BASES DE DATOS
TEMA 1

Introducción a las Bases de Datos

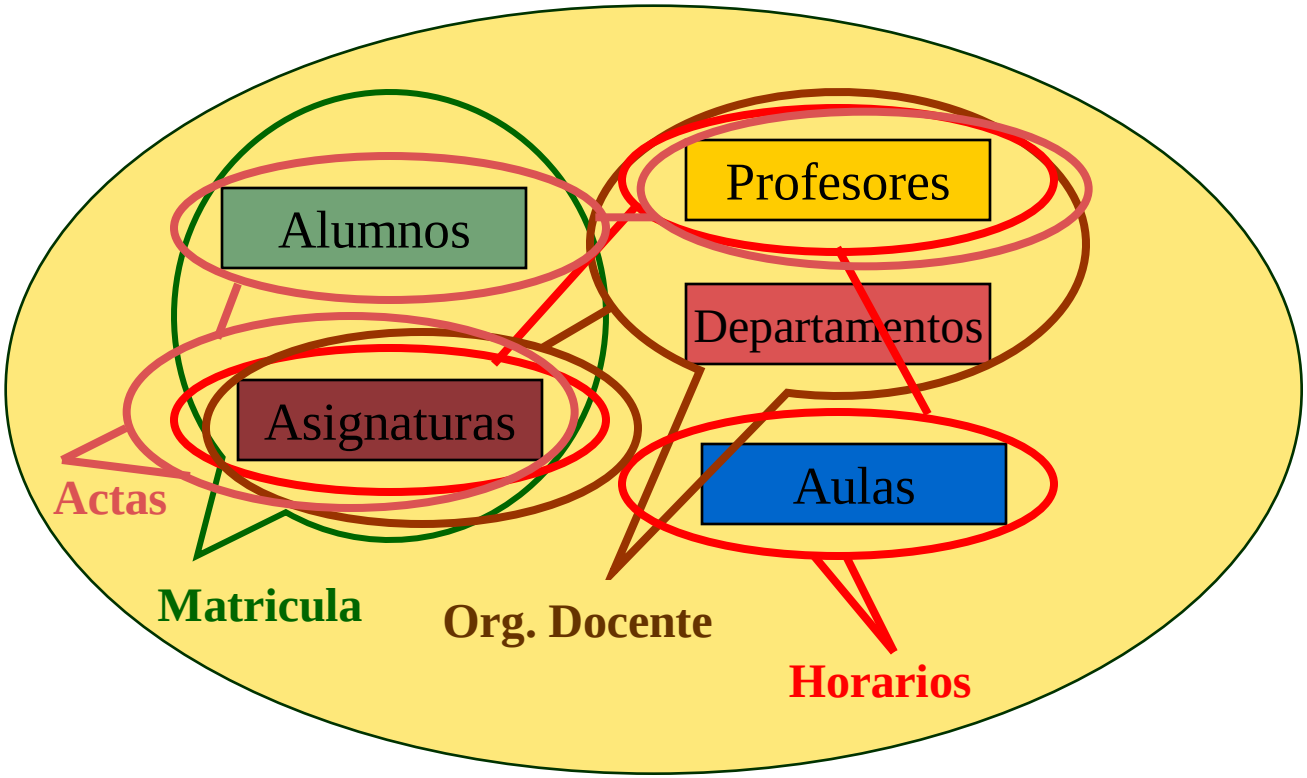
1.1 y 1.2 Introducción y Concepto intuitivo de base de datos.

- Ejemplo: Datos referentes a organización docente
- Aplicaciones posibles:
  - Matricula: alumnos, asignaturas, calificaciones.
  - Organización docente: profesores, asignaturas, departamentos.
  - Actas: asignaturas, profesores, alumnos.
  - Horarios: profesores, asignaturas, aulas.
- Problemas de una estructura basada en archivos de registros: *los datos pertenecen a las aplicaciones*
  - Redundancia. Datos repetidos en archivos distintos.
  - Inconsistencia. Motivadas por la redundancia, aparecen incoherencias.
  - No reusabilidad. Fuerte acoplamiento entre la estructura de datos y la aplicación.
  - Seguridad en el acceso a los datos.

