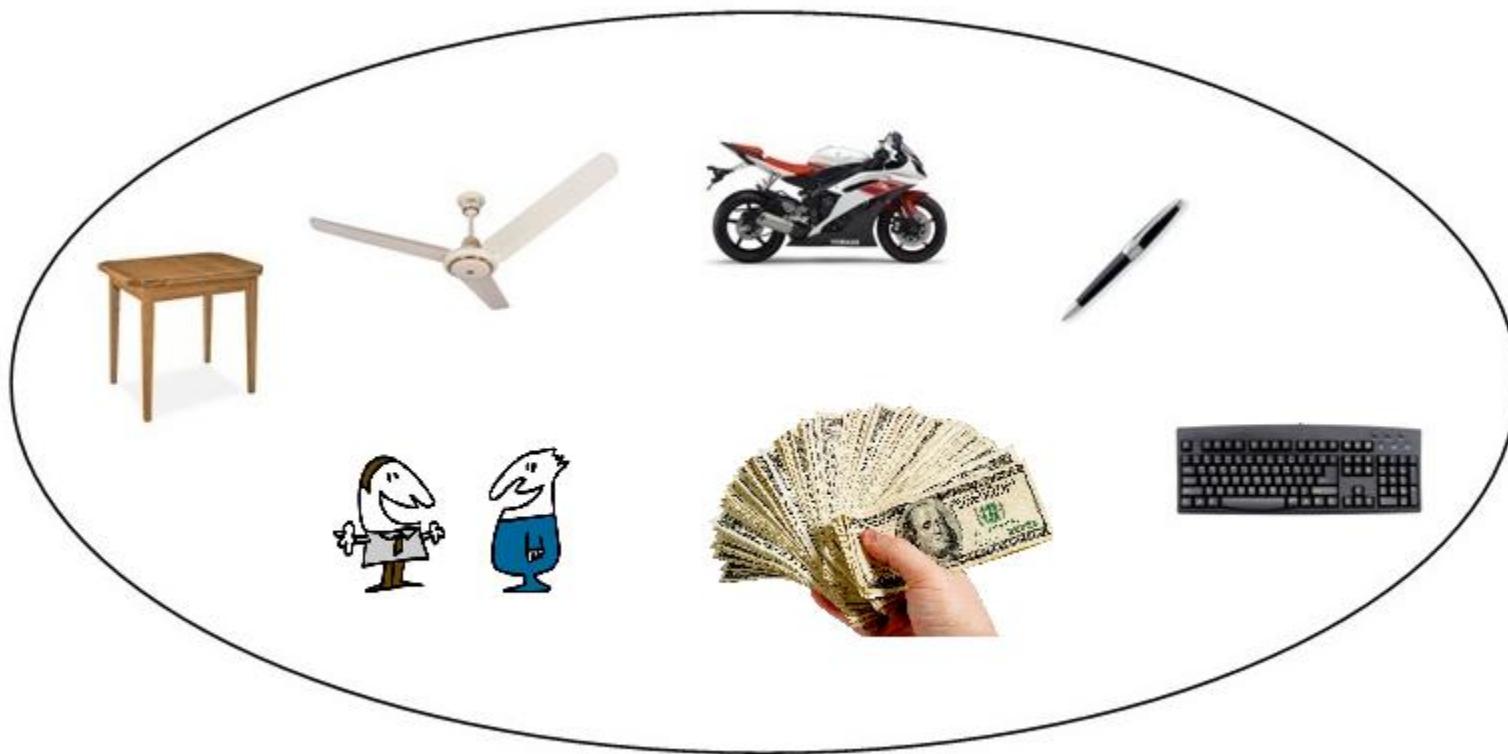


Bases de Datos I

Introducción a las bases de datos
Conceptos básicos

¿Qué es una base de datos?

- Conjunto de datos relacionados/relevantes a aspectos del mundo real.



SABD (Sistema Administrador de Bases de Datos)

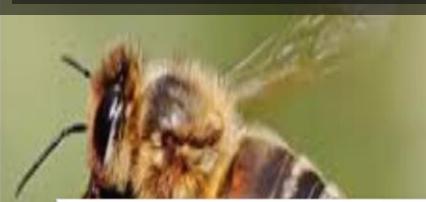
- Conjunto de programas que permiten a los usuarios crear y dar mantenimiento a las bases de datos.
- Es el software que facilita la comunicación de los usuarios con la BD por medio de un lenguaje de consultas.
- Garantiza la integridad y seguridad de los datos, así como la recuperación de la base de datos en caso de fallas.

SABD

Crear



Comunicación



Seguridad



Mantener



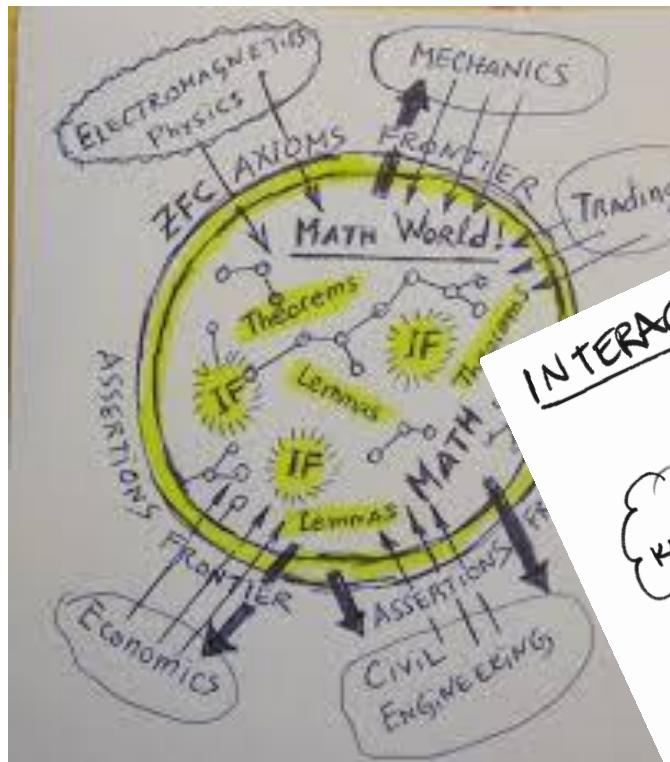
SQL



Integridad



SABD: Software que facilita procesos



Definición



Interacción



Construcción



Manipulación

SABD: Definición

- Tipos de datos, estructuras, restricciones, catálogo o diccionario de datos metadatos.



SABD: Construcción

- Almacena los datos en un medio controlado por el SABD.



SABD: Manipulación

- Actualización de la BD para reflejar cambios en el universo de la base de datos.



SABD: Interacción

- Permite que varios usuarios y programas accedan a la base de datos de forma simultánea.

INTERACTION DESIGN



Funciones del SABD

- Proteger:
 - Funcionamiento defectuoso del hardware o software (caídas).
 - Acceso no autorizado o malintencionado.
- Mantener:
 - Evolución según cambian las necesidades en el tiempo.
 - Las necesidades cambian con el tiempo.

