

## 1. Introducció

Abans de l'aparició de les aplicacions informàtiques, les empreses empraven com a úniques eines de gestió de dades, els fitxers amb calaixos, carpetes i fitxers de cartó.

Els sistemes d'informació actuals es basen en Bases de Dades (BBDD) i Sistemes Gestors de Bases de Dades (SGBD) que s'han convertit en elements imprescindibles de la vida quotidiana de la societat moderna.

### 1.1. Les dades i la seva representació

Una base de dades, des del punt de vista informàtic, podem dir que és un sistema format per un conjunt de dades emmagatzemades que permeten, que un conjunt de programes les puguin manipular.

- DADES: Són representacions informàtiques de la informació disponible, relativa als objectes del món real.
- INFORMACIÓ: Quan les dades s'organitzen i es tracten, aconseguim la informació.
  - Món real: Constituït pels objectes materials i no materials, de la realitat.
  - Món conceptual: coneixements o informacions que obtenim mitjançant l'observació de la part del món real de cada persona.
  - Món de les representacions: Format per les representacions informàtiques, de les del món conceptual, necessàries per poder fer feina.
- ENTITATS: Les entitats són els objectes del món real que conceptualitzem. Són identificables, és a dir, distingibles. Pot tenir dos significats diferents:
  - Entitat tipus: Tipus genèric d'entitat, per exemple, els cotxes, en general.
  - Entitat instància: Conceptualització d'un objecte concret del món real, com ara un cotxe concret, distingible gràcies a algun atribut.
- ATRIBUTS: Són les propietats de les entitats que ens interessin. Els valors que un atribut determinat pot prendre s'anomena domini.
  - Valor nul: El valor d'un atribut pot ser desconegut o inexistent. L'expressió valor nul indica que no hi ha cap valor associat a un atribut determinat d'una entitat instància concreta.
  - Atribut identificatiu clau: Tot atribut o conjunt d'atributs que permeten identificar inequívocament les instàncies d'una entitat s'anomenen claus. Tot atribut identificador és, al mateix temps, una clau.
- MÓN DE LES REPRESENTACIONS: La representació informàtica més freqüent en l'àmbit de les BD és la representació tabular, la qual s'implementa habitualment en fitxers que s'estructuren en registres i camps. Les BD només són conjunts de fitxers interrelacionats.
  - Cada taula representa una entitat genèrica.
  - Cada columna representa un atribut.
  - Cada cel·la emmagatzema el valor que tingui l'atribut de l'entitat instància.

## 1.2. Sistemes de fitxers i evolució cap als SGBD

- EVOLUCIÓ DE L'EMMAGATZEMAMENT DE LA INFORMACIÓ
  - Fitxers manuals: Sistema de fitxers manuals, que era un conjunt de carpetes etiquetades i ordenades i classificades dins d'un armari o dins d'un arxiu.
  - Sistemes de fitxers: Les dades necessàries es guardaven en fitxers a l'ordinador i aquestes aplicacions accedien a ells per obtenir els informes.
  - Sistemes gestors de Bases de dades: Qualsevol canvi en l'estructura de les dades no afecta els programes d'aplicació que utilitzen aquestes dades.
- FITXERS: Per l'emmagatzematge de la informació de forma permanent s'utilitzen dispositius d'emmagatzematge massiu anomenats memòries secundàries,
  - Dades: Són els fets o aspectes que necessitem guardar per obtenir algun tipus d'informació. Per exemple, un fitxer amb dades de clients d'una tenda.
  - Camp: És un caràcter o conjunt de caràcters que tenen significat específic. Per exemple, crearem un camp per guardar el nom del client, un altre per guardar el seu telèfon...
  - Registre: Conjunt de camps relacionats seguint una certa lògica, que descriuen, una persona, un lloc, una cosa...
  - Fitxer: És un conjunt de registres relacionats. A l'exemple, el fitxer estaria format per totes les dades dels clients de la tenda.
- DISPOSITIUS D'EMMAGATZEMATGE D'INFORMACIÓ
  - Dispositius d'emmagatzematge per mitjà magnètic: Aquests dispositius són els més antics i utilitzats a gran escala.
    - Discs durs externs.
    - Discs durs interns.
  - Dispositius d'emmagatzematge per mitjà òptic: Aquests dispositius fan servir com a suport per emmagatzemar les dades en la superfície reflectant. Per llegir aquest disc necessitem una unitat lectora o gravadora instal·lada en el nostre equip de treball.
  - Dispositius d'emmagatzematge per mitjà electrònic (basades en memòria flash):
    - Memòria USB: (universal serial bus).
    - Targetes de memòria.
    - Unitats d'estat sòlid (SSD).
- TIPUS DE FITXERS