1.1 y 1.2 Introducción y Concepto intuitivo de base de datos.

Concepto intuitivo de base de datos

- Una alternativa:



Tema 1: Introducción

MOTIVACION

- **Complejidad de un sistema de archivos para:**
  - Crear un archivo con una estructura determinada.
  - Consultar o actualizar un archivo imponiendo diferentes condiciones.
  - Modificar dinámicamente la estructura de un archivo.
  - Proteger ciertas filas o columnas de usuarios no autorizados.
  - Permitir el acceso desde distintas aplicaciones, con distintos lenguajes y sistemas operativos.
- **Solución:**

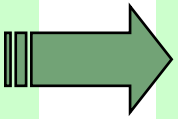
Utilizar un sistema de gestión de BD

1.3 DATO OPERATIVO

Definición: Pieza de información que necesita una empresa para su funcionamiento.

- ITEM BÁSICO: Elementos acerca de los que se puede pedir información (sustantivos).
- ATRIBUTOS: Características de los items básicos (adjetivos).
- RELACIONES, VINCULOS o CONEXIONES: Conexiones lógicas entre items.

Datos Operativos Gestión Docente:
Asignaturas
Profesores
Aulas
Asignaturas que imparte cada profesor



Esquema Lógico Gestión Docente:
Item básico 'Asignatura'
Atributos: Nombre, Créditos
Item básico: 'Profesor'
Atributos: Nombre, Dni
Item básico: 'Aula'
Vinculo entre 'Asignatura' y 'Profesor'
Atributo: Curso

1.3 DATO OPERATIVO

Cuando se determinan y clasifican de esta forma todos los datos operativos, se obtiene el ESQUEMA LÓGICO de la Base de Datos.

Para obtener el Esq. Lógico:

I. Identificar todos los items básicos (profesores, asignaturas, grupos, aulas,..)
II. Identificar todos los vínculos entre ellos (profesor imparte grupo, grupo pertenece a asignatura,..)
III. Caracterizar items y vínculos con los correspondientes atributos.

El Esquema Lógico representa la estructura de datos de la empresa expresada en términos conceptuales.

1.4 CONCEPTO DE BBDD Y DE SGBD

Compartir información entre varios usuarios

Persistente en soporte informático

Concepto intuitivo de base de datos

- *Fondo **común** de información **almacenada en una computadora** para que cualquier persona o programa **autorizado pueda acceder a ella**, **independientemente del lugar de su procedencia y del uso que haga**.*

Acceso Restringido

Procesamiento de los datos ajeno a la propia BBDD

OPERACIONES:

Estructura de Datos.	*MODIFICAR*
Información.	*INSERTAR*
	BORRAR
	CONSULTAR

1.4 CONCEPTO DE BBDD Y DE SGBD

Definiciones iniciales:

- Base de datos:

  Definición conceptual: Está constituida por una instancia de un esquema lógico junto con las instancias de los datos operativos que dicho esquema organiza.

  Conjunto de datos comunes a un "proyecto" almacenados sin redundancia para ser útiles a diferentes aplicaciones.

- Sistema de Gestión de Bases de Datos (DBMS ó SGBD):

  Conjunto de elementos software con capacidad para definir, mantener y utilizar una base de datos.

