

# Bases de datos

Jaume Raventós Moret  
Dolors Costal Costa  
Carme Martín Escofet  
Ernest Abadal  
Lluís Codina

PID\_00198121

Material docente de la UOC


**Jaume Raventós Moret**

Ingeniero químico por el Instituto Químico de Sarriá. Ha desarrollado su carrera profesional como formador informático en empresas y centros de formación, y como consultor TIC diseñando y desarrollando aplicaciones de gestión y sistemas de base de datos. Colabora con los Estudios de Ciencias de la Información y la Comunicación de la UOC desde el año 1999 como consultor en la asignatura *Sistemas informáticos* y últimamente en la asignatura *Bases de datos*. Participa como profesor en varios programas formativos TIC en el Área de Promoción Económica del Ayuntamiento de Hospitalet de Llobregat.


**Dolors Costal Costa**

Doctora en Informática por la Universidad Politécnica de Cataluña. Profesora titular del Departamento de Ingeniería de Servicios y Sistemas de Información de la Universidad Politécnica de Cataluña, asignada a la Facultad de Informática de Barcelona.


**Carme Martín Escofet**

Doctora en Informática por la Universidad Politécnica de Cataluña. Profesora de la asignatura *Bases de datos* en la Facultad de Informática de Barcelona. También ha sido profesora en la Facultad de Matemáticas y Estadística y en la Escuela Universitaria Politécnica de Vilanova i la Geltrú y es personal docente colaboradora de los Estudios de Informática y Multimedia de la UOC.


**Ernest Abadal**

Doctor en Ciencias de la Información, licenciado en Filosofía y diplomado en Biblioteconomía y Documentación. Profesor titular de la Facultad de Biblioteconomía y Documentación de la Universidad de Barcelona. Su ámbito de docencia e investigación son las tecnologías aplicadas a la documentación y, de una manera especial, las publicaciones digitales. Ha escrito varios libros y numerosos artículos sobre estas cuestiones. Es director de la revista *BiD: textos universitaris de biblioteconomia i documentació*, miembro del consejo asesor de la revista *El profesional de la información* y de Hipertext.net y revisor de diferentes revistas científicas de su especialidad.


**Lluís Codina**

Profesor titular del Departamento de Comunicación de la Universidad Pompeu Fabra y director de la Unitat de Suport a la Qualitat i a la Innovació Docent (USQUID) de la Facultad de Comunicación. Imparte docencia en los Estudios de Periodismo y de Comunicación Audiovisual. Participa en másteres oficiales y en programas de doctorado del Departamento de Comunicación. Coordina el Grupo de Investigación en Documentación Digital y el de Comunicación Interactiva y forma parte del Grupo de Investigación en Periodismo, los dos grupos reconocidos de la UPF y el último por la Generalitat de Cataluña. Fundador y codirector del primer máster que se imparte íntegramente en la web sobre documentación digital, el anuario Hipertext.net y coeditor del weblog sobre documentación digital Servei d'Alerta. Participa en el máster interuniversitario de la Universidad de Barcelona y de la Universidad Pompeu Fabra sobre gestión de contenidos digitales.

La revisión de este material docente ha sido coordinada por la profesora:  
Mercè Vázquez García (2013)

Primera edición: febrero 2013

© Jaume Raventós Moret, Dolors Costal Costa, Carme Martín Escofet, Ernest Abadal, Lluís Codina

Todos los derechos reservados

© de esta edición, FUOC, 2013

Av. Tibidabo, 39-43, 08035 Barcelona

Realización editorial: Eureka Media, SL

Diseño: Manel Andreu

Depósito legal: B-589-2013



Los textos e imágenes publicados en esta obra están sujetos –excepto que se indique lo contrario– a una licencia de Reconocimiento-NoComercial-SinObraDerivada (BY-NC-ND) v.3.0 España de Creative Commons. Podéis copiarlos, distribuirlos y transmitirlos públicamente siempre que citéis el autor y la fuente (FUOC. Fundació per a la Universitat Oberta de Catalunya), no hagáis de ellos un uso comercial y ni obra derivada. La licencia completa se puede consultar en <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.es>

