

Introducción

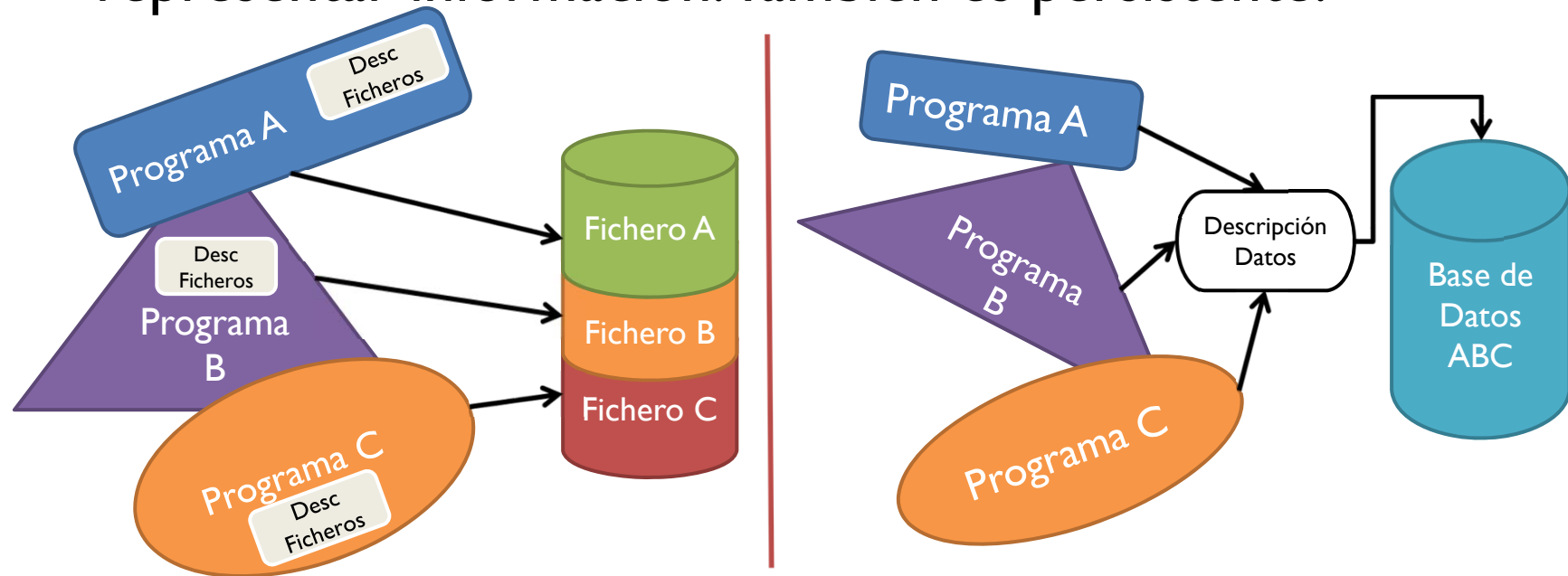
BASES DE DATOS

Profesor: Héctor Gómez Gauchía

**Materiales: Mercedes García Merayo
Luis Garmendia**

Ficheros vs. Bases de datos

- ▶ **Fichero**: almacenamiento persistente de datos usado para representar información: Ficheros de texto, documentos XML....
- ▶ **Base de datos**: colección de datos relacionados usada para representar información. También es persistente.



Sistemas de Archivos

- ▶ Información almacenada en archivos del sistema operativo.
- ▶ Usuarios manipulan la información mediante programas de aplicación
 - ▶ Un programa para actualizar los datos
 - ▶ Un programa para añadir información
 - ▶ Un programa para generar informes...
- ▶ Estos programas de aplicación se desarrollan en respuesta a las necesidades.
 - ▶ Nuevos archivos
 - ▶ Nuevos programas de aplicación

Sistemas de Archivos

- ▶ **Redundancia e inconsistencia de datos.**
 - ▶ Archivos y programas de aplicación creados por diferentes programadores en diferentes formatos y lenguajes.
 - ▶ Información puede estar **duplicada** en diferentes archivos.
 - ▶ **Inconsistencia** de datos: actualizaciones de ficheros no sincronizadas.
- ▶ **Dificultad en el acceso a los datos.**
 - ▶ Nuevas necesidades de acceso requieren nuevos desarrollos
- ▶ **Aislamiento de datos.**
 - ▶ Diferentes archivos con diferentes formatos.

Sistemas de Archivos

- ▶ **Problemas de integridad**

- ▶ Restricciones de consistencia controladas en diferentes programas de aplicación. Difícil actualización o inclusión de nuevas restricciones que afectan a diferentes archivos.

