

Bases de dades

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

PID_00198119

Material docent de la UOC

**Jaume Raventós Moret**

Enginyer químic per l'Institut Químic de Sarrià. Ha desenvolupat la seva carrera professional com a formador informàtic en empreses i centres de formació i com a consultor TIC dissenyant i desenvolupant aplicacions de gestió i sistemes de base de dades. Col·labora amb els Estudis de Ciències de la Informació i la Comunicació de la UOC des de l'any 1999 com a consultor en l'assignatura *Sistemes informàtics* i darrerament en l'assignatura *Bases de dades*. Participa com a professor en diversos programes formatius TIC en l'Àrea de Promoció Econòmica de l'Ajuntament de l'Hospitalet de Llobregat.

**Dolors Costal Costa**

Doctora en Informàtica per la Universitat Politècnica de Catalunya. Professora titular del Departament d'Enginyeria de Serveis i Sistemes d'Informació de la Universitat Politècnica de Catalunya, adscrita a la Facultat d'Informàtica de Barcelona.

**Carme Martín Escofet**

Doctora en Informàtica per la Universitat Politècnica de Catalunya. Professora de l'assignatura Bases de dades a la Facultat d'Informàtica de Barcelona. També ha estat professora a la Facultat de Matemàtiques i Estadística i a l'Escola Universitària Politècnica de Vilanova i la Geltrú i és personal docent col·laboradora dels Estudis d'Informàtica i Multimèdia de la UOC.

**Ernest Abadal**

Doctor en Ciències de la Informació, llicenciat en Filosofia i diplomad en Biblioteconomia i Documentació. Professor titular de la Facultat de Biblioteconomia i Documentació de la Universitat de Barcelona. El seu àmbit de docència i recerca són les tecnologies aplicades a la documentació i, d'una manera especial, les publicacions digitals. Ha escrit diversos llibres i nombrosos articles sobre aquestes *qüestions*. És director de la revista *BiD: textos universitaris de biblioteconomia i documentació*, membre del consell assessor de la revista *El professional de la informació* i d'*Hipertext.net* i revisor de diverses revistes científiques de la seva especialitat.

**Lluís Codina**

Professor titular del Departament de Comunicació de la Universitat Pompeu Fabra i director de la Unitat de Suport a la Qualitat i a la Innovació Docent (USQUID) de la Facultat de Comunicació. Imparteix docència en els Estudis de Periodisme i de Comunicació Audiovisual. Participa en màsters oficials i en programes de doctorat del Departament de Comunicació. Coordina el Grup de Recerca de Documentació Digital i Comunicació Interactiva i forma part del Grup de Recerca de Periodisme, tots dos grups reconeguts per la UPF i l'últim per la Generalitat de Catalunya. Fundador i codirector del primer màster que s'imparteix íntegrament al web sobre documentació digital, l'anuari *Hipertext.net* i coeditor del weblog sobre documentació digital *Servei d'Alerta*. Participa en el màster interuniversitari de la Universitat de Barcelona i de la Universitat Pompeu Fabra sobre Gestió de continguts digitals.

La revisió d'aquest material docent ha estat coordinada per la professora:
Mercè Vázquez García (2013)

Primera edició: febrer 2013

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

Tots els drets reservats

© d'aquesta edició, FUOC, 2013

Av. Tibidabo, 39-43, 08035 Barcelona

Disseny: Manel Andreu

Realització editorial: Eureka Media, SL

Dipòsit legal: B-588-2013



Els textos i imatges publicats en aquesta obra estan subjectes -llevat que s'indiqui el contrari- a una llicència de Reconeixement-NoComercial-SenseObraDerivada (BY-NC-ND) v.3.0 Espanya de Creative Commons. Podeu copiar-los, distribuir-los i transmetre'ls públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>

Introducció

L'assignatura *Bases de dades* té com a objectiu introduir-vos en el món de les bases de dades, el mecanisme que habitualment s'utilitza per a emmagatzemar dades de manera permanent i gestionar-les convenientment. Un cop definits alguns conceptes bàsics, repassarem l'evolució que ha seguit la gestió de dades i descriurem les estructures de dades més utilitzades en els sistemes d'informació.