  Un sistema de gestión de bases de datos debe permitir:

  1. *Tratamiento de estructuras de almacenamiento (crear, modificar, eliminar y recuperar).*
  2. *Acceder a los datos de forma eficiente y segura.*
  3. *Organizar el acceso multiusuario, gestión de permisos, concurrencia,…*

1.5. ELEMENTOS DE UNA BD

- **Datos:**
  - Integrados (sin redundancia) y organizados mediante un esquema lógico.
  - Compartidos (útiles a varias aplicaciones).
  - Control de acceso.
- **Software**
  Programas de Aplicación ⟺ **SGBD.** Gestionar la estructura y la información de la BD.
- **Hardware:**
  - BD centralizada (cliente/servidor).
  - BD distribuida (varios niveles de ejecución, serv. de aplicaciones).
- **Usuarios.** Objetivo: Satisfacer las necesiadades del usuario.
  - Usuario terminal. Directamente o mediante prog. de aplicación.
  - Programador de aplicaciones. Desarrolla prog. de aplicación sobre la BD para uso de los usuarios terminales. Interactúa con el esquema lógico que le concierne.
  - Administrador (DBA ó DBM).Gestionar los componentes del SGBD. Garantizar operatividad y seguridad.

1.6. INDEPENDENCIA DE DATOS

Independencia:

Los datos se organizan independientemente de las aplicaciones que los vayan a usar y de los ficheros en los que vayan a almacenarse.

- Independencia Física:

El almacenamiento físico de los datos debe ser independiente del diseño lógico de la BD a todos los niveles.

Representación de campos
Organización en registros
Organización en archivos
Mecanismos de acceso
Distribución física

Esto permite:

- Realizar cambios en estructura física:

Optimizar recursos
Cambios hardware
Optimización accesos

- Descargar a las aplicaciones de problemas físicos

1.6. INDEPENDENCIA DE DATOS

- Independencia Lógica:

Existen dos tipos de estructuras lógicas:

Esquema lógico general
Vistas de usuario

Cada aplicación debe poder organizar los datos según sus propios esquemas y acceder a los datos que le son necesarios. Las vistas de usuarios permanecerán inmutables a cambios en la estructura lógica de la BD.

↓

Aumento de seguridad y fiabilidad.

Menos problemas para las aplicaciones.

Posibilidad de cambios en los esquemas por parte de las aplicaciones y por parte de los administradores.

1.7. OBJETIVOS DE UN SGBD

1. Diseño y utilización orientada al usuario:

Los datos y aplicaciones deben ser accesibles a los usuarios de la manera más amigable posible

- **Soportar un modelo de datos teórico.**
- **Soportar facilidades de definición.**
- **Soportar lenguajes de acceso y modificación.**

2. Centralización:

Los datos deben gestionarse de forma centralizada e independiente de las aplicaciones.

↓

Administrador de la Base de datos

Utilidades de administración

1.7. OBJETIVOS DE UN SGBD

3. No redundancia:

   Los datos no deben estar duplicados

   → Gestión de accesos concurrentes

4. Consistencia (Integridad semántica): Evitar errores humanos

   Los datos deben ser consistentes (sin fallos lógicos)

   → Mecanismos de mantenimiento de integridad

5. Fiabilidad: Control y recuperación de errores

   Los datos deben estar protegidos contra fallos catastróficos

   → Mecanismos de mantenimiento de recuperación y relanzamiento de transacciones

1.7. OBJETIVOS DE UN SGBD

6. Seguridad. Evitar alteraciones no autorizadas.

No todos los datos deben ser accesibles a todos los usuarios

- Mecanismos de gestión de usuarios y privilegios.
- Mecanismos de protección de información basado en identificación de usuarios.

1.8. VENTAJAS EN LA UTILIZACIÓN DE UN SGBD

- **Frente al usuario:**

  Usuario terminal. Estructura base común, accesible y reutilizable por diferentes usuarios y aplicaciones.

  Programador de aplicaciones. elimina problemas:

  - Depuración de errores.
  - Independencia Lógica y Física → Evita reescribir programas.
  - Mantenimiento en general (copias de seguridad, recuperación de fallos, etc..)

  Administrador de B.D. no existiría.