- ▶ **Problemas de atomicidad**

- ▶ Transacciones que afectan a más de un archivo requieren asegurar la consistencia. Difícil en sistemas de archivos.

- ▶ **Problemas de acceso concurrente**

- ▶ Acceso a los datos desde diferentes programas de aplicación que no han sido coordinados.

- ▶ **Seguridad**

- ▶ Diferentes niveles de acceso a los datos

Propiedades Bases de Datos: ACID



Volveremos durante el curso a ver estas propiedades

- ▶ **Atomicidad.** *Atomicity*
 - ▶ Todas las operaciones de la transacción se realizan adecuadamente en la base de datos o ninguna de ellas.
- ▶ **Consistencia.** *Consistency*
 - ▶ La ejecución aislada de la transacción conserva la consistencia de la base de datos.
- ▶ **Aislamiento.** *Isolation*
 - ▶ Aunque se ejecuten varias transacciones concurrentemente, el sistema garantiza que para cada par de transacciones T_i y T_j , se cumple que para los efectos de T_i , o bien T_j ha terminado su ejecución antes de que comience T_i , o bien que T_j ha comenzado su ejecución después de que T_i termine.
- ▶ **Durabilidad.** *Durability*
 - ▶ Tras la finalización con éxito de una transacción, los cambios realizados en la base de datos

Sistemas de Archivos

- ▶ Los Sistemas Gestores de Bases de Datos ofrecen una solución integral a todos estos problemas
- ▶ Los ficheros se prefieren a los SGBD cuando:
 - ▶ La base de datos considerada y las aplicaciones asociadas son simples, bien definidas y sin visos de cambio.
 - ▶ Requisitos de tiempo real que no pueden cubrir los SGBD.
 - ▶ No se requiere acceso concurrente.

¿Qué es una Base de Datos?

“Colección de datos → Ej: Universidad

... relacionados entre sí por su significado

... que modelan el mundo real (minimundo)

... con un propósito específico”

- ▶ Datos: Hechos conocidos que se pueden grabar y que tienen un significado implícito.
 - ▶ Agenda telefónica
 - ▶ Citas consulta médica
 - ▶ Referencias bibliográficas...

Ejemplo de Base de Datos **Universidad**

- ▶ **Propósito.** Mantener información de estudiantes, cursos y calificaciones.
- ▶ **Minimundo.** Diferentes aspectos de un entorno universitario
- ▶ **Entidades** (Elementos). Estudiantes, asignaturas, departamentos, profesores, calificaciones, requisitos...
- ▶ **Relaciones entre los elementos del minimundo (ó universo de discurso).**
 - ▶ Estudiantes se matriculan de una o varias asignaturas
 - ▶ Cada asignatura es ofertada por un departamento
 - ▶ Las asignaturas pueden tener ciertos requisitos.
 - ▶ Los profesores pertenecen a un único departamento