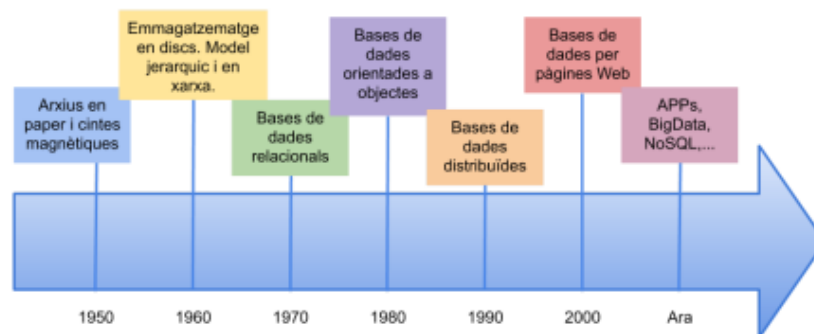
Segons la forma d'accés.

  - Seqüencial: Els registres es llegeixen un darrere l'altre des del començament del fitxer fins que s'arriba al final del fitxer.
  - Directe: Permet seleccionar un registre i accedir directament a ell mitjançant una clau.
  - Indexat: Per seleccionar un registre consultam prèviament de forma seqüencial dins d'una taula que conté la clau més alta i l'adreça de començament de cada bloc de registres.
  - Dinàmic: Permet l'accés directe o per índex a un registre i a partir d'aquest s'accedeix de forma seqüencial.

## - INCONVENIENTS DELS FITXERS

- Inconsistència de la informació: Un error en l'actualització d'aquesta informació pot provocar inconsistències en les dades del fitxer.
- Redundància: Aquest problema és quan es tenen dades que no aporten informació i que es poden calcular a partir d'altres dades. Per exemple, l'import amb IVA de les factures del fitxer de factures de proveïdors.
- Rigidesa de cerca: Cada fitxer té una organització determinada, segons el tipus d'accés per al qual s'ha definit.
- Dependència dels programes: Qualsevol canvi en l'estructura del fitxer representa modificar els programes que l'utilitzen.

## - EVOLUCIÓ DELS SISTEMES D'EMMAGATZEMATGE DE LA INFORMACIÓ



### 1.3. Sistemes Gestors de Bases de Dades

#### - SISTEMES GESTORS DE BASES DE DADES

Les seves característiques principals són:

- Abstracció: Els SGBD estalvien als usuaris detalls de l'emmagatzemament físic de les dades.
- Independència de dades: És la capacitat de modificar l'estructura física o lògica d'una BD sense haver de canviar les aplicacions que hi accedeixen.
- Redundància: No hi ha dades repetides. Problema d'espai però, sobretot, del manteniment de la qualitat de les dades.
- Integritat i consistència de dades: Els SGBD han d'assegurar el manteniment de la qualitat de les dades. Si hem permès que hi hagi redundància, és necessari que sigui el SGBD que vigili que aquesta informació que apareix repetida s'actualitzi de forma coherent.
- Seguretat: Han de garantir que la informació està segura, protegida d'accessos no autoritzats.