A partir dels problemes que presenten els fitxers de dades tradicionals, justificarem l'ús de les bases de dades, de les quals exposarem objectius, característiques, avantatges i inconvenients. Així mateix, identificarem els elements que componen un sistema de base de dades, fent especial atenció als diferents tipus d'usuari.

Explicarem el concepte de *model de dades* el qual permet aconseguir el procés d'abstracció que condueix del món real al món de les dades, distingint entre l'esquema i l'estat de la base de dades, i introduïrem els diferents tipus d'abstracció que faciliten la representació de les dades en el disseny de bases de dades. Mostrarem com s'organitzen les dades de manera persistent en els suports d'emmagatzematge secundari i veurem com el sistema de gestió de bases de dades (SGBD) hi accedeix seguint l'esquema general del procés d'execució d'una consulta.

Per entendre les característiques dels SGBD n'analitzarem l'arquitectura des de diferents punts de vista i ens centrarem en l'arquitectura de tres nivells. L'anàlisi de les funcions que ha de fer l'SGBD ens ajudarà a determinar els components que ha de tenir per a complir-les: *llenguatges* de base de dades, que permeten definir-la i manejar-la; *interfícies*, que permeten a cada categoria d'usuari especificar-hi les sol·licituds; el *nucli*, format per diferents mòduls de programari que duen a terme tasques específiques; *utilitats* de valor afegit i aplicacions integrades, que permeten simplificar les tasques dels usuaris; i el *diccionari de dades*, que conté informació sobre les dades que els mòduls de l'SGBD necessiten consultar.

Amb relació a l'anàlisi de les bases de dades, encara que hi ha altres enfocaments, ens centrarem en les bases de dades relacionals, ja que actualment el model relacional és el més emprat. Concretament, d'aquest model estudiarem els fonaments teòrics i alguns llenguatges que s'utilitzen per a treballar-hi: l'àlgebra relacional i el llenguatge estàndard SQL.

En la darrera part del material aprofundirem en l'estudi de les bases de dades documentals a partir de les quatre operacions que s'hi poden fer: producció i administració, en què ens centrarem en l'estudi de l'estructura i les característiques dels programes informàtics (SGBD) que permeten crear i explotar bases de dades documentals; distribució, en què introduïrem el concepte d'*interfície de consulta* d'una base de dades documental i donarem indicacions per elaborar-la; disseny, en què ens centrarem en la metodologia per a la creació de bases de dades docu-

mentals; i finalment avaluació, en què descriurem els indicadors fonamentals per a avaluar una base de dades documental.

Aquest tipus de bases de dades es caracteritzen per facilitar l'accés a documents de tipus cognitiu (llibres, articles de revista, informes de tota mena, etc.). Contenen grans quantitats de text que s'emmagatzema de manera poc estructurada, sense fer servir les taules característiques del model relacional, i disposen d'índexs analítics i àmplies funcionalitats de cerca per a permetre la recuperació del contingut. El seu ús i extensió, però, són menors que el de les bases de dades relacionals.

Objectius

Els materials didàctics inclosos en aquesta assignatura s'orienten a aconseguir que assolis els objectius següents:

1. Conèixer els problemes que representa la gestió les dades.
2. Identificar els objectius i les característiques de les bases de dades.
3. Distingir els elements que estructuraven un sistema de base de dades.
4. Saber utilitzar l'abstracció per a modelitzar bases de dades.
5. Conèixer els mitjans d'emmagatzematge de bases de dades.
6. Comprendre l'arquitectura, el funcionament i els components dels sistemes de gestió de bases de dades.
7. Conèixer els fonaments del model de dades relacional.
8. Saber consultar una base de dades mitjançant les operacions de l'àlgebra relacional.
9. Saber utilitzar les sentències del llenguatge SQL per a definir i manipular bases de dades.
10. Conèixer l'estructura i el funcionament dels programes informàtics (SG-BD) que permeten la creació de bases de dades documentals.
11. Conèixer l'estructura i el funcionament dels programes informàtics que permeten la distribució de bases de dades al Web.
12. Determinar els recursos tecnològics necessaris per a la creació i distribució de bases de dades documentals.
13. Conèixer els elements bàsics de les interfícies web de consulta de bases de dades.
14. Conèixer i aplicar la metodologia per al disseny i la creació de bases de dades documentals.
15. Analitzar i avaluar bases de dades a partir d'indicadors.