## Introducción

La asignatura *Bases de datos* tiene como objetivo introducirnos en el mundo de las bases de datos, el mecanismo que habitualmente se utiliza para almacenar datos de manera permanente y gestionarlas convenientemente. Una vez definidos algunos conceptos básicos, repasaremos la evolución que ha seguido la gestión de datos y describiremos las estructuras de datos más utilizadas en los sistemas de información.

A partir de los problemas que presentan los ficheros de datos tradicionales, justificaremos el uso de las bases de datos, de las cuales expondremos sus objetivos, características, ventajas e inconvenientes. Asimismo, identificaremos los elementos que componen un sistema de base de datos, poniendo especial atención a los diferentes tipos de usuario.

Explicaremos el concepto de *modelo de datos* el cual permite conseguir el proceso de abstracción que conduce del mundo real al mundo de los datos, distinguiendo entre el esquema y el estado de la base de datos, e introduciremos los diferentes tipos de abstracción que facilitan la representación de los datos en el diseño de bases de datos. Mostraremos cómo se organizan los datos de forma persistente en los apoyos de almacenamiento secundario y veremos cómo el sistema de gestión de bases de datos (SGBD) accede al mismo siguiendo el esquema general del proceso de ejecución de una consulta.

Para entender las características de los SGBD, analizaremos su arquitectura desde diferentes puntos de vista y nos centraremos en la arquitectura de tres niveles. El análisis de las funciones que tiene que hacer el SGBD nos ayudará a determinar los componentes que tiene que tener para cumplirlas: *lenguajes* de base de datos, que permiten definirlos y manejarlos; *interfaces*, que permiten a cada categoría de usuario especificar las solicitudes; y el *núcleo*, formado por diferentes módulos de software que realizan tareas específicas utilidades de valor añadido y aplicaciones integradas, que permiten simplificar las tareas de los usuarios; y el diccionario de datos, que contiene información sobre los datos que los módulos del SGBD necesitan consultar.

Con relación al análisis de las bases de datos, aunque hay otros enfoques, nos centraremos en las bases de datos relacionales, ya que actualmente el modelo relacional es el más utilizado. Concretamente de este modelo estudiaremos los fundamentos teóricos y algunos lenguajes que se utilizan para trabajarlos: el álgebra relacional y el lenguaje estándar SQL.

En la última parte del material profundizaremos en el estudio de las bases de datos documentales a partir de las cuatro operaciones que se pueden hacer: producción

y administración, en las que nos centraremos en el estudio de la estructura y las características de los programas informáticos (SGBDD) que permiten crear y explotar bases de datos documentales; distribución, en la que introduciremos el concepto de *interfaz de consulta* de una base de datos documental y daremos indicaciones para elaborarla; diseño, en el que nos centraremos en la metodología para la creación de bases de datos documentales; y evaluación, en la que describiremos los indicadores fundamentales para evaluar de una base de datos documental.

Este tipo de bases de datos se caracterizan por facilitar el acceso a documentos de tipo cognitivo (libros, artículos de revista, informes de todo tipo, etc.). Contienen grandes cantidades de texto que se almacena de manera poco estructurada, sin utilizar las tablas características del modelo relacional, y disponen de índices analíticos y amplias funcionalidades de busca para permitir la recuperación del contenido. No obstante, su uso y extensión son menores que el de las bases de datos relacionales.