- **Frente al sistema:**

  - Control centralizado: fiabilidad, consistencia, seguridad.
  - Escalabilidad: A nivel de capacidad de procesamiento y rendimiento.
  - Criterios de asignación de recursos.
  - Generación de nuevas aplicaciones.
  - Equilibrio entre requerimientos conflictivos.

Resumen Histórico

- **Prehistoria: ficheros y sistemas de acceso (1950-65):**
  - **IBM, Honeywell y General Electric (Finales de los 50 y 60): Productos basados en la definicion de datos.**
  - **1960. COBOL. Desarrollado por el grupo CODASYL. Todo programa escrito en Cobol tiene una parte (DATA DIVISION) donde se describen los datos y ficheros que se van a utilizar.**
  - **1965. Se mejora para permitir trabajar con estructuras de datos complejas.**
  - **1969. DBTG (subgrupo de CODASYL) introduce DDL y DML para SGBD en red. Modelo de datos basado en una red con tipos de registro como nodos y las relaciones uno-a-muchos como arcos.**

    Precedentes: Principios de los 60. IDS (General Electric).

    Otros sistemas basados en modelo en red: IDMS de Cullinet (1970), DMS 1100 de Sperry (1971) e IDS-2 de Honeywell (1975).
  - **IMS de IBM. Basado en modelo jerárquico. (Estructura arbórea).**

Estos modelos (navegacionales) presentan el problema de carecer de independencia física.

Resumen Histórico

- **Desarrollo del modelo relacional (1970-1990):**
  - **(1970-78) Desarrollo del modelo teórico**
    - **Definición del modelo relacional (E.F. Codd 1970-72)**
    - **Basado en Algebra y Calculo relacional, calculo de predicados y teoría de conjuntos.**
    - **Primeros problemas de diseño**
  - **(1975- 1980) Desarrollo de los primeros SGBD relacionales**
    - **SQL (1975) a partir del modelo teórico.**
    - **System/R de IBM (1976).**
    - **ORACLE, DB2, Ingres Universidad de Berkley, Informix (integrado a posteriori por IBM), SQL Server de Microsoft.**
  - **(1980-199-) Desarrollo de los grandes sistemas relacionales**
    - **SQL comercial, desarrollo de generadores de aplicaciones.**
    - **Sistemas distribuidos, estructuras cliente/servidor (lenguajes visuales)**

El éxito de este modelo reside en que la representación y el acceso a la información están basados en criterios semánticos y no en estructurales.

Mediante SQL defino qué datos busco y no como acceder a lo que busco.

Existe un único mecanismo para representar la información: la relación.

Resumen Histórico

- **Nuevos modelos de representación de información :**
- **Algunos problemas en el modelo relacional**
  - **Problemas en la representación de estructuras complejas.**

Bases de datos orientadas a objetos. Surgen a fines de los 80. GemStone, Gbase y Vbase. La complejidad en la consulta y la falta de un lenguaje estandar los ha relegado a ámbitos específicos: CAD, VLSI, multimedia,...

  - **Evolución del modelo relacional a sistemas de bases de datos objeto-relacionales. Añade características de la orientación a objetos sin renunciar a la compatibilidad con el modelo tradicional.**
    - **Ejemplos: Oracle (a partir de la V.9), DB2 Universal Database de IBM y PostgreSQL.**

Resumen Histórico

Evolución actual:

- No solo manejar información, incorporando el manejo de conocimiento mediante las siguinetes técnicas:
- Bases de datos deductivas.
- Técnicas para extracción de conocimiento de bases de datos (Data Mining).
- Tratamiento semántico de la información.
- Representación y manipulación de la imprecisión y de la incertidumbre en bases de datos.
- Estandares para representación e intercambio de datos como XML.
- Arquitecturas redundantes y escalables (clusters y grid de computadores) que buscan alta disponibilidad y crecimiento en la capacidad de procesamiento de los SGBDs.