Ejemplo de Base de Datos

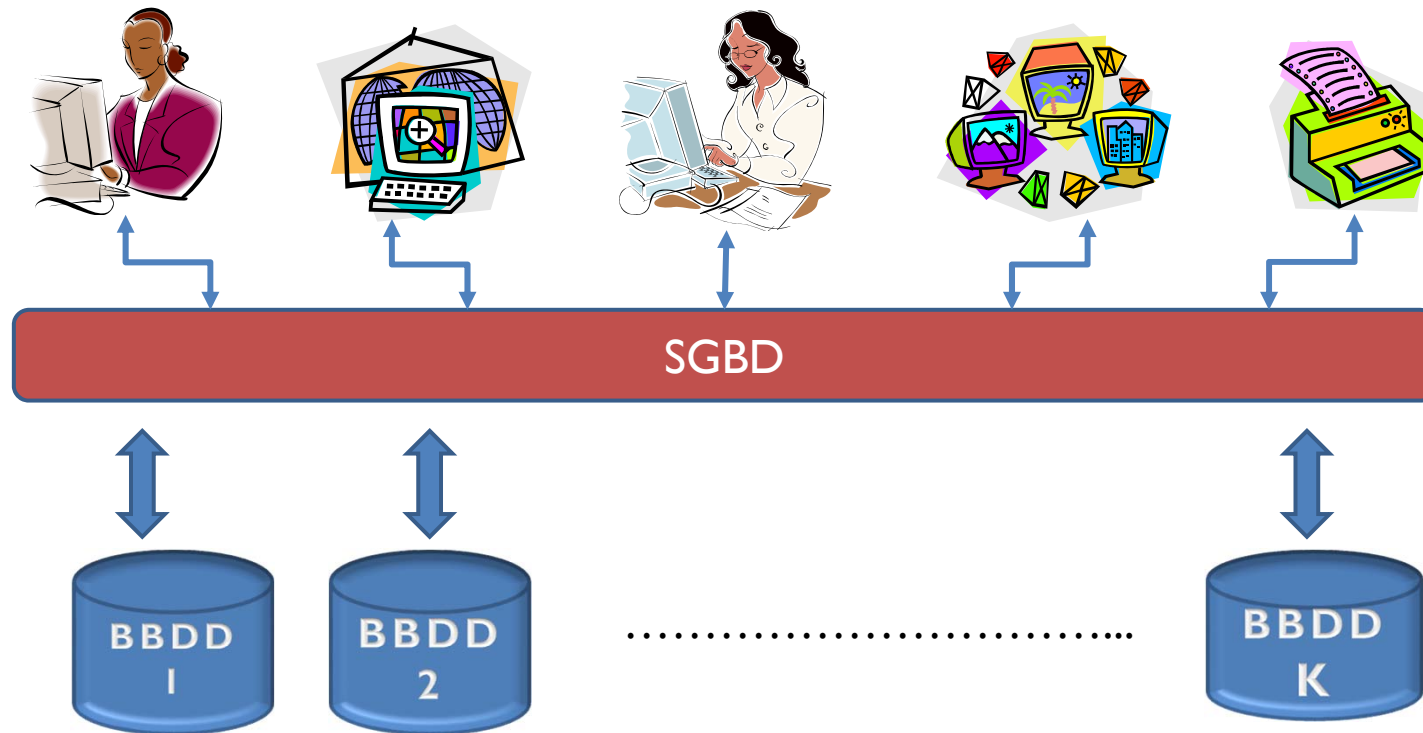
- ▶ La manipulación de la base de datos implica la consulta, creación, borrado y actualización ...
 - ▶ Listado de las asignaturas de las que se ha matriculado Luis López
 - ▶ Alumnos que se han matriculado en la asignatura de Bases de Datos
 - ▶ Nota media de la asignatura Teoría de Autómatas en la convocatoria de Septiembre
 - ▶ Cambiar la nota de Luis en la asignatura de Lógica Formal
 - ▶ Anular la matricula de Ana Álvarez
 - ▶ Dar de alta una nueva asignatura en un departamento

Sistema Gestor de Base de Datos

- ▶ Una base de datos debe reflejar los cambios que se producen en el “*minimundo*” (*aspecto del mundo real*) que representa.
- ▶ Sistema Gestor de Base de Datos (SGBD)
 - ▶ Colección de programas que permite a los usuarios crear y mantener una base de datos. Facilita las tareas de definición, construcción, manipulación y compartición de las bases de datos entre varios usuarios y aplicaciones.
- ▶ Sistema de base de datos: Base de Datos + SGBD