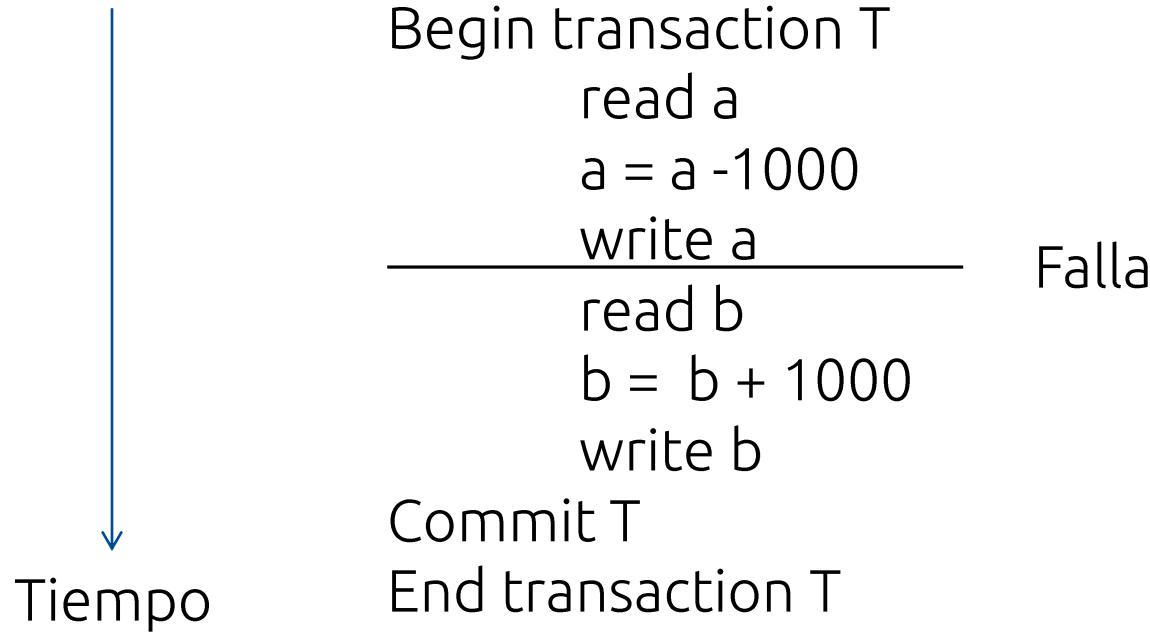
Otras funcionalidades del SABD

- Optimización:
 - Rapidez en el retorno/almacenamiento de información.
- Control de concurrencia:
 - Acceso a información de varios usuarios al mismo tiempo.
- Recuperación:
 - Mecanismos para la recuperación de las bases de datos.
- Seguridad:
 - Mecanismos para la encriptación de la información almacenada y restringir el acceso a los datos.
- Integridad:
 - Mecanismos para la corrección y complementación de los datos en una base de datos.

Recuperación

- Ejemplo de la transacción transferencia de fondos.
- ¿Qué pasa si en medio de la transacción se produce una falla?
 - Transacciones de ejecución atómica.
 - Se realiza del todo o no se realiza.
 - Si la transacción queda a medias, se borra lo que se alcanzó a hacer dejando el estado de la base de datos como el anterior previo al inicio de la transacción fallida.

Recuperación



Control de concurrencia

- Disponibilidad para que varios usuarios autorizados puedan acceder la información de forma simultánea.
- Sirve de arbitraje en caso de conflictos cuando más de un usuario intenta acceder de forma simultánea un mismo registro o tupla.
- La solución del conflicto se da por medio de protocolos o conjuntos de reglas que permiten la concurrencia pero se aseguran de dejar la BD en un estado consistente después de la intervención de los usuarios involucrados en la concurrencia.

Otras funcionalidades del SABD – Seguridad

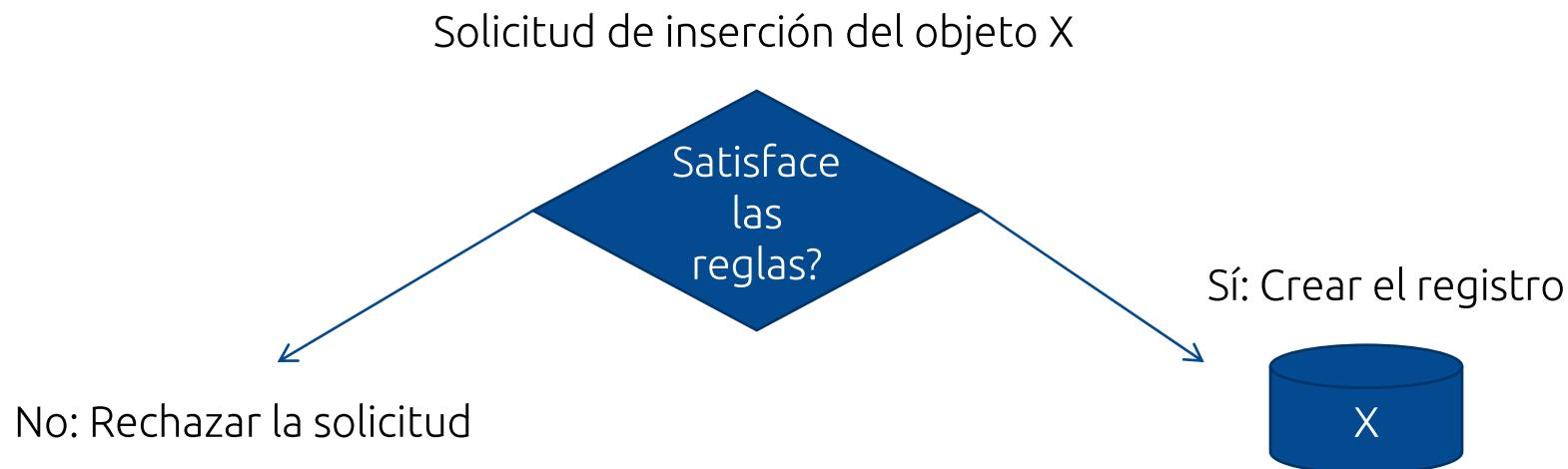
- Usuario y contraseña.
- Encriptación.
- Control de flujo de datos por medio de clases de seguridad.
- Control de inferencia de datos.
- Busca que un usuario no pueda inferir datos privados a partir de datos que sí tiene acceso.

Integridad

- Garantiza que la definición de reglas de integridad ayuden a la verificación sintáctica y semántica de los datos. Ejemplo:
 - El salario de los empleados siempre es mayor o igual al salario mínimo 200 mil colones.
 - La edad de las personas no puede ser menor a 18 años (para una entidad financiera por ejemplo).

Otras funcionalidades del SABD – Integridad

- Impedir la creación de registros inconsistentes.



En resumen, el SABD

- Operaciones de inserción, modificación, borrado y consulta de filas en las tablas de la base de datos.
- Creación, modificación, borrado de tablas y otros objetos en la base de datos.
- Control del accesos concurrente a los datos en un ambiente multiusuario.
- Optimización de consultas.
- Control de acceso a los datos (seguridad).
- Recuperación de la base de datos a un estado consistente ante la presencia de una falla.
- Definición de las características para el almacenamiento de las filas de las tablas.

Sistema de archivos

- La introducción en 1964 de la familia Sistema 360 de IBM impulsó que la industria y comercio utilizaran la computación como herramienta para mejoramiento de la productividad y toma de decisiones.
- Aparición de la cinta magnética a mediados de los años 40.
- Con este dispositivo se podía guardar grandes cantidades de datos a costo razonable.