Continguts

Mòdul didàctic 1

Sistemes de base de dades

Jaume Raventós Moret

1. Alguns conceptes bàsics
2. Evolució de la gestió de dades
3. Estructures de dades
4. Problemes de la gestió de fitxers de dades
5. Objectius i característiques de les bases de dades
6. Avantatges i inconvenients de les bases de dades
7. Elements d'un sistema de base de dades
8. Usuaris de les bases de dades
9. Model, esquema i estat de la base de dades
10. Tipus d'abstracció en el disseny de bases de dades
11. Arquitectura dels SGBD
12. Estructura global d'un sistema de base de dades
13. Emmagatzematge de bases de dades
14. Accés de l'SGBD a les dades
15. Funcions i components de l'SGBD

Mòdul didàctic 2

El model relacional i l'àlgebra relacional

Dolors Costal Costa

1. Introducció al model relacional
2. Estructura de les dades
3. Operacions del model relacional
4. Regles d'integritat
5. L'àlgebra relacional

Mòdul didàctic 3

El llenguatge SQL

Carme Martín Escofet

1. Sentències de definició de dades
2. Sentències de manipulació de dades
3. Sentències de control
4. Subllenguatges especialitzats

Mòdul didàctic 4

Sistemes de gestió documental i bases de dades documentals

Ernest Abadal, Lluís Codina

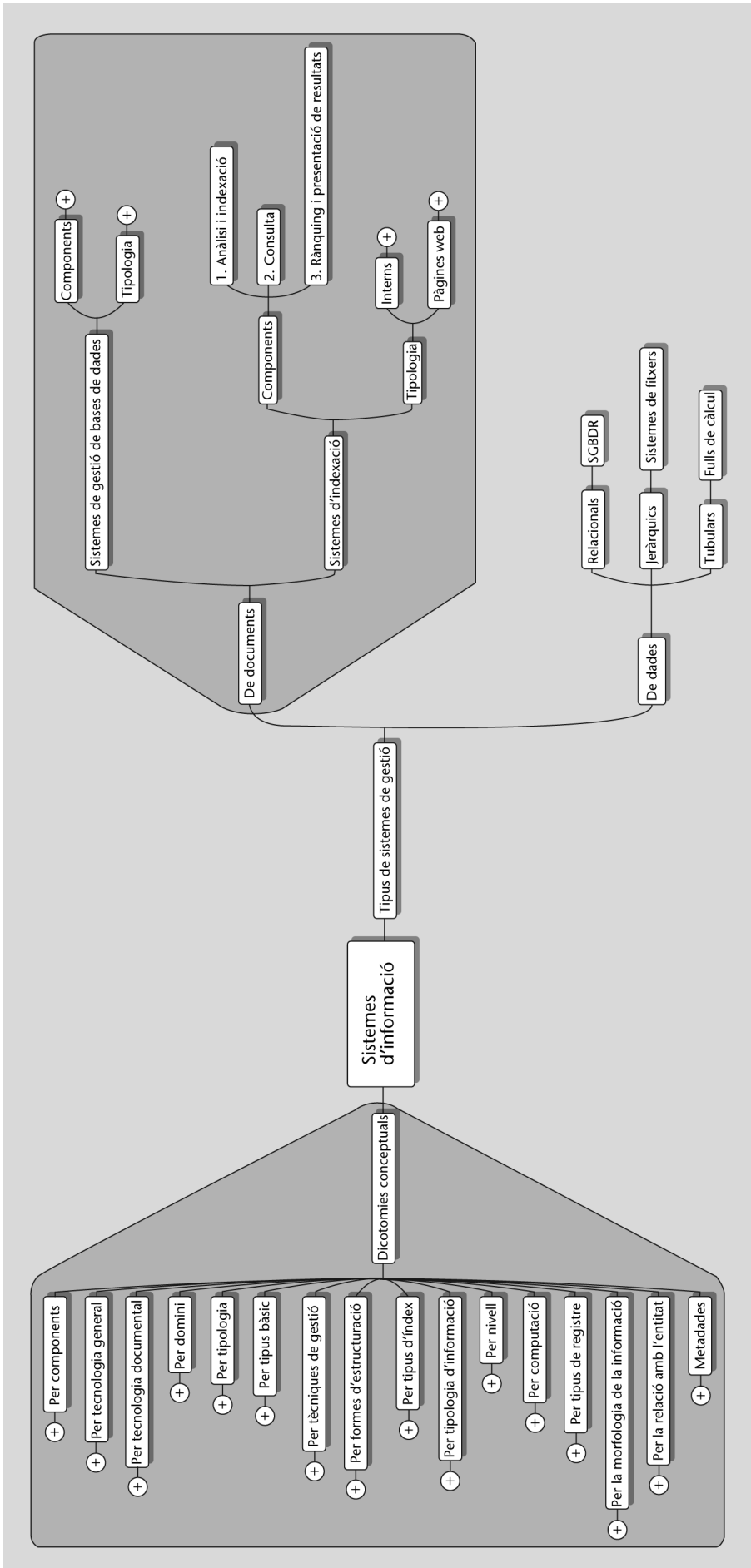
1. Producció i administració de bases de dades
2. Distribució de bases de dades
3. Metodologia per a la creació de bases de dades documentals
4. Avaluació de bases de dades