## Objetivos

Los materiales didácticos incluidos en esta asignatura se orientan a conseguir que el estudiante alcance los objetivos siguientes:

1. Conocer los problemas que representa la gestión los datos.
2. Identificar los objetivos y las características de las bases de datos.
3. Distinguir los elementos que estructuran un sistema de base de datos.
4. Saber utilizar la abstracción para modelar bases de datos.
5. Conocer los medios de almacenamiento de bases de datos.
6. Comprender la arquitectura, el funcionamiento y los componentes de los sistemas de gestión de bases de datos.
7. Conocer los fundamentos del modelo de datos relacional.
8. Saber consultar una base de datos mediante las operaciones del álgebra relacional.
9. Saber utilizar las sentencias del lenguaje SQL para definir y manipular bases de datos.
10. Conocer la estructura y el funcionamiento de los programas informáticos (SGBDD) que permiten la creación de bases de datos documentales.
11. Conocer la estructura y el funcionamiento de los programas informáticos que permiten la distribución de bases de datos en la web.
12. Determinar los recursos tecnológicos necesarios para la creación y distribución de bases de datos documentales.
13. Conocer los elementos básicos de las interfaces web de consulta de bases de datos.
14. Conocer y aplicar la metodología para el diseño y la creación de bases de datos documentales.
15. Analizar y evaluar bases de datos a partir de indicadores.

## Contenidos

### Módulo didáctico 1

#### **Sistemas de base de datos**

Jaume Raventós Moret

1. Algunos conceptos básicos
2. Evolución de la gestión de datos
3. Estructuras de datos
4. Problemas de la gestión de ficheros de datos
5. Objetivos y características de las bases de datos
6. Ventajas e inconvenientes de las bases de datos
7. Elementos de un sistema de base de datos
8. Usuarios de las bases de datos
9. Modelo, esquema y estado de la base de datos
10. Tipos de abstracción en el diseño de bases de datos
11. Arquitectura de los SGBD
12. Estructura global de un sistema de base de datos
13. Almacenamiento de bases de datos
14. Acceso del SGBD a los datos
15. Funciones y componentes del SGBD

### Módulo didáctico 2

#### **El modelo relacional y el álgebra relacional**

Dolors Costal Costa

1. Introducción al modelo relacional
2. Estructura de los datos
3. Operaciones del modelo relacional
4. Reglas de integridad
5. El álgebra relacional

### Módulo didáctico 3

#### **El lenguaje SQL**

Carme Martín Escofet

1. Sentencias de definición de datos
2. Sentencias de manipulación de datos
3. Sentencias de control
4. Sublenguajes especializados

### Módulo didáctico 4

#### **Sistemas de gestión documental y bases de datos documentales**

Ernest Abadal, Lluís Codina

1. Producción y administración de bases de datos
2. Distribución de bases de datos
3. Metodología para la creación de bases de datos documentales
4. Evaluación de bases de datos