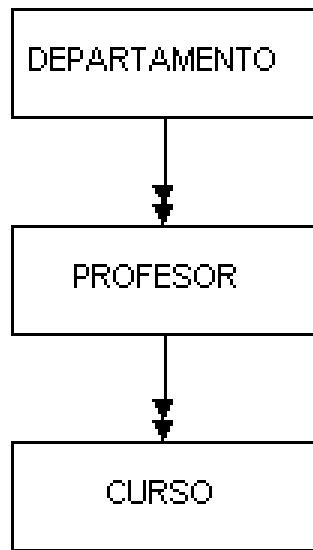
Sistema de archivos

- Se da la noción del concepto de archivo secuencial en el cual el acceso a los registros se hace de la misma forma en que los registros fueron almacenados, uno después del otro sin dejar espacios.
- Con el descenso en los costos de almacenamiento, aumento de capacidad de almacenamiento y la invención de los discos magnético o discos duros se desarrollaron los primeros sistemas de archivos.

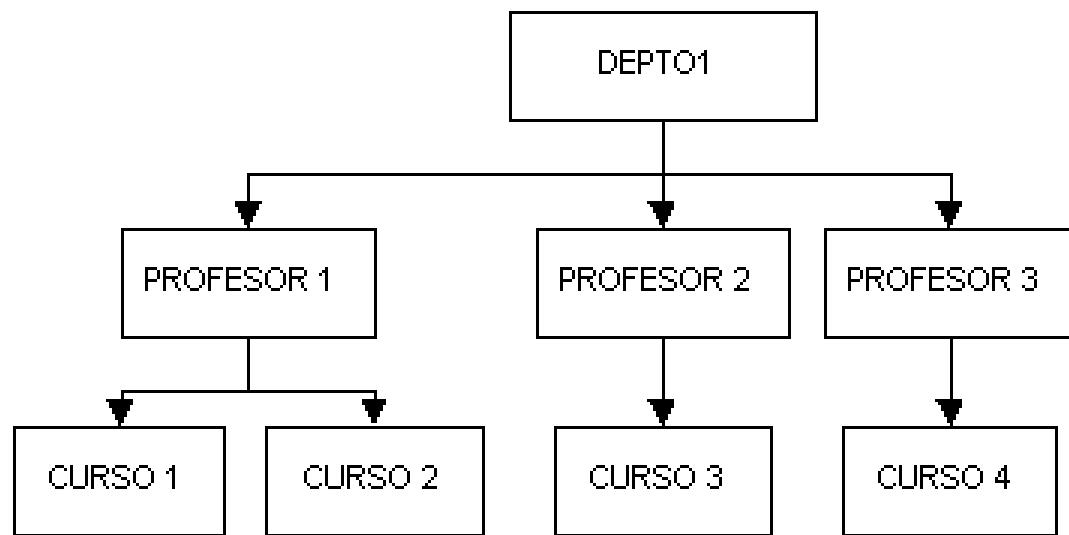
Primera generación: bases de datos jerárquicas

- Problema, un nódulo hijo no puede tener 2 padres.
- Esto impide representar las relaciones muchos a muchos sin tener mucha redundancia.

Estructura lógica

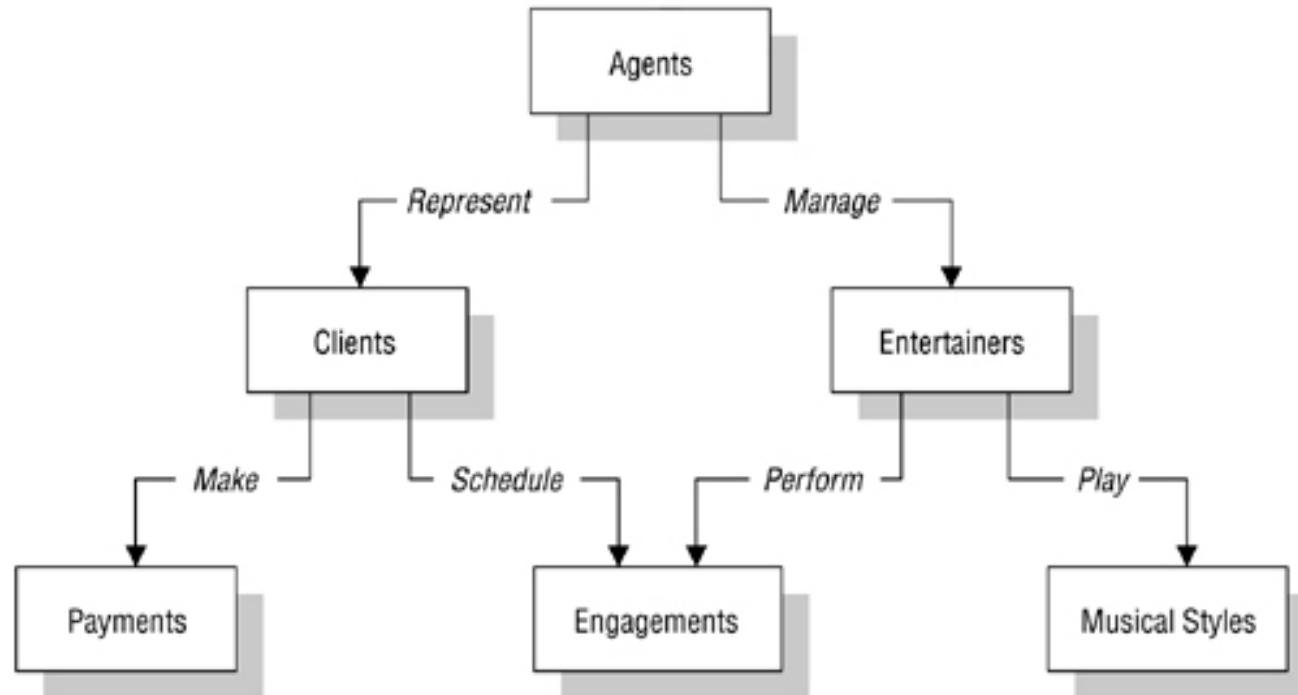


Ejemplo de base de datos



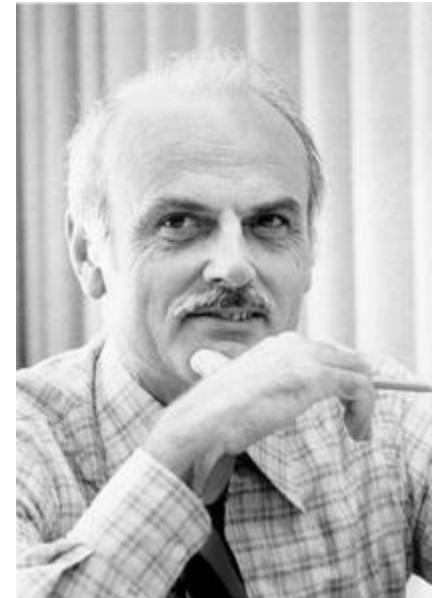
Segunda generación: bases de datos en red

- Mismo esquema jerárquico pero sin la limitación de que un hijo solo puede tener 1 parent.
- Un hijo puede tener más de un parent por medio de punteros, incrementa la complejidad.