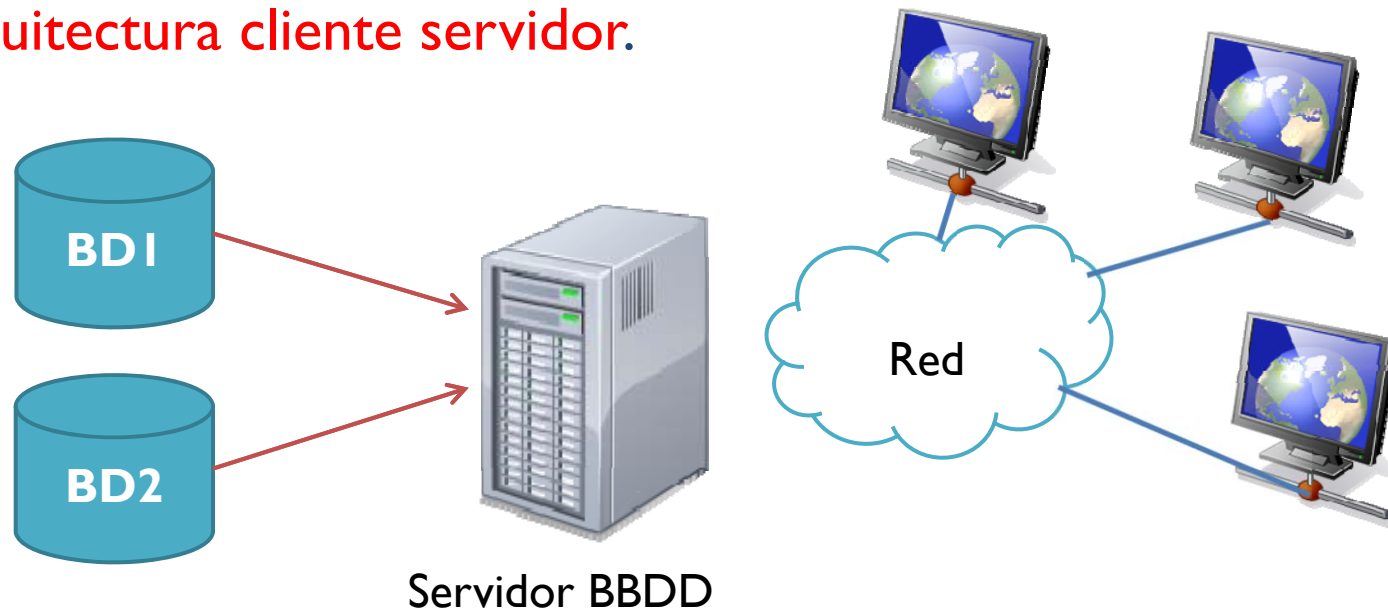
Sistema Gestor de Base de Datos

- Un SGBD es un sistema software capaz de manejar grandes colecciones de datos relacionados, compartidos de forma consistente y segura.



Sistema Gestor de Bases de Datos

- ▶ Tienen como objetivo proporcionar un entorno práctico y eficiente para almacenamiento y recuperación de datos.
 - ▶ Debe garantizar fiabilidad ante caídas del sistema.
 - ▶ Congruencia de los datos ante accesos simultáneos.
 - ▶ Habitualmente están implementados mediante con una **arquitectura cliente servidor**.



Comportamiento del SGBD con los Datos

- ▶ Un SGBD es un sistema software capaz de manejar grandes colecciones de datos relacionados, compartidos de forma consistente y segura.
- ▶ Independencia de datos. Proporciona una vista abstracta de los datos, omitiendo detalles de representación y almacenamiento.
- ▶ Integridad y seguridad de los datos. Un SGBD permite establecer restricciones de integridad en la manipulación de los datos y controles de acceso para diferentes tipos de usuarios.
- ▶ Acceso eficiente a los datos mediante el uso de diferentes técnicas de almacenamiento y recuperación.
- ▶ Centralización de la administración de datos. Expertos que establecen la organización de los datos, los grupos de usuarios...
- ▶ Acceso concurrente a los datos de forma transparente para los usuarios.



Pasos en el Diseño de una Base de Datos

- ▶ 1.- Recopilación y análisis de requisitos.

Interactuar con los expertos del dominio para caracterizar los requisitos de los datos de los usuarios de la base de datos.

- ▶ 2.- Diseño Conceptual (Modelo Conceptual).

Traducción de los requisitos a un *modelo de datos*.

- ▶ Descripción de la estructura de la base de datos mediante los conceptos del modelo elegido.
- ▶ La estructura son los tipos de datos, las relaciones y las restricciones.
- ▶ No se especifican detalles de almacenamiento físico.

Pasos en el Diseño de una Base de Datos

- ▶ 3.- Diseño Lógico (Modelo Lógico).

Se traduce el esquema conceptual de la base de datos al modelo de implementación de datos del SGBD elegido.

- ▶ 3.1.- Normalización de relaciones.

Análisis de las relaciones para detectar posibles problemas.

- ▶ 4.- Diseño Físico (Modelo Interno)

