digital de la Información. De Bits a Bibliotecas Digitales y la Web*. Ra-Ma, 2002.

*El capítulo 8, páginas 201–220* presenta una introducción al contexto de las bases de datos documentales, análisis documental y lenguajes documentales.

[BR99a] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley, 1999.

*El capítulo 7, páginas 165–173* estudia el procesamiento del análisis documental para la selección de los términos a incorporar en un lenguaje documental; además, presenta las características informáticas del tesoro como lenguaje de control de términos.

[Liz92] L. Lizasoain. *Bases de Datos en CD-ROM*. Paraninfo, 1992.

*El capítulo 3, páginas 37–46* presenta la definición y tipologías de las bases de datos documentales, así como los protocolos de acceso.

*El capítulo 4, páginas 51–57* ofrece referencias de bases de datos documentales comerciales distribuidas mediante CD-ROM y conexión en línea.

*El capítulo 7, páginas 113–123* presenta las cuestiones básicas del análisis documental y la introducción a los lenguajes documentales.

### Bibliografía complementaria

[Dew88] [TR90] [Man93] [She94] [Tra97]

### Otras referencias bibliográficas

REFERENCIAS
-------------

[SY73] [BE80] [Cod95] [Cod96] [GR96] [Bae96]