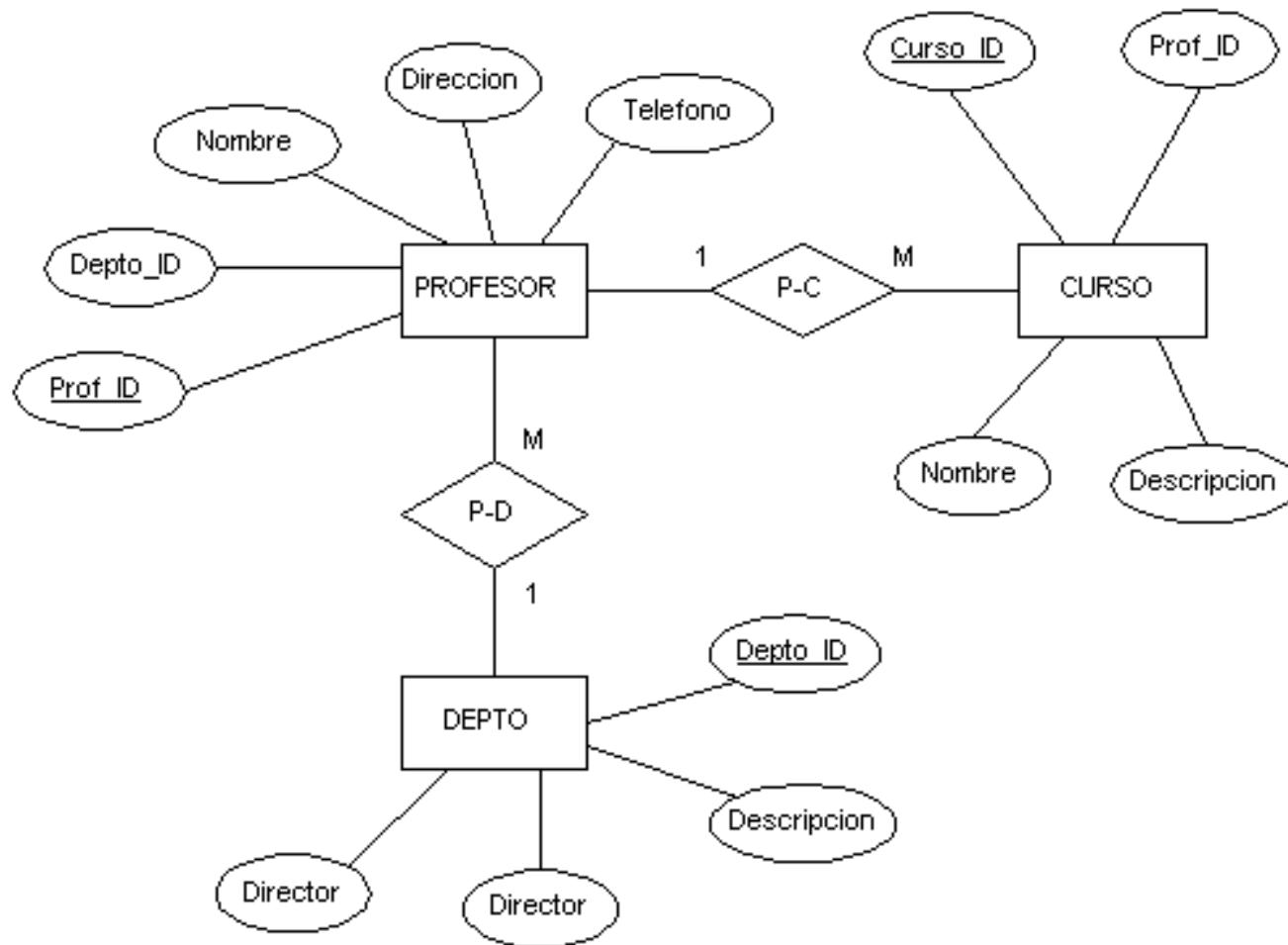


Tercer generación: modelo relacional

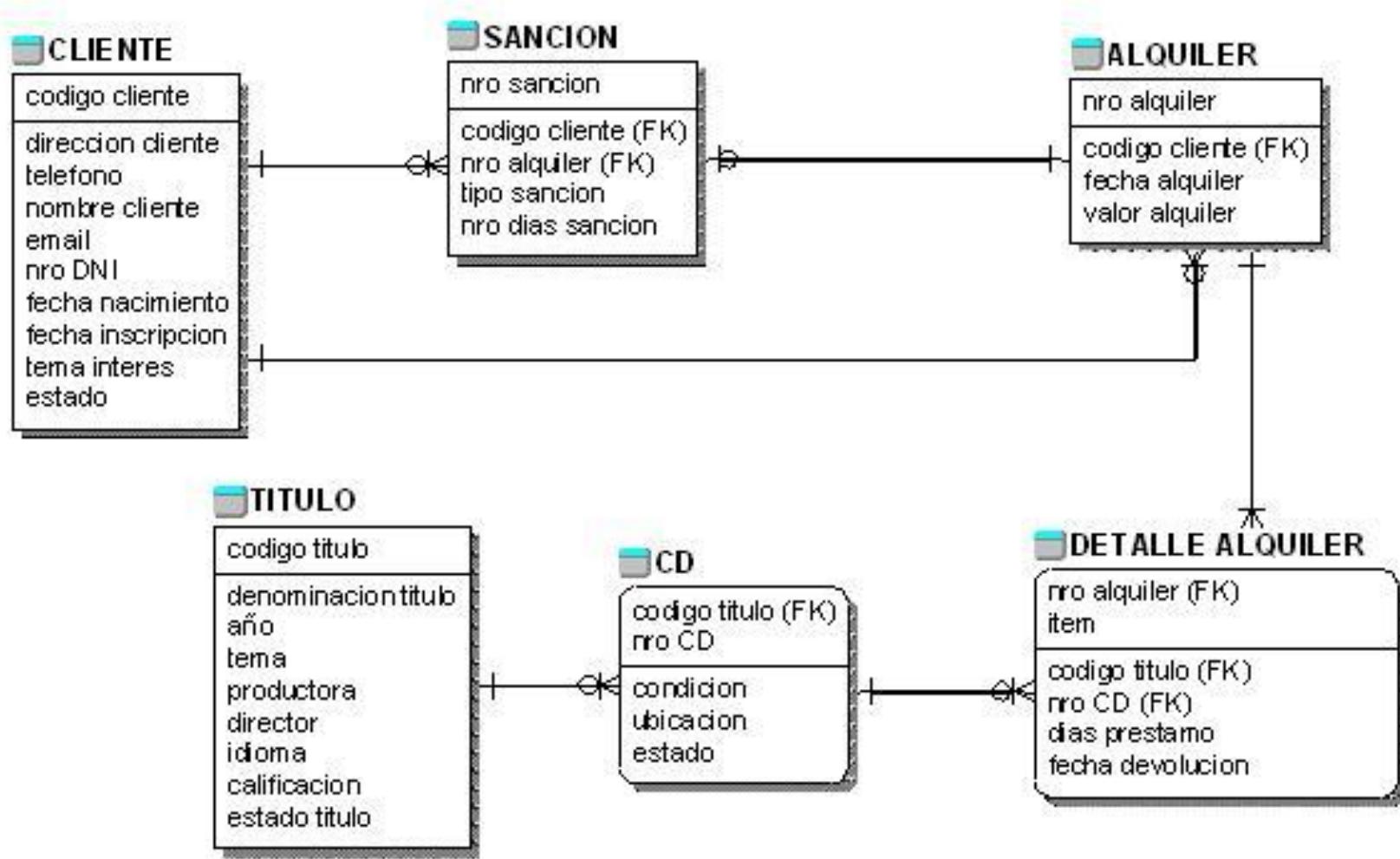
- En 1970, Edgar F. Codd propone un modelo relacional basado en teoría de conjuntos y análisis de predicados.
- La idea consistía en representar los datos como tablas bidimensionales.