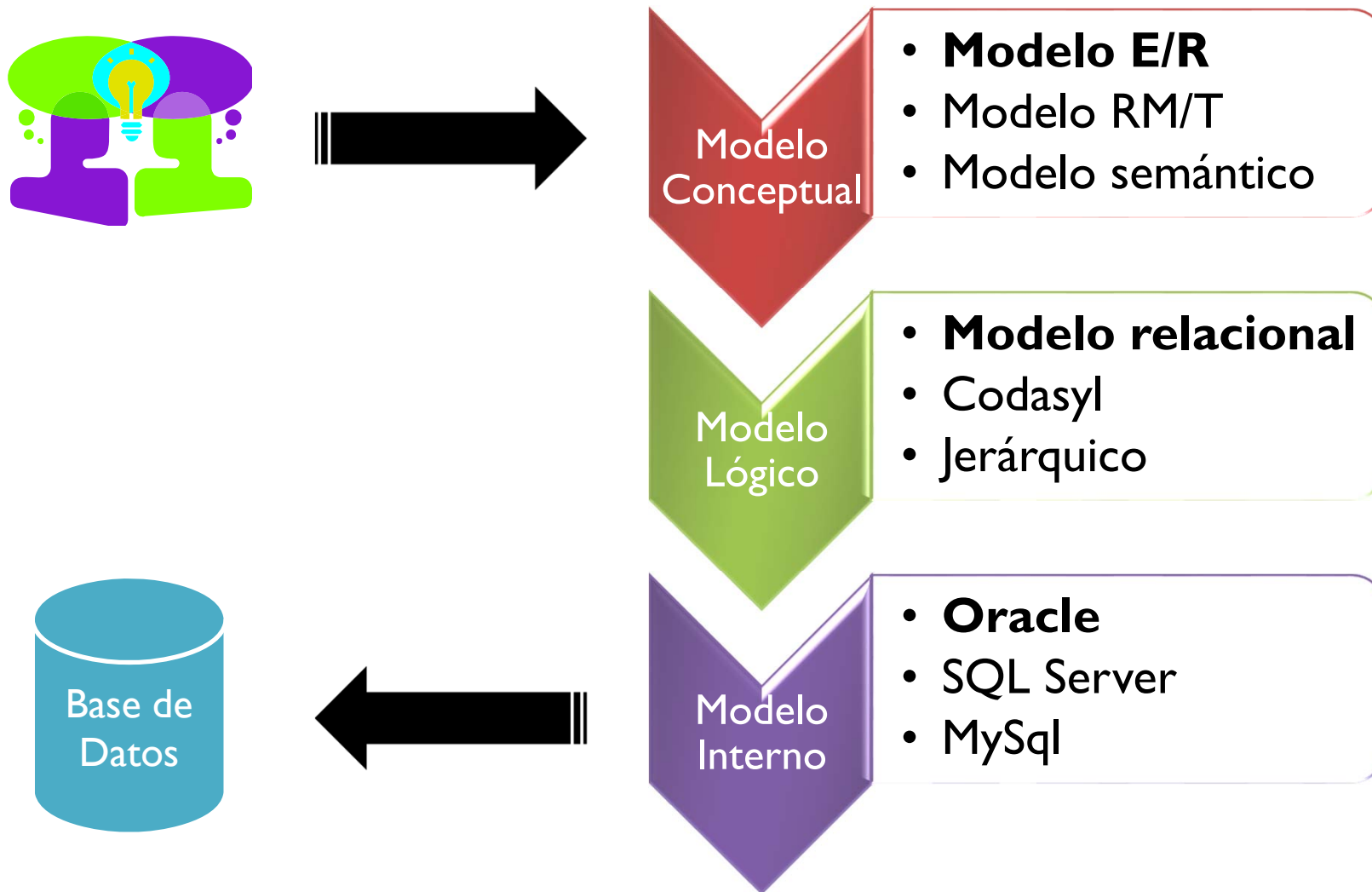
Refinamiento del diseño de la base de datos para alcanzar un rendimiento óptimo (organización de archivos, índices...)

- ▶ 5.- Diseño de Seguridad.

Descripción de roles y accesos a los diferentes elementos de la base de datos.



Diseño de una Base de Datos



Modelo de Datos Entidad-Relación

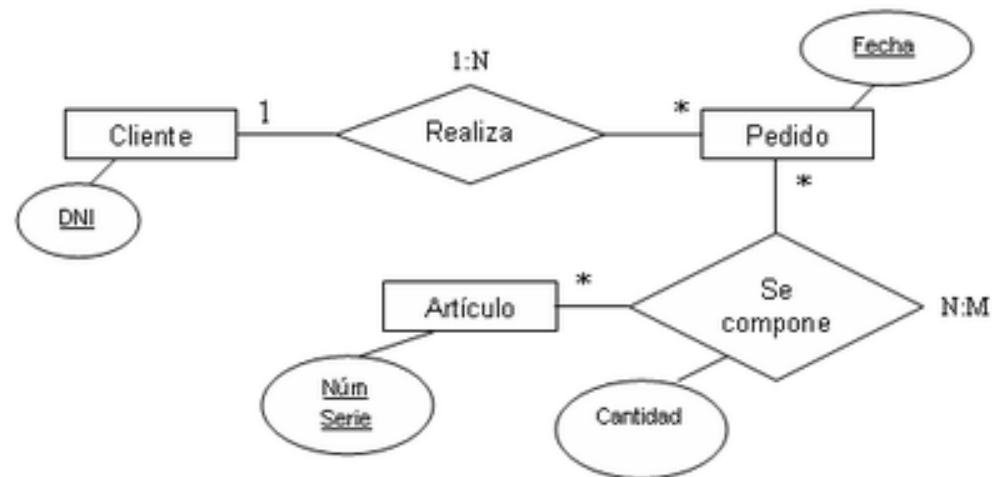
- ▶ Existen diferentes modelos de datos, pero el más utilizado por su sencillez y eficiencia es el **modelo Entidad-Relación**.
- ▶ Este modelo representa la realidad a través de **entidades**, que son objetos que existen y que se distinguen de otros por sus características.
 - ▶ *Un alumno se distingue de otro por sus características particulares como lo es el nombre, o el numero de control asignado al entrar a una institución educativa...*
- ▶ Las características de las entidades se llaman **atributos**.
 - ▶ *Nombre, dirección, teléfono, grado, grupo, etc. son atributos de la entidad alumno.*
- ▶ A su vez una entidad se puede asociar o relacionar con más entidades a través de **relaciones**.





