

# *Data warehouse*

Magatzems de dades i models  
multidimensionals

Àngels Rius Gavídia  
Montse Serra Vizern  
Alberto Abelló Gamazo  
José Samos Jiménez  
Josep Vidal Portolés  
Josep Curto Díaz

PID\_00181200

**Àngels Rius Gavídia**

Llicenciada en Informàtica per la Universitat Politècnica de Catalunya. Professora associada adscrita a l'Escola Universitària Politècnica de Vilanova i la Geltrú des de 1996. El seu àmbit de recerca s'enmarca dins del camp de les bases de dades. Actualment professora pròpia dels Estudis d'Informàtica i Multimèdia de la Universitat Oberta de Catalunya.

**Montse Serra Vizern**

Enginyera en Informàtica i màster en Combinatòria i Comunicació Digital per la Universitat Autònoma de Barcelona. Professora del programa especial Citius d'Iniciació Professional a l'Empresa de la Universitat Autònoma de Barcelona. El seu àmbit de recerca s'emmarca dins del camp de l'ensenyament en CTS (ciència, tecnologia i societat). Actualment és professora pròpia dels Estudis d'Informàtica i Multimèdia de la Universitat Oberta de Catalunya.

**Alberto Abelló Gamazo**

Doctor i enginyer en Informàtica per la Universitat Politècnica de Catalunya. Professor associat al departament de Llenguatges i Sistemes Informàtics d'aquesta universitat.

**José Samos Jiménez**

Doctor en Informàtica per la Universitat Politècnica de Catalunya. Professor titular del Departament de Llenguatges i Sistemes Informàtics de la Universitat de Granada, assignat a l'Escola Tècnica Superior d'Enginyeria Informàtica.

**Josep Vidal Portolés**

Enginyer tècnic en Informàtica de Gestió i diplomat professor d'EGB. Professor associat a la Universitat de Girona, en l'especialitat de sistemes d'informació. Cap de projectes i desenvolupament del gestor de dades d'una entitat financera (Caixa de Girona).

**Josep Curto Díaz**

MBA per l'IE Business School, màster en Business intelligence i màster en Direcció i gestió en sistemes i tecnologies de la informació per la Universitat Oberta de Catalunya i llicenciat en Matemàtiques per la Universitat Autònoma de Barcelona. Consultor a la Universitat Oberta de Catalunya especialitzat en intel·ligència de negoci (*business intelligence*). Analista sènior a IDC Research Spain.

La revisió d'aquest material docent ha estat coordinada per la professora: Àngels Rius Gavídia (2012)

Segona edició: febrer 2012

© Àngels Rius Gavídia, Montse Serra Vizern, Alberto Abelló Gamazo, José Samos Jiménez, Josep Vidal

Portolés, Josep Curto Díaz

Tots els drets reservats

© d'aquesta edició, FUOC, 2012

Av. Tibidabo, 39-43, 08035 Barcelona

Disseny: Manel Andreu

Realització editorial: Eureka Media, SL

Dipòsit legal: B-1.018-2012

*Cap part d'aquesta publicació, incloent-hi el disseny general i la coberta, no pot ser copiada, reproduïda, emmagatzemada o transmesa de cap manera ni per cap mitjà, tant si és elèctric com químic, mecànic, òptic, de gravació, de fotocòpia o per altres mètodes, sense l'autorització prèvia per escrit dels titulars del copyright.*

## Introducció

Com a conseqüència immediata i donada la tendència actual de les noves tecnologies i Internet, l'assignatura de *Models multidimensionals i magatzems de dades* es presenta com un pas endavant dins l'àmbit d'estudi de les bases de dades.

En primera instància es defineix el concepte de magatzem de dades a partir de les propietats que el caracteritzen i que alhora el diferencien respecte de les bases de dades convencionals. La característica principal que distingeix el magatzem de dades de la base de dades tradicional és l'objectiu pel qual ha estat concebut: oferir suport a la presa de decisions. Perquè la informació continguda en el magatzem pugui servir d'ajut a la presa de decisions, és necessari que el magatzem sigui dissenyat seguint un model de dades específic: el model de dades multidimensional.

Un sistema d'informació que tingui com a objectiu donar suport a la presa de decisions, sigui quina sigui la naturalesa de l'organització, no es pot basar únicament en el magatzem de dades. Cal la integració d'un conjunt de components que girin al voltant d'un eix central: el magatzem de dades. Tots aquests elements que hem citat anteriorment configuren el que s'anomena *factoria de la informació corporativa* (FIC).

Construir una FIC no resulta gens fàcil, és costós tant en recursos com en la seva pròpia concepció, i requereix un canvi cultural dins la corporació. Però si s'aconsegueix de dur a terme fins al final i s'ha dissenyat correctament, la gran inversió feta *a priori* s'arribarà a traduir en beneficis que col·locaran la institució en una posició competitiva dins el mercat actual. Un cop construïda la FIC, el magatzem de dades pot donar suport a noves estratègies de negoci (*business intelligence*, ERP, CRM), i també a aplicacions (OLAP, EIS, mineria de dades, *webhousing*) destinades a l'explotació de les dades amb un nivell de detall molt acurat.