Tercer generación: modelo relacional



Tercer generación: modelo relacional



Comparación de sistema de archivos y SABD

- Entidad financiera que capta dinero por medio de cuentas de ahorro.
- Sistema de archivos:
 - Archivo: Cliente_Ahorros.

[2-345-345, Bruce Wayne, Gótica, 200.000]

...

...

...

Comparación de sistema de archivos y SABD

- Una nueva ley le permite a la entidad financiera brindar el servicio de cuentas corrientes.
- Sistema de archivos:
 - Archivo Cliente_Cuentas_Corrientes.
 - Información: número de la cuenta, nombre del cliente, cantón, provincia, teléfono y saldo.

[001-254675-03, Bruce Wayne, Gótica, NY, 2758940, 500.000]

...

...

...

- El señor Wayne se pasa de casa a Los Ángeles en California y por lo tanto debe actualizar el domicilio.

Desventajas del sistema de archivos

- Redundancia de información:
 - El nombre y dirección se almacenan más de una vez.
- Inconsistencia de información:
 - En caso que se actualice la dirección en un registro y en otro no.
- El proceso de modificación de programas existentes es muy costoso.

Ventajas del sistema de bases de datos

- La consulta de la base de datos se hace de forma directa.
- Diferentes usuarios pueden tener acceso a la base de datos de forma simultánea.
- Se reduce la redundancia de datos.
- Permite independencia de las aplicaciones respecto a la estructura física de los datos almacenados.
- Se minimizan las inconsistencias en los datos (con un buen diseño).
- El proceso de modificación de programas se reduce considerablemente.
- Se tienen controles centralizados para mantener la seguridad, privacidad e integridad de datos.

Propiedades de una bases de datos

- Se dice que una base de datos almacena únicamente hechos, que son un reflejo de universo, llamado Universo de Discurso (UoD) determinado.
- Los cambios introducidos en el mini mundo se reflejan en la base de datos.
- Las funciones que el gestor de base de datos se limita a facilitar son, fundamentalmente, las de edición y consulta de los datos.



Propiedades de una bases de datos

- Conjunto de datos lógicamente coherente con significado inherente.
- No es una base de datos un conjunto aleatorio de datos.
- Una base de datos se diseña, construye y rellena con datos para un propósito específico.
- Dispone de un conjunto de usuarios y aplicaciones preconcebidas para satisfacer necesidades de estos usuarios.



Conceptos básicos

Consulta

- Recuperan datos.

Transacción

- Lectura y escritura de datos.

Aplicación

- Programa con el que se accede a la base de datos.

Conceptos básicos – Esquemas

- Descripción de la BD.
- Se especifica en la etapa de diseño de la BD.
- No se espera que cambie bruscamente ni con mucha frecuencia.

Conceptos básicos – Diagramas de esquemas

- Un esquema visualizado se llama diagrama de esquema:

ESTUDIANTE

NOMBRE	NumEstudiante	Clase	Especialidad
--------	---------------	-------	--------------

CURSO

NombreCurso	NumCurso	Horas	Departamento
-------------	----------	-------	--------------

PRERREQUISITO

NumCurso	NumPrerrequisito
----------	------------------

SECCIÓN

IDSeccion	NumCurso	Semestre	Profesor
-----------	----------	----------	----------

INFORME_CALIF

NumEstudiante	IDSeccion	Nota
---------------	-----------	------

Conceptos básicos – Estructura de esquemas

- Cada objeto instanciado del esquema.

Conceptos básicos – Ssnapshots/Instancias

- Estado de la base de datos en un momento concreto.

Transacción

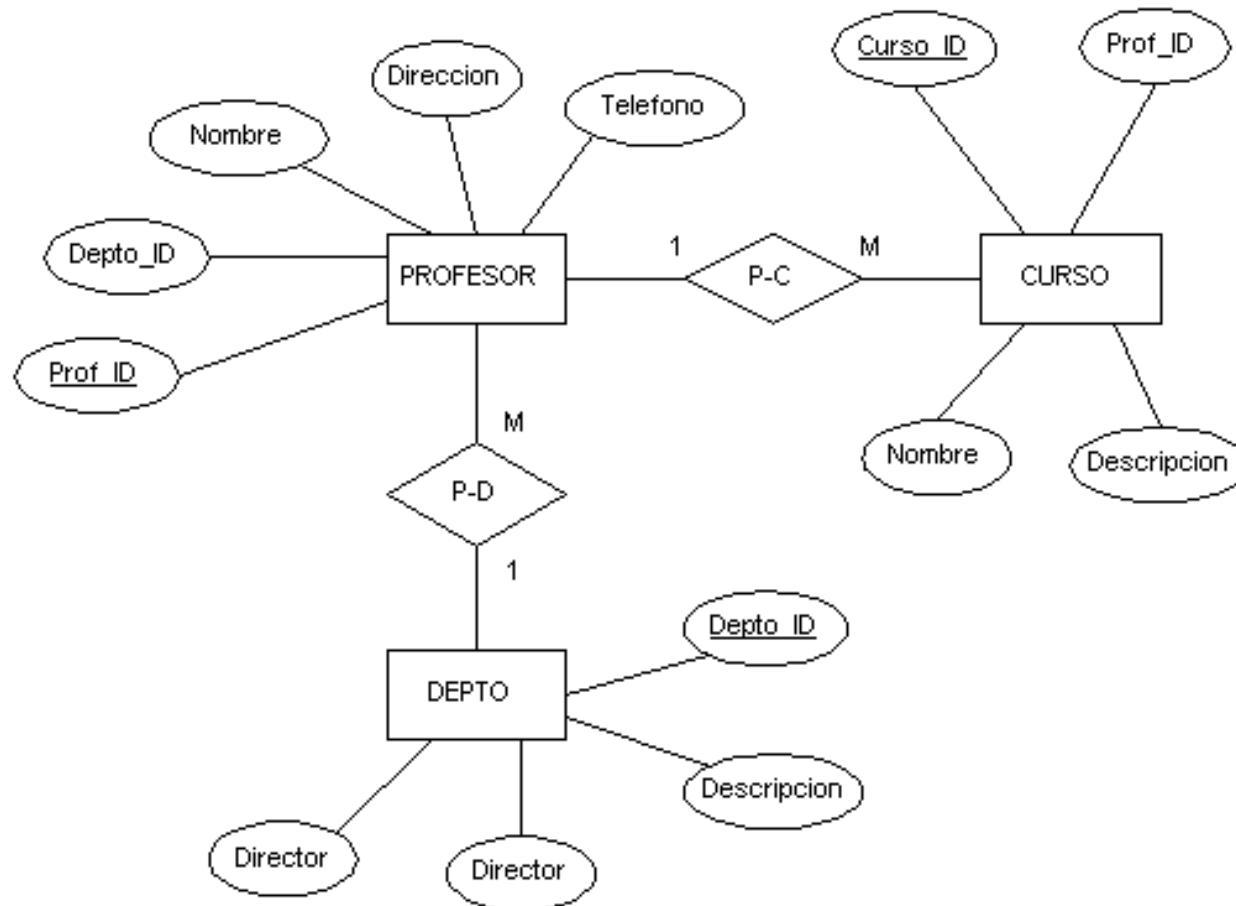
- Conjunto de operaciones (insertar, modificar, borrar, consultar) que forman una unidad indivisible de trabajo, la cual es aplicada completamente o no lo es del todo.

