

FUND. BASES DE DATOS - TEMA 1

CONCEPTO INTUITIVO DE BASE DE DATOS

Casi todas las empresas requieren de aplicaciones para poder gestionar y acceder a información. Que los datos pertenezcan a las aplicaciones implica que pueda generarse:

- **Redundancia:** una misma información está almacenada en más de un sitio
- **Inconsistencia:** datos no actualizados o que se contradicen (un mismo paciente tiene dos edades diferentes o ha estado en dos consultas simultáneamente)
- **No hay reutilización** de datos

Para organizar un conjunto de información, la integramos y compartimos en diferentes estructuras. Llevamos a cabo esta idea utilizando un software especial que nos permite crear y usar almacenes de datos centralizados.

Concepto intuitivo de BD: fondo común de información almacenada en una computadora para que cualquier persona o programa (usuario) autorizado (control de acceso) pueda acceder a ella, independientemente del lugar de procedencia y del uso que haga de la misma. Una BD cobra sentido si hay varios usuarios.

Operaciones que se pueden realizar con una BD

El objetivo es poder gestionar datos de manera transparente (sin tener que programar código para acceder a los ficheros). Lo que queremos hacer será “CRUD”:

- **Insertar** datos (Create)
- **Obtener** datos previamente insertados (Read)
- **Modificar** datos existentes (Update)
- **Borrar** datos existentes (Delete)

SISTEMAS DE GESTIÓN DE BD

Base de Datos (BD): conjunto de datos comunes a un “proyecto” almacenados sin redundancia para ser útiles a diferentes aplicaciones.

Sistema de Gestión de BD (SGBD): conjunto de elementos software con capacidad para definir, mantener y utilizar una BD. En inglés DBSM. Entre otros, debe permitir:

- Definir estructuras de almacenamiento
- Acceder a los datos de forma eficiente y segura
- Organizar la actualización de los datos y el acceso multiusuario

Los elementos que están involucrados en una BD son los siguientes:

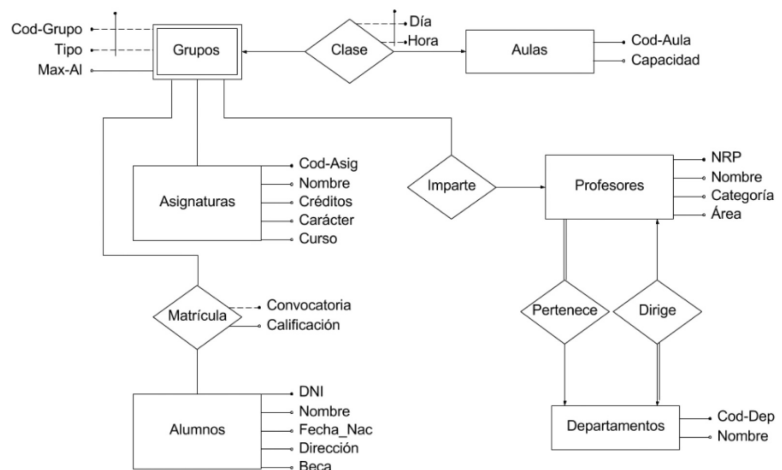
1. **Datos:** que estarán integrados (sin redundancia) y compartidos (útiles a varias aplicaciones)
2. **Hardware:** del que dependerá tener una BD centralizada o una BD distribuida (para problemas más complejos)
3. **Software (SGBD):** programas para definir las estructuras y gestionar la información de la BD
4. **Usuarios:** normalmente son las aplicaciones, no las personas. Este grupo engloba al usuario final, al programador de aplicaciones y al administrador (DBA, DBM)

A continuación, seguimos definiendo conceptos relacionados con una BD.

Dato operativo: Pieza de información que necesita una organización para su funcionamiento. Consta de:

- **Ítem básico:** elemento sobre el que se puede pedir información (~ sustantivos)
ej: **estudiante, asignatura, profesor**
- **Atributo:** característica de los ítems básicos (~ adjetivos)
ej: **nombre, apellidos, DNI, código asignatura**
- **Relación:** conexión lógica entre ítems (~ verbo)
ej: **profesor imparte asignatura, estudiante matriculado en asignatura**

Al determinar y clasificar de esta forma todos los datos operativos, se obtiene el **esquema lógico** de la BD. Para representarlos, hacemos uso de notaciones gráficas. Nosotros usaremos el **diagrama E/R**.