#### - ARQUITECTURA DELS SGBD

- Esquemes i nivells: L'esquema de la BD és un element fonamental de l'arquitectura d'un SGBD que permet independitzar-ho de la BD, canviar el disseny de la BD, el seu esquema, sense haver de canviar el SGBD.
  - Nivell Lògic: Descriu els conceptes que formen la base de dades en forma de models o esquemes.
  - Esquemes externs: Són les diverses visions lògiques que els processos usuaris (programes d'aplicació i usuaris directes) tindran de les parts de la BD que utilitzaran.

- Esquema Conceptual: En ell s'hi expliciten les dades, les interrelacions i les restriccions que s'han establert. Permet treballar amb les dades de manera més senzilla.
- Nivell Intern o Físic: Contindrà la descripció de l'organització física de la BD. Implica un coneixement a baix nivell de la implementació física de l'organització de les dades i el seu accés.
- Independència de les dades:
  - Independència física: Quan els canvis en l'organització física de la BD no afecten el nivell lògic.
  - Independència lògica: Quan els usuaris no es veuen afectats pels canvis en el nivell lògic.
- ETAPES DE DISSENY D'UNA BDD
  - Etapa 0 o prèvia: "Anàlisi de requeriments"  
Aquest pas consisteix a entendre quines dades hauran de ser emmagatzemades dins la Base de Dades, quines aplicacions hauran de ser construïdes al voltant d'ella i quines d'elles són més freqüents.
  - Etapa del "Disseny conceptual"  
És l'etapa on s'obté una estructura de la informació de la futura Base de Dades. L'objectiu d'aquesta etapa és elaborar un model conceptual del problema.
  - Etapa del "Disseny lògic"  
El seu objectiu és transformar els esquemes obtinguts a l'etapa anterior, en un conjunt de dades pròpies del model de dades escollit, per això es treballa amb el model abstracte de dades obtingut al final de l'etapa de disseny conceptual.
  - Etapa del disseny "físic"  
En aquesta etapa serà necessari conèixer la funcionalitat de SGBD escollit. També és en aquest moment on dissenyen els mecanismes de seguretat.

- SISTEMES DE FITXERS TRADICIONALS VS SGBD

	Fitxers	Bases de dades
<b>Entitats tipus</b>	Les entitats instància d'un fitxer pertanyen a una sola entitat tipus.	Les BD contenen entitats instància d'infinat d'entitats tipus interrelacionades.
<b>Interrelacions</b>	El sistema no interrelaciona fitxers.	El sistema té previstes eines per interrelacionar fitxers.
<b>Redundàncies</b>	És necessari crear fitxers a mida de cada aplicació, amb totes les dades necessàries, encara que estiguin repetides en altres fitxers.	Tècnicament, totes les aplicacions poden treballar amb la mateixa BD, la qual cosa evita la redundància de dades i els riscos que comporta.
<b>Inconsistències</b>	És possible que els valors d'unes mateixes dades en diferents fitxers no coincideixin, si els programadors no les han actualitzat degudament.	Si les interrelacions estan ben dissenyades, les dades només han d'estar emmagatzemades en la BD un sol cop. Per tant, no hi ha risc d'inconsistències.
<b>Obtenció de dades</b>	Si no hi ha una aplicació que obtingui les dades que volem, o bé s'ha de fer un programa a mida, o bé s'ha d'aprofitar la sortida d'un programa amb objectius similars, i fer els càlculs necessaris manualment.	Permeten obtenir qualsevol conjunt de dades, segons les necessitats, dels del seu propi entorn de treball, sense haver d'escriure, compilar i executar cap nou programa d'aplicació contra la BD.