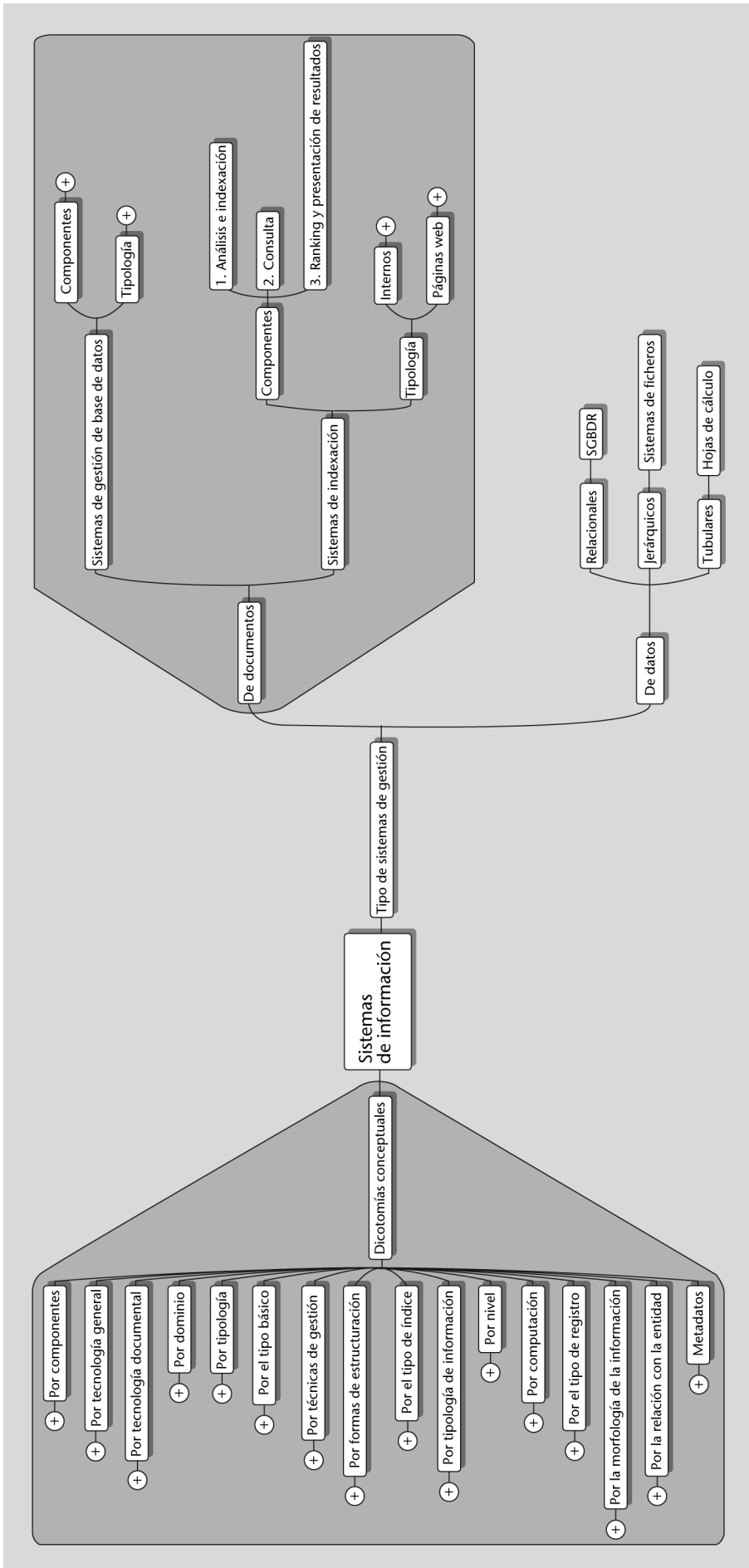
## Glosario

### Nomenclatura, definición y ejemplos de sistemas de información y de bases de datos.

Concepto	Acrónimo	Elementos	Definición	Otras denominaciones	Ejemplos
Sistema de información		contenido (datos e información) software recursos materiales (hardware o archivadores) personas	Sistema de elementos relacionados entre sí que procesan datos e información para el cumplimiento de determinados objetivos de una organización. Las actividades (entrada, almacenamiento, procesamiento y salida de información) no necesariamente deben estar informatizadas.	<i>Information system</i>	<b>Sistema de información manual:</b> sistema de archivo de libros en una biblioteca, sus tareas de organización y las personas que lo gestionan o lo usan. <b>Sistema de información informatizado:</b> el mismo, pero soportado por un ordenador con el programa correspondiente.
Sistema informático		software hardware personas	Sistema formado por el conjunto de los componentes de hardware y software que permite el proceso de datos. Puede formar parte de un sistema de información, pero no necesariamente.	<i>Computer system</i>	Un ordenador con sus programas, dispositivos periféricos y la persona que los maneja.
Sistema de base de datos	<b>SBD</b>	contenido software hardware personas	Sistema de información basado en una base de datos que incluye el software del sistema gestor de base de datos.	<i>Database system</i>	BD científicas o académicas (Scopus, Web of Knowledge, ACM Digital Library), BD de tesis doctorales, BD periodísticas, portales de revistas, hemerotecas y repositorios digitales.
Base de datos	<b>BD</b>	contenido (datos)	Conjunto estructurado de datos que representan entidades del mundo real de interés para un proyecto o una organización. Eventualmente, entre estos conjuntos de datos se pueden definir interrelaciones. El objetivo es organizar la información de manera que se pueda acceder a ella y administrar los datos con la máxima facilidad.	<i>Database</i>	Gestión informatizada (GI) de reparación de vehículos (clientes, coches, reparaciones, piezas), GI de actividades deportivas (alumnos, cursos, participantes, equipos, actividades, calendario), GI de préstamo bibliotecario.
Base de datos relacional	<b>BDR</b>	contenido (datos)	Conjunto de datos estructurado de acuerdo con el modelo relacional.	<i>Relational database</i>	BD de alquiler de coches, BD de cuentas bancarias, BD de ventas, BD de gestión hotelera (clientes, habitaciones, calendario).