Pasamos ahora a estudiar los **objetivos de un SGBD**:

- **Independencia** de los datos.
- Diseño y utilización **orientada al usuario**, los datos y aplicaciones deben ser accesibles a los usuarios de la manera más amigable posible. Para ello:
 - + Se escoge un modelo de datos teórico (el más exitoso es el de la tabla)
 - + Se busca que sea fácil de implementar y para ello ha de estar perfectamente definido
 - + Se implementan lenguajes de acceso y modificación

- **Centralización:** los datos deben gestionarse de forma centralizada e independiente de las aplicaciones.

- **No redundancia:** los datos no deben estar duplicados si no hay una razón de peso para ello. Esto supone una nueva complicación: gestionar accesos concurrentes a un mismo dato.

- **Consistencia:** los datos no deben tener fallos lógicos. Deben respetar las reglas definidas en la organización (ej: **dos distintas personas no pueden tener el mismo DNI, no puede haber una asignatura sin nombre...**). Para ello, se desarrollan mecanismos de mantenimiento de integridad.

Además, cualquier operación debe llevar a la BD de un estado válido a otro válido. De aquí nacen las transacciones, operaciones que se efectúan en su totalidad o que no se realizan en absoluto.

- **Fiabilidad:** los datos han de estar protegidos frente a fallos catastróficos, como que se vaya la luz mientras se hace una transacción bancaria, o que se incendie el edificio donde guardo físicamente mis datos. Para ello, se desarrollan mecanismos de recuperación y relanzamiento de transacciones.

- **Seguridad:** no todos los datos han de ser accesibles a todos los usuarios. Para ello, se desarrollan mecanismos de gestión de usuarios y privilegios y mecanismos de protección de datos.

VENTAJAS DE UTILIZAR UN SGBD

Para el usuario:

- Usuario final: puede acceder a los datos
- Programador de aplicaciones: elimina problemas de a) diseño lógico y físico, b) depuración de errores y c) mantenimiento en general (copias de seguridad, recuperación de fallos...)
- Administrados de BD: puede adoptar y ejecutar decisiones a todos los niveles

Para el sistema:

- Control centralizado: fiabilidad, consistencia, seguridad...
- Criterios de uniformización (los datos organizados con un mismo patrón)
- Generación de nuevas aplicaciones
- Equilibrio entre requerimientos conflictivos

CONCEPTO DE INDEPENDENCIA

Independencia: los datos se organizan independientemente de las aplicaciones que los vayan a usar y de los archivos en los que vayan a almacenarse. Se clasifica en física y lógica.

Independencia física: el diseño lógico de la BD ha de ser, a todos los niveles, independiente del almacenamiento físico de los datos. Esto significa que los datos no se acceden por localización, sino por semántica (ej: en vez de “acceder posición X de memoria” es “acceder datos de los estudiantes”).

Esto nos permite realizar cambios en la estructura física sin alterar la lógica de las aplicaciones. Además, estas últimas no tendrán que gestionar aspectos relativos al almacenamiento.

Independencia lógica: cada aplicación debe poder organizar los datos según sus propios esquemas y acceder a los datos que le son necesarios y le conciernen (vista de usuario).

La independencia lógica persigue que los cambios en el esquema lógico general no afecten a las vistas de usuario. De este modo, las aplicaciones no necesitarían ser modificadas ante un cambio de este tipo. Sin embargo, no siempre es posible conseguir esta independencia, pues ciertos cambios lógicos requieren modificaciones en las vistas de usuario.

Esta independencia asegura a) un aumento de seguridad y fiabilidad, b) menos problemas para las aplicaciones y c) posibilitar a los desarrolladores de las aplicaciones y al DBM cambiar los esquemas.

- FIN DEL TEMA 1 -