Volveremos durante el curso a ver el resto de transpas

Modelo de Datos Entidad-Relación



✖ Volveremos durante el curso a ver el resto de transpas

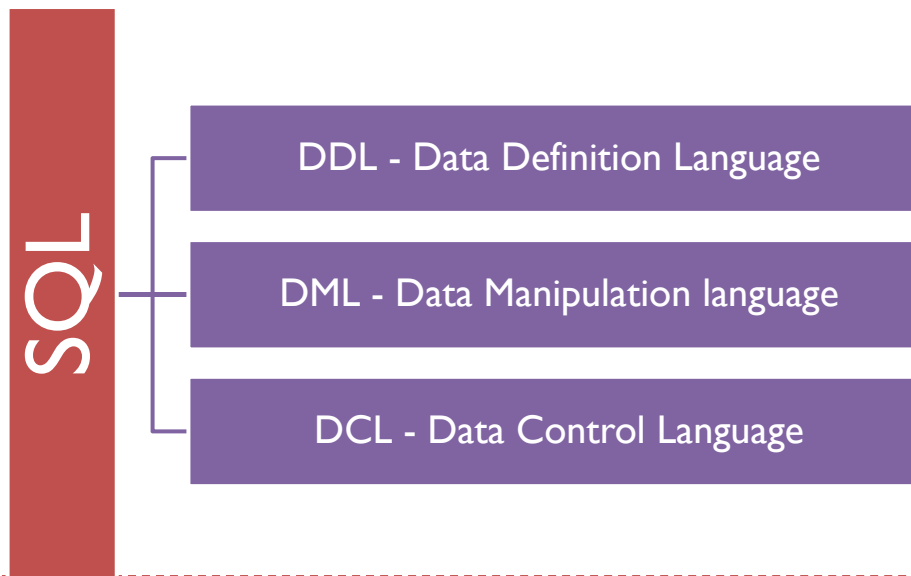
Modelo Relacional

- ▶ Codd propuso en 1970 el **modelo relacional**.
 - ▶ *Esquema de una Base de Datos*: diseño completo creado por el diseñador
 - ▶ *Instancia de una Base de Datos*: información almacenada en un instante
 - ▶ Una base de datos es una colección de **relaciones**
 - ▶ Cada relación es una tabla con filas y columnas

Dni	Apellidos	Nombre	Dirección	CC
08932583P	López Amor	María	Calle Loma 2	004024090248091245
74657235U	García Mancha	Luis	Plaza Retamar 4	857345784683639848
37651947T	Pozo Mares	Fernando	Calle Fernando el Santo 56	765736572365473483
78538465G	Salinas Pérez	Antonio	Calle Ferrocarril 14	376581248912481736

Structured Query Language (SQL)

- ▶ *Lenguaje de definición de datos (DDL)* para especificar el esquema de la base de datos
- ▶ *Lenguaje de manipulación de datos (DML)* para expresar las consultas y modificaciones de la base de datos.
- ▶ *Lenguaje de Control de Datos (DCL)* permite controlar el acceso a los datos contenidos en la base de datos.



Lenguaje de definición de datos (DDL)

- ▶ Los lenguajes de definición de datos sirven para crear y definir los esquemas de la BD.
- ▶ El SGBD utiliza un conjunto de tablas denominado **diccionario de datos** para almacenar:
 - ▶ Esquema de la base de datos
 - ▶ Estructura de almacenamiento y los métodos de acceso utilizados
 - ▶ Restricciones de integridad
 - Restricciones de dominio
 - Integridad referencial
 - Aserciones
 - ▶ Autorización