<b>Aïllament de dades</b>	Les dades estan disperses i aïllades en diferents arxius, la qual cosa dificulta el desenvolupament de les aplicacions.	Totes les dades són en la mateixa BD, interconnectades, la qual cosa en facilita l'obtenció.
<b>Integritat de dades</b>	Els programes han d'implementar totes les restriccions sobre les dades, afegint el codi font corresponent. El manteniment és complicat quan la informació es conté en diferents fitxers utilitzats per diferents aplicacions.	La BD s'encarrega directament d'implementar les restriccions sobre les dades. Els programes no han d'incorporar codi font addicional per garantir-les.
<b>Atomicitat</b>	Alguns conjunts d'operacions sobre les dades s'han d'executar de manera indivisible (o tots o cap), independentment de les fallades que el sistema pugui presentar (com ara per un tall de subministrament elèctric). Però això és molt difícil de garantir amb un sistema d'informació basat en fitxers.	Les BD incorporen la tècnica de les transaccions per tal de garantir fàcilment l'execució atòmica d'una pluralitat de processos sobre les dades.

<b>Accés concurrent</b>	L'actualització simultània de dades d'un mateix fitxer per part de diferents usuaris o aplicacions en pot provocar fàcilment la inconsistència.	Amb la tècnica del bloqueig, les BD garanteixen automàticament la consistència de les dades, malgrat que més d'un usuari o més d'una aplicació les vulguin actualitzar simultàniament.
<b>Seguretat</b>	Habitualment, cada fitxer serveix per a un sol usuari o una sola aplicació (sobretot simultàniament), i ofereix una visió única del món real. Però no sempre tots els usuaris que utilitzen un fitxer haurien de tenir accés a totes les dades que conté.	Una BD pot ser compartida per molts usuaris de diferents tipus (fins i tot, simultàniament), els quals poden tenir diferents visions (vistes) del món real, en funció del seu perfil i dels permisos que s'hagin de concedir en cada cas.

#### 1.4. Models de Bases de Dades

Els models són els instruments que s'apliquen a una parcel·la del món real per obtenir una estructura de dades a la que anomenem esquema.

Un model de dades és un conjunt d'eines conceptuais que permeten descriure les dades, les seves relacions i les regles d'integritat que han de complir.

Els models de BD es poden agrupar en tres grups:

- MODELS LÒGICS BASATS EN OBJECTES: Aquests models s'apliquen en els dos nivells superiors, és a dir al nivell conceptual i a l'extern.
  - MODEL ENTITAT-RELACIÓ: Es caracteritza perquè està format per un conjunt d'elements anomenats entitats.
  - MODEL ORIENTAT A OBJECTES: Els models estrictament orientats a objectes defineixen les BD en termes d'objectes, de les seves propietats i de les seves operacions. Les operacions de cada classe s'especifiquen en termes de procediments predefinits, anomenats mètodes.
    - CLASSES
    - VARIABLES
    - MÈTODES
    - COMENTARIS
- MODELS LÒGICS BASATS EN REGISTRES: Aquests models s'apliquen en els dos nivells inferiors, és a dir al nivell conceptual i al físic.
  - MODEL JERÀRQUIC.
  - MODEL EN XARXA.
  - MODEL RELACIONAL: Les dades s'estructuren en representacions tabulars, anomenades taules, que representen entitats tipus del món conceptual, i que estan formades per files i columnes. Les columnes formen els camps i les files són els registres. En el models de dades jeràrquic i en el model de dades en xarxa, les dades s'estructuraven gràcies a dos elements: els registres i les interrelacions. Però el model relacional només consta d'un element: les relacions o taules.

Avantatges:

- Evita la duplicitat de registres.
- Vetlla per la integritat referencial.
- Afavoreix la comprensibilitat.

- MODELS FÍSICS DE DADES: El model físic de dades representa com es construirà el model a la base de dades físicament al sistema informàtic. Un model físic de base de dades mostra totes les estructures de taula.

Els models relacionals no són vàlids per alguns programes que necessiten tractar un volum de dades molt gran, o dades relativament desestructurades.

- BASES DE DADES DOCUMENTALS: Són aquelles que fan servir un conjunt de programes que gestionen dades a partir de documents o arxius amb diferents estructures.