Glossari

Nomenclatura, definició i exemples de sistemes d'informació i de bases de dades

Concepte	Acrònim	Elements	Definició	Altres denominacions	Exemples
sistema d'informació		contingut (dades i informació) programari recursos materials (maquinari o arxivadors) persones	Sistema d'elements relacionats entre si que processen dades i informació per al compliment de determinats objectius d'una organització. Les activitats (entrada, emmagatzematge, processament i sortida d'informació) no necessàriament han d'estar informatitzades.	<i>information system</i>	Sistema d'informació manual: sistema d'arxivament de llibres en una biblioteca, les seves tasques d'organització i les persones que el gestionen o l'usen. Sistema d'informació informatitzat: el mateix, però suportat per un ordinador amb el programari corresponent.
sistema informàtic		programari maquinari persones	Sistema format pel conjunt dels components de maquinari i programari que permet el processament de dades. Pot formar part d'un sistema d'informació, però no necessàriament.	<i>computer system</i>	Un ordinador amb els seus programes, dispositius perifèrics i la persona que els maneja.
sistema de base de dades	SBD	contingut programari maquinari persones	Sistema d'informació basat en una base de dades que inclou el programari del sistema gestor de base de dades.	<i>database system</i>	BD científiques o acadèmiques (Scopus, Web of Knowledge, ACM Digital Library), BD de tesis doctorals, BD periodístiques, portals de revistes, hemeroteques i repositoris digitals.
base de dades	BD	contingut (dades)	Conjunt estructurat de dades que representen entitats del món real d'interès per a un projecte o una organització. Eventualment, entre aquests conjunts de dades es poden definir interrelacions. L'objectiu és organitzar la informació de manera que es pugui accedir i administrar les dades amb la màxima facilitat.	<i>database</i>	Gestió informatitzada (GI) de reparació de vehicles (clients, cotxes, reparacions, peces), GI d'activitats esportives (alumnes, cursos, participants, equips, activitats, calendari), GI de préstec bibliotecari.
base de dades relacional	BDR	contingut (dades)	Conjunt de dades estructurat d'acord amb el model relacional.	<i>relational database</i>	BD de lloguer de cotxes, BD de comptes bancaris, BD de vendes, BD de gestió hotelera (clients, habitacions, calendari).
base de dades documental	BDD	contingut (documents o referències)	Conjunt de dades i/o documents estructurat d'acord amb el model documental.	<i>documental database</i>	BD de premsa (<i>El País</i>), BD de recursos d'Internet (Intute), BD de fotografies (AGE Fotostock), BD de recursos i articles de revistes d'educació (ERIC).
sistema de recuperació	SR	programari	Sistema que facilita la recuperació de dades o informació (documents)	<i>retrieval system</i>	
sistema de recuperació d'informació	SRI	programari	Sistema que permet la recuperació d'informació d'un fons documental per mitjà d'una interfície de consulta i un llenguatge d'interrogació amb operadors que permeten optimitzar la cerca.	<i>information retrieval system</i>	Motor de cerca (Google, Yahoo), BD referencial (LISA).
sistema de gestió de bases de dades	SGBD	programari	Sistema destinat a la creació, la gestió, el manteniment i l'explotació de bases de dades.	<i>database management system</i>	Els exemples de SGBDR i SGD.