Ejemplo DDL

```
CREATE TABLE CITIES (  
    CITY_ID      CHAR(6) NOT NULL,  
    CITY_NAME    CHAR(2) NOT NULL,  
    COUNTRY_CODE CHAR(6) NOT NULL,  
    MILES_TO_LA  SMALLINT,  
    POPULATION   INT,  
    HEMISPHERE   CHAR(1),  
    AVG_INCOME   DECIMAL(9,2),  
    LAST_CENSUS  DATE,  
    LAST_UPDATED TIMESTAMP NOT NULL WITH DEFAULT,  
    CITY_DESCRIPT VARCHAR(300),  
    PRIMARY KEY (CITY_ID),  
    FOREIGN KEY (COUNTRY_CODE) REFERENCES COUNTRY  
);
```


Ejemplo DDL

```
ALTER TABLE CITIES
ADD CONSTRAINT POP_MILES CHECK (
    POPULATION BETWEEN 1 AND 50000000
    AND MILES_TO_LA >=-1 AND
    MILES_TO_LA <=13500
);

DROP TABLE CITIES;
```

Lenguajes de Manipulación de Datos (DML)

- ▶ Los lenguajes de manipulación de datos sirven para la inserción (INSERT), eliminación (DELETE), consulta (SELECT) y modificación (UPDATE) de la información de la base de datos.
- ▶ Lo utilizan los usuarios finales, los programas de aplicación y cualquiera que lo necesite.

```
SELECT NOMBRE, APES FROM ALUMNOS  
WHERE CALIF > 5;
```

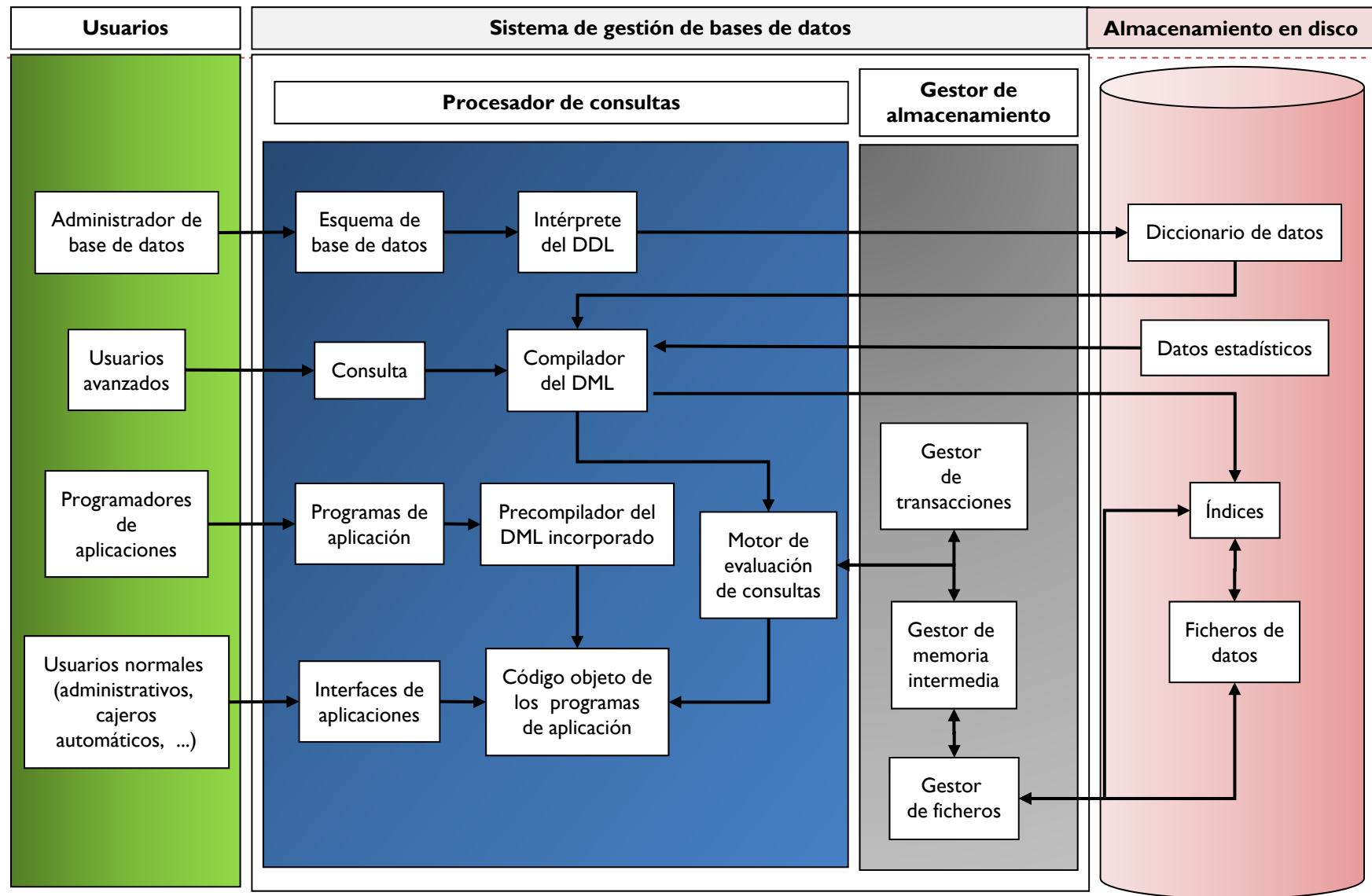
Lenguaje de Control de Datos (DCL)