Transferencia bancaria

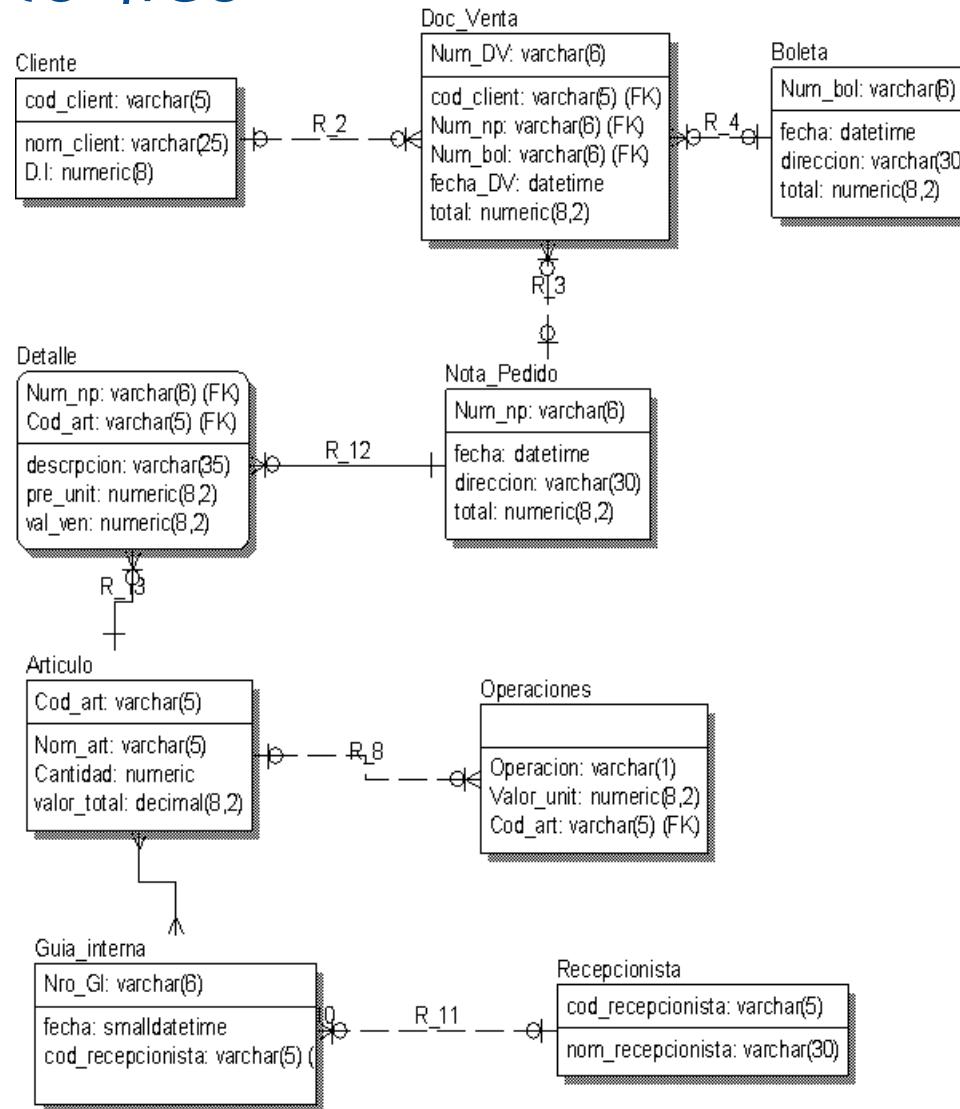
- Suponga que se hace una transferencia de fondos (\$ 1000.00) de la cuenta A de un cliente de un Banco a la cuenta B de otro cliente
- ¿Cuántas operaciones son necesarias?