Concepte	Acrònim	Elements	Definició	Altres denominacions	Exemples
sistema de gestió de bases de dades relacionals	SGBDR	programari	Sistema de gestió de bases de dades que suporta el model relacional i permet operacions basades en aquest model.	<i>relational database management system</i>	Programari propietari: dBase, Access, Oracle, Informix, SQL Server, DB2 Programari lliure (codi obert): PostgreSQL, Sybase, MySQL
sistema de gestió documental	SGD	programari	Sistema d'informació que incorpora i gestiona documents als quals proporciona accés en el decurs del temps.	<i>document retrieval system</i> <i>text retrieval system</i>	Els exemples d'SGBD, d'SGB i de sistemes d'indexació.
sistema de gestió de bases de dades documentals	SGBDD	programari	Sistema de gestió de bases de dades que suporta el model documental.	<i>documental database management system</i>	CDS/ISIS, Inmagic, Knosys, FileMaker
sistema de gestió bibliogràfica	SGB	programari	Sistema de gestió documental especialitzat en l'emmagatzematge i la recuperació de referències bibliogràfiques, i la generació de bibliografies amb diferents formats.	<i>bibliographic management system</i>	Sistema d'escriptori: Reference Manager, ProCite Sistema en línia: RefWorks, EndNoteZotero, Connotea, CiteUlike, Mendeley
sistema d'indexació		programari	Sistema de gestió documental que genera índexs analítics (fitxers invertits) a partir del contingut de documents textuais.	motor de cerca <i>Text retrieval software</i> <i>full-text retrieval system</i> <i>Text information management system</i>	AskSam, Autonomy, Google Custom Search, Greenstone



Font: per veure el gràfic original i desplegar tot el contingut vegeu <http://tinyurl.com/diagramasgd>.

Bibliografia

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

Beaulieu, A. (2010). *Aprende SQL* (2a. 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". *Anuari SOCADI de Documentació i Informació* (pàg. 195-209). Barcelona: SOCADI.

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

Cuadra, D. i altres (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* (7a. ed.). Madrid: Prentice-Hall.

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

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

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

Informix Software, Inc. (1996). *Informix Guide to SQL Syntax Version 7.2* (vol. 1 y 2). Menlo Park (Califòrnia): 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. i altres (2006). *Tecnología y diseño de bases de datos* (1a. ed.). Madrid: Ra-Ma.

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

Raya, F. (1987). *Database design for information retrieval: a conceptual approach*. Nova 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, pàg. 9-23). Madrid : Consejo Superior de Investigaciones Científicas, Instituto de Información y Documentación en Ciencia y Tecnología.

Silberschatz, A.; Korth, H. F.; Sudarshan, S. (2006). *Fundamentos de bases de datos* (5a. ed.). Madrid [etc.]: McGraw-Hill.

Villanueva, E. (1996, gener-juny). "Bases de datos y bibliotecología: cómo deshacer la innecesaria incomunicación". A: *Investigación bibliotecológica* (vol. 10, núm. 20, pàg. 27-32). Mèxic, 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.