Base de datos documental	<b>BDD</b>	contenido (documentos o referencias)	Conjunto de datos y/o documentos estructurado de acuerdo con el modelo documental.	<i>Documental database</i>	BD de prensa ( <i>El País</i> ), BD de recursos de Internet (Intute), BD de fotografías (AGE Fotostock), BD de recursos y artículos de revistas de educación (ERIC).
Sistema de recuperación	<b>SR</b>	software	Sistema que facilita la recuperación de datos o información (documentos)	<i>Retrieval system</i>	

Concepto	Acrónimo	Elementos	Definición	Otras denominaciones	Ejemplos
Sistema de recuperación de información	<b>SRI</b>	software	Sistema que permite la recuperación de información de un fondo documental por medio de una interfaz de consulta y un lenguaje de interrogación, con operadores que permiten optimizar la búsqueda.	<i>Information retrieval system</i>	Motor de búsqueda (Google, Yahoo), BD referencial (LISA).
Sistema de gestión de bases de datos	<b>SGBD</b>	software	Sistema destinado a la creación, la gestión, el mantenimiento y la explotación de bases de datos.	<i>Database management system</i>	Los ejemplos de SGBDR y SGD.
Sistema de gestión de bases de datos relacionales	SGBDR	software	Sistema de gestión de bases de datos que soporta el modelo relacional y permite operaciones basadas en este modelo.	<i>Relational database management system</i>	<b>Software propietario:</b> dBase, Access, Oracle, Informix, SQL Server, DB2 <b>Software libre (código abierto):</b> PostgreSQL, Sybase, MySQL
Sistema de gestión documental	SGD	software	Sistema de información que incorpora y gestiona documentos a los que proporciona acceso en el transcurso del tiempo.	<i>Document retrieval system</i> <i>Text retrieval system</i>	Los ejemplos de SGBD, de SGB y de sistemas de indexación.
Sistema de gestión de bases de datos documentales	<b>SGBDD</b>	software	Sistema de gestión de bases de datos que soporta el modelo documental	<i>documental database management system</i>	CDS/ISIS, Inmagic, Knosys, FileMaker
Sistema de gestión bibliográfica	<b>SGB</b>	software	Sistema de gestión documental especializado en el almacenamiento y recuperación de referencias bibliográficas, y la generación de bibliografías con diferentes formatos.	<i>Bibliographic management system</i>	<b>Sistema de escritorio:</b> Reference Manager, ProCite <b>Sistema en línea:</b> RefWorks, EndNoteZotero, Connotea, CiteUlike, Mendeley
Sistema de indexación		software	Sistema de gestión documental que genera índices analíticos (ficheros invertidos) a partir del contenido de documentos textuales.	Motor de búsqueda <i>Text retrieval software</i> <i>Full-text retrieval system</i> <i>Text information management system</i>	AskSam, Autonomy, Google Custom Search, Greenstone





Fuente: para ver el gráfico original y desplegar todo el contenido véase <http://tinyurl.com/diagramasgd>.