Esperem que el contingut d'aquesta assignatura us ensenyi el camí cap a nous horitzons dins l'àmbit de les bases de dades que us puguin ser útils en el vostre futur professional.

## Objectius

Els objectius que es pretén que els estudiants assoleixin amb aquesta assignatura són els següents:

- 1.** Arribar a un nivell de coneixements alt per poder dissenyar un magatzem de dades, triar la millor arquitectura que el suporti i saber-lo administrar.
- 2.** Conèixer els diferents tipus d'aplicacions i eines per a una òptima explotació del magatzem de dades (EIS, OLAP).
- 3.** Saber triar les millors tècniques per extreure informació del magatzem de dades, que servirà per a arribar a conclusions importants que permetran d'emprendre un nou pla d'acció que faci més competitiva l'empresa.
- 4.** Saber analitzar l'entorn de negoci de l'empresa: quins mòduls funcionals necessarien, model client/servidor, quines bases de dades s'haurien de compartir, tipus d'interfície, etc.
- 5.** Saber com s'han d'implementar processos de negoci d'alt nivell en les empreses, processos que proporcionin un rendiment competitiu.
- 6.** Poder proposar solucions integrades i fer entendre als clients que la implantació d'un ERP és un projecte complex a causa del seu profund impacte en els processos de l'empresa (canvi cultural de l'empresa, canvi en els processos de negoci, canvi en l'organització i canvi en la disciplina de treball).
- 7.** Tenir els coneixements necessaris per a poder seleccionar de manera adequada proveïdors i productes tenint en compte la problemàtica de l'empresa (Oracle, SAP, Baan Company, PeopleSoft, Computer Associates, Microstrategy, etc.).

## Continguts

### Mòdul didàctic 1

#### **Introducció a l'emmagatzematge de dades**

Àngels Rius Gavídia, Montse Serra Vizern i Josep Curto Díaz

1. Què és un magatzem de dades?
2. Evolució històrica
3. Característiques d'un magatzem de dades
4. Objectius d'un magatzem de dades
5. Comparativa: magatzem de dades i bases de dades operacionals

### Mòdul didàctic 2

#### **La factoria d'informació corporativa**

Alberto Abelló Gamazo, José Samos Jiménez i Josep Curto Díaz

1. Usuaris i fons d'informació dels magatzems de dades
2. Els magatzems de dades departamentals
3. El magatzem de dades corporatiu
4. El magatzem de dades operacional
5. El component d'integració i transformació
6. Les metadades
7. La factoria d'informació corporativa

### Mòdul didàctic 3

#### **Construcció de la factoria d'informació corporativa**

José Samos Jiménez i Josep Curto Díaz

1. Punt de partida: suport a la presa de decisions en les organitzacions
2. Variants de la FIC
3. Estratègies per a la construcció de la FIC
4. Desenvolupament del component d'integració i transformació
5. Desenvolupament del magatzem de dades corporatiu
6. Incorporació del magatzem de dades operacional

### Mòdul didàctic 4

#### **Disseny multidimensional**

Alberto Abelló Gamazo

1. Les necessitats dels analistes i les eines OLAP
2. Components del model multidimensional
3. Disseny conceptual
4. Disseny lògic
5. Consultes amb SQL'99
6. Disseny físic

### Mòdul didàctic 5

#### **Ús dels magatzems de dades**

Josep Vidal Portolés i Josep Curto Díaz

1. Beneficis dels magatzems de dades
2. Eines d'exploració dels magatzems de dades
3. Estratègies de negoci

## **Bibliografia**

**Devlin, B.** (1997). *Data Warehouse from Architecture Implementation*. EUA: Addison Wesley Longman Inc.

**Inmon, W. H.** (1999). *Building the Operational Data Store*. EUA: John Wiley & Sons Inc.

**Inmon, W. H.** (2005). *Building the Data Warehouse* (4a. ed.). EUA: John Wiley & Sons Inc.

**Inmon, W. H.; Hackathorn, R. D.** (1994). *Using the Data Warehouse*. Nova York: Wiley.

**Inmon, W. H.; Imhoff, C.; Sousa, R.** (1998). *Corporate Information Factory*. EUA: John Wiley & Sons Inc.

**Inmon, W. H.; Strauss, D.; Neushloss, G.** (2008). *DW 2.0: The Architecture for the next generation of Data Warehousing*. EUA: Morgan Kaufman Series.

**Kimball, R.** (2009). *Data Warehouse Toolkit Classics: The Data Warehouse Toolkit; The Data Warehouse Lifecycle Toolkit; The Data Warehouse ETL* (2a. ed.). Nova York: John Wiley & Sons Inc.