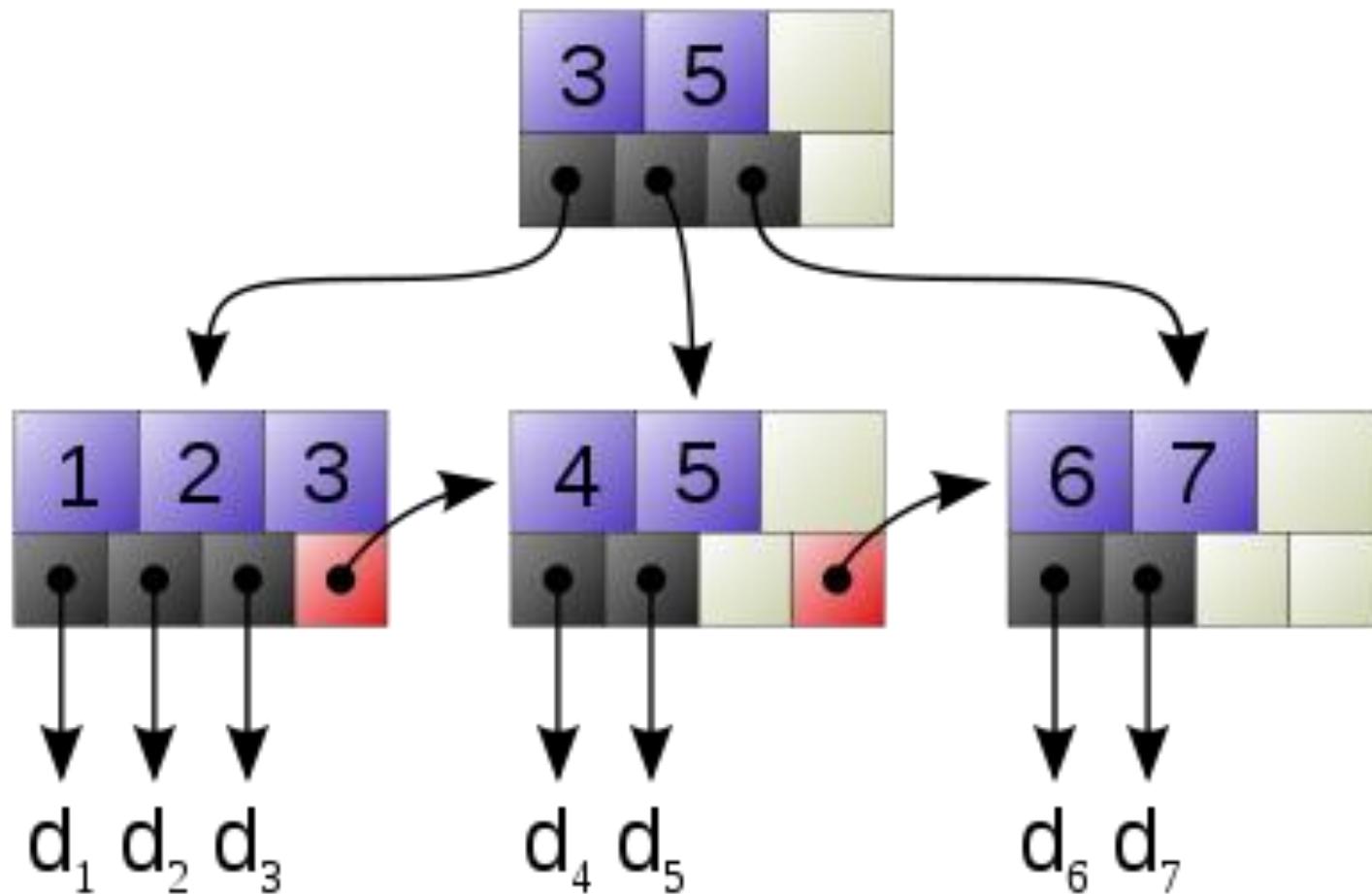
Modelo conceptual



Modelo lógico



Modelo físico



Modelo de datos

- Un modelo de datos es una definición lógica, independiente y abstracta de los objetos, operadores y demás que en conjunto constituyen la máquina abstracta con la que interactúan los usuarios.
- Los objetos nos permiten modelar la estructura de los datos.
- Los operadores permiten modelar su comportamiento.

Modelo de datos

- Desarrollo de modelos y metodologías para representar la información del mundo real.
- Es una definición lógica independiente y abstracta ("sencilla") de los objetos presentes en el área que se pretende automatizar.
- Puede ser implementado físicamente de diversas maneras.

Modelo de datos

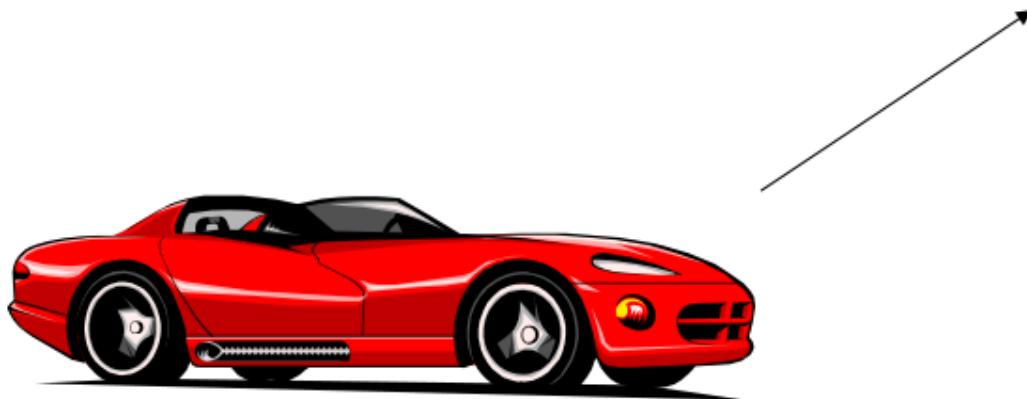
- La implementación de determinado modelo de datos es una realización física, en una máquina real, de los componentes de la máquina abstracta que en conjunto constituyen ese modelo.
- El modelo es aquello que los usuarios tienen que conocer; la implementación es lo que los usuarios no tienen que conocer.

Modelo de datos

- Conjunto de conceptos que pueden ser usados para describir la estructura de la BD que permiten abstraer la información.
- Por estructura entiéndase los tipos de datos, relaciones y restricciones (constraints) que almacenan la data.
- El modelo relacional es un modelo de datos, no el único, hay otros modelos.

Categorías de modelos de datos – Conceptual (alto nivel)

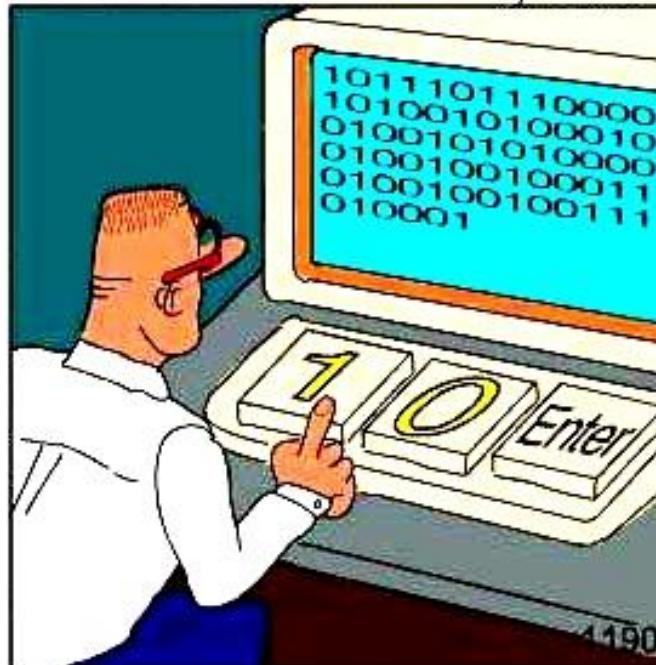
- Provee conceptos parecidos a como los usuarios los conciben. Se le llaman también modelos de datos basados en entidad o basados en objetos.
- Utilizan conceptos como entidades, atributos y relaciones.
- Una entidad representa un objeto o concepto del mundo real.



- **Atributos:**
 - color
 - velocidad
 - ruedas
 - motor

Categorías de modelos de datos – Físico (bajo nivel)

- Proveen conceptos que describen a detalle cómo la data es almacenada en la computadora.
- Están pensados para el programador, no para los usuarios finales.



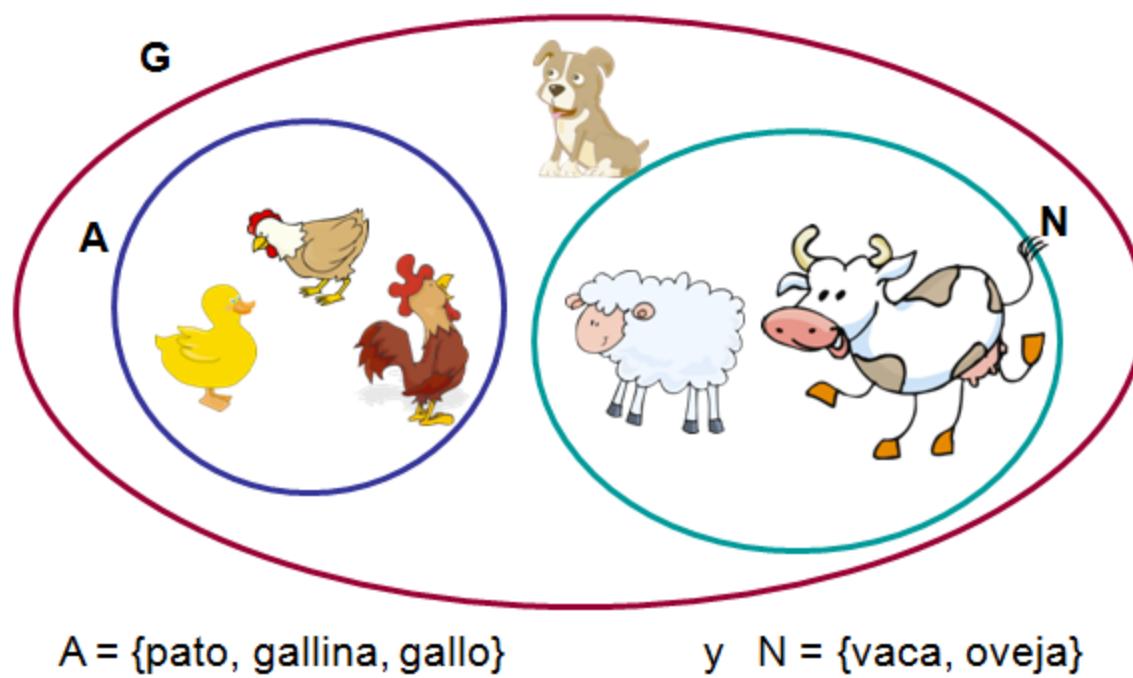
REAL Programmers code in BINARY.