## Bibliografía

**Abadal, E.; Codina, L.** (2005). *Bases de datos documentales: características, funciones y método*. Madrid: Síntesis.

**Beaulieu, A.** (2010). *Aprende SQL* (2.<sup>a</sup> ed.). Madrid: Anaya Multimedia.

**Celma, M.; Casamayor, J. C.; Mota, L.** (2003). *Bases de datos relacionales*. Madrid: Pearson/Prentice-Hall.

**Codina, L.** (1998). "Metodología de análisis de sistemas de información y diseño de bases de datos documentales: aspectos lógicos y funcionales". *En: Anuari SOCADI de documentació i informació* (pág. 195-209). Barcelona: SOCADI.

**Connolly, T. M.; Begg, C. E.** (2005). *Sistemas de bases de datos: un enfoque práctico para diseño, implementación y gestión* (4.<sup>a</sup> ed.). Madrid: Pearson.

**Cuadra, D. y otros** (2008). *Desarrollo de bases de datos: Casos prácticos desde el análisis a la implementación*. Madrid: Ra-Ma.

**Date, C. J.** (2001). *Introducción a los sistemas de bases de datos* (7.<sup>a</sup> ed.). Madrid: Prentice-Hall.

**Date, C. J.; Darwen, H.** (1997). *A guide to the SQL Standard* (4.<sup>a</sup> ed.). Reading (Massachussets): Addison-Wesley.

**Elmasri, R.; Navathe, S. B.** (2007). *Fundamentos de sistemas de bases de datos* (5.<sup>a</sup> ed.). Madrid [etc.]: Pearson Addison Wesley..

**García-Molina, H.; Ullman, J. D.; Widom, J.** (2002). *Database Systems: The Complete Book*. Upper Saddle River (Nueva Jersey): Prentice-Hall.

**Informix Software, Inc.** (1996). *Informix Guide to SQL Syntax Version 7.2* (vol. 1 y 2). Menlo Park (California): Informix Press.

**Melton, J.; Simon, A. R.** (2002). *SQL: 1999. Understanding Relational Language Components*. San Francisco: Morgan Kaufmann.

**Peña, R.** (2002). *Gestión digital de la información: de bits a bibliotecas digitales y la web*. Madrid: Ra-Ma.

**Piattini, M. G. y otros** (2006). *Tecnología y diseño de bases de datos* (1.<sup>a</sup> ed.). Madrid: Ra-Ma.

**Ramakrishnan, R.; Gehrke, J.** (2003). *Database management systems* (3.<sup>a</sup> ed.). Boston [etc.]: McGraw-Hill.

**Raya, F.** (1987). *Database design for information retrieval: a conceptual approach*. Nueva York [etc.]: John Wiley & Sons.

**Rodríguez Yunta, L.** (1998). "Evaluación e indicadores de calidad en bases de datos". *Revista española de documentación científica* (vol. 21, núm. 1, pág. 9-23). Madrid: Consejo Superior de Investigaciones Científicas: Instituto de Información y Documentación en Ciencia y Tecnología

**Silberschatz, A.; Korth, H. F.; Sudarshan, S.** (1998). *Fundamentos de bases de datos* (5.<sup>a</sup> ed.). Madrid [etc.]: McGraw-Hill.

**Villanueva, E.** (1996, enero-junio). "Bases de datos y bibliotecología: cómo deshacer la innecesaria incomunicación". A: *Investigación bibliotecológica* (vol. 10, núm. 20, pág. 27-32). México, DF: Centro Universitario de Investigaciones Bibliotecológicas, UNAM.

**Willitts, J.** (1992). *Database design and construction: an open learning course for students and information managers*. Londres: Library Association.