- ▶ El lenguaje de control de datos permite a los administradores de bases de datos
 - ▶ Imponer la seguridad de acceso a datos restringiendo el acceso a usuarios y grupos.
 - ▶ Definir las operaciones admitidas a cada usuario/rol.

```
GRANT
    SELECT,
    UPDATE (SALARY)
ON BLUEDEP
TO CLERK;
```

```
REVOKE
    SELECT,
    UPDATE (LASTNAME, FIRSTNAME)
ON TABLE EMPL
FROM PETER;
```

Estructura Global de un SGBD



Usuarios de bases de datos

- ▶ **Usuarios normales**

Usan la BD a través de aplicaciones.

- ▶ **Programadores de aplicaciones.**

Interaccionan con la BD a través de llamadas que usan DML.

- ▶ **Usuarios avanzados.**

Consultas en un lenguaje de consultas de bases de datos.

- ▶ **Usuarios especializados**

Escriben aplicaciones de bases de datos especializadas.

Usuarios de bases de datos

- ▶ Administrador de la base de datos.
 - ▶ Definición del esquema.
 - ▶ Definición de la estructura de almacenamiento y del método de acceso.
 - ▶ Modificación del esquema y la organización física.
 - ▶ Concesión de derechos de acceso.
 - ▶ Mantenimiento

Gestión de transacciones y concurrencia

- ▶ Una **transacción** es una colección de operaciones que lleva a cabo una única función lógica en una aplicación de base de datos.
 - ▶ Unidad de atomicidad (todo o nada) y consistencia
 - ▶ Unidad de trabajo
- ▶ El gestor de transacciones asegura que la BD permanece en un estado consistente a pesar de fallos del sistema
- ▶ El gestor de control de **concurrencia** controla la interacción entre las transacciones concurrentes para garantizar la consistencia de la base de datos.

Gestor de Almacenamiento

- ▶ Módulo de programa que proporciona la interfaz entre los datos de bajo nivel de la BD y los programas de aplicación y las consultas remitidas al sistema.
- ▶ Es responsable de las siguientes tareas:
 - ▶ Interacción con el gestor de archivos
 - ▶ Almacenamiento, recuperación y actualización de los datos.
- ▶ Componentes:
 - ▶ Gestor de autorizaciones e integridad
 - ▶ Gestor de archivos
 - ▶ Gestor de la memoria intermedia

Procesamiento de consultas

- ▶ Traduce las actualizaciones y las consultas, en el nivel lógico, en una secuencia eficiente de operaciones en el nivel físico
- ▶ Intérprete del DDL
 - ▶ Registra las definiciones en el diccionario de datos
- ▶ Compilador del DML
 - ▶ Traducción de instrucciones DML a un plan de evaluación (puede haber varios) que consiste en instrucciones de bajo nivel que entiende el motor de evaluación de consultas
 - ▶ Optimización de consultas: plan de evaluación menor coste
- ▶ Motor de evaluación de consultas
 - ▶ Ejecuta las instrucciones de bajo nivel generadas por el compilador de DML

Evolución histórica

50-60	<ul style="list-style-type: none">• Cintas magnéticas acceso secuencial• Tarjetas perforadas.
60-70	<ul style="list-style-type: none">• Discos magnéticos acceso directo a los datos• <i>Bases de datos jerárquicas y en red</i>• 1970: Codd define modelo relacional
80-90	<ul style="list-style-type: none">• Prototipos de bases de datos relacionales evolucionan hacia sistemas comerciales (System R => SQL / DS (IBM))• SQL se convierte en el estándar industrial• Sistemas de <i>bases de datos paralelos y distribuidos</i>• Sistemas de <i>bases de datos orientados a objetos</i>
90	<ul style="list-style-type: none">• Aplicaciones de <i>minería de datos y de toma de decisiones</i>• Aparición del comercio electrónico en la Web
2000	<ul style="list-style-type: none">• XML y XQuery como tecnologías de bases de datos• Administración de bases de datos automática