Categorías de modelos de datos – Implementación (representación)

- Provee conceptos que pueden ser entendidos por usuarios finales pero similar a cómo la data es organizada en el computador.
- Esconde algunos detalles de almacenamiento de la data.

Modelo relacional

- Modelo de datos basado en lógica de predicados y teoría de conjuntos.



Modelo relacional

- Las ideas fundamentales del modelo de datos relacional aparecieron en forma pública en artículo de Codd de 1970.
- La base de datos se puede ver como un conjunto de estructuras llamadas relaciones, donde los datos se organizan en forma tabular.
- Los SABD son muy confiables y poderosos a pesar de su simplicidad.

Modelo relacional

- El modelo relacional se basa en el concepto de relación.
- Los datos se representan en forma matricial.
- Por ejemplo, en un sistema de información de ventas una entidad de la cual se almacenan datos, es CLIENTE.

Modelo relacional

- Descripción de datos:
 - Entidades y relaciones se representan en forma de tablas.
 - Las tablas reciben el nombre de relación.
 - Las filas (tuplas) contienen datos sobre cada entidad.
 - Las columnas corresponden a los atributos de las entidades.

Modelo relacional

- Operaciones:
 - Unión, intersección, diferencia, producto cartesiano, selección, proyección y reunión.
- Restricciones de integridad:
 - Integridad de entidad e integridad referencial(mediante el uso de claves).

Ejemplo modelo relacional

Estudiante

	Codigo	Nombre	Cedula	Fecha de Nacimiento
01	Mario Benedetti	1-876-192	30/04/1990	
02	María Borges	2-452-985	21/05/1978	
03	Horario Quiroga	4-567-321	31/12/1985	
04	Gabriel García	7-593-495	04/08/1989	

Curso

	Codigo	Nombre	Creditos	Horas
	IC-4300	Bases de Datos	4	4
	IC-2101	POO	3	4
	IC-2001	Estructuras de Datos	4	4
	IC-2103	Cálculo	3	3

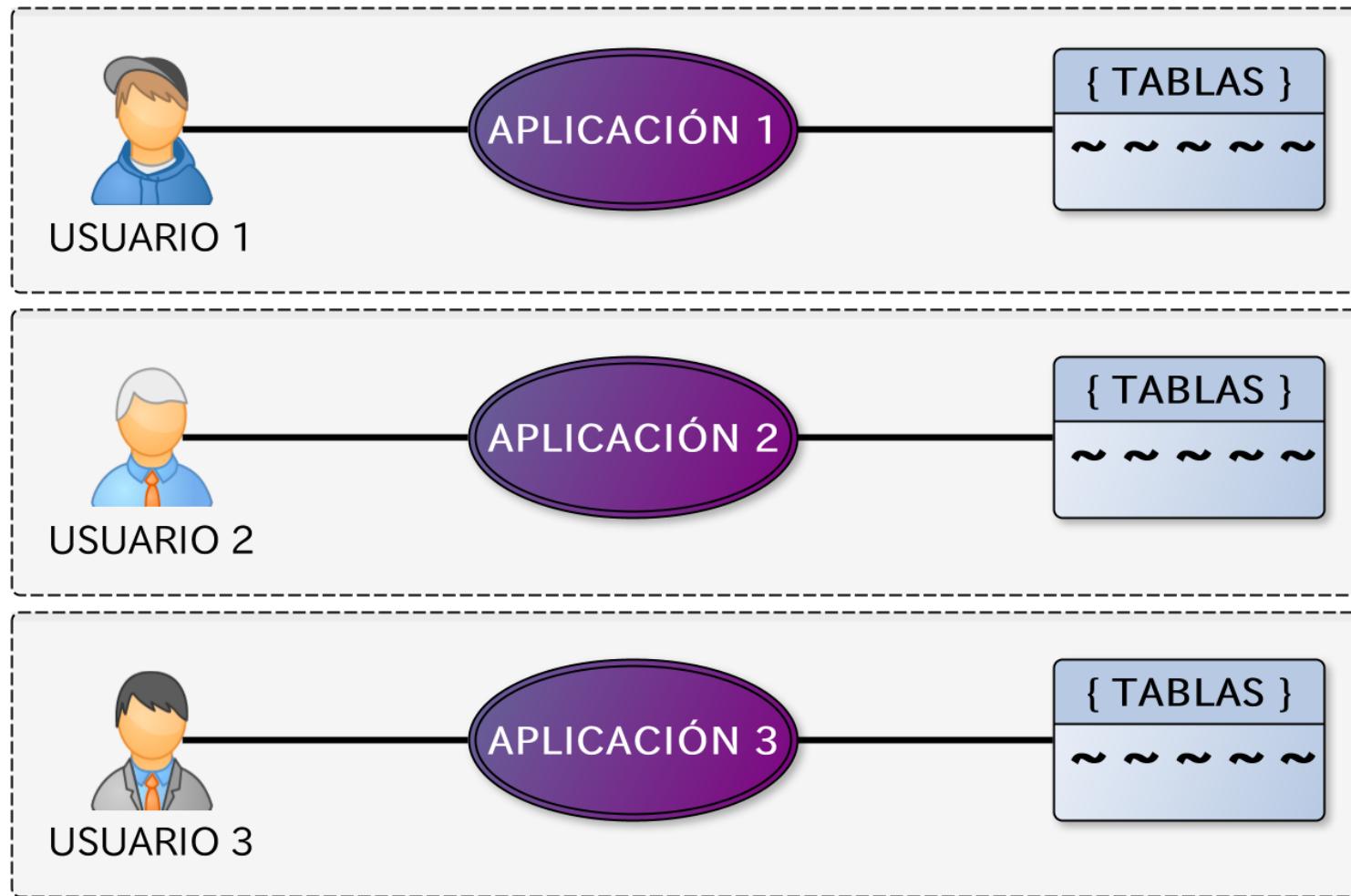
Ejemplo modelo relacional

Entidades	Estudiante	Curso
Atributos	Código, Nombre, Cédula, Fecha de Nacimiento	Código, Nombre, Créditos, Horas
Relación	Cursa	
Cardinalidad	Muchos a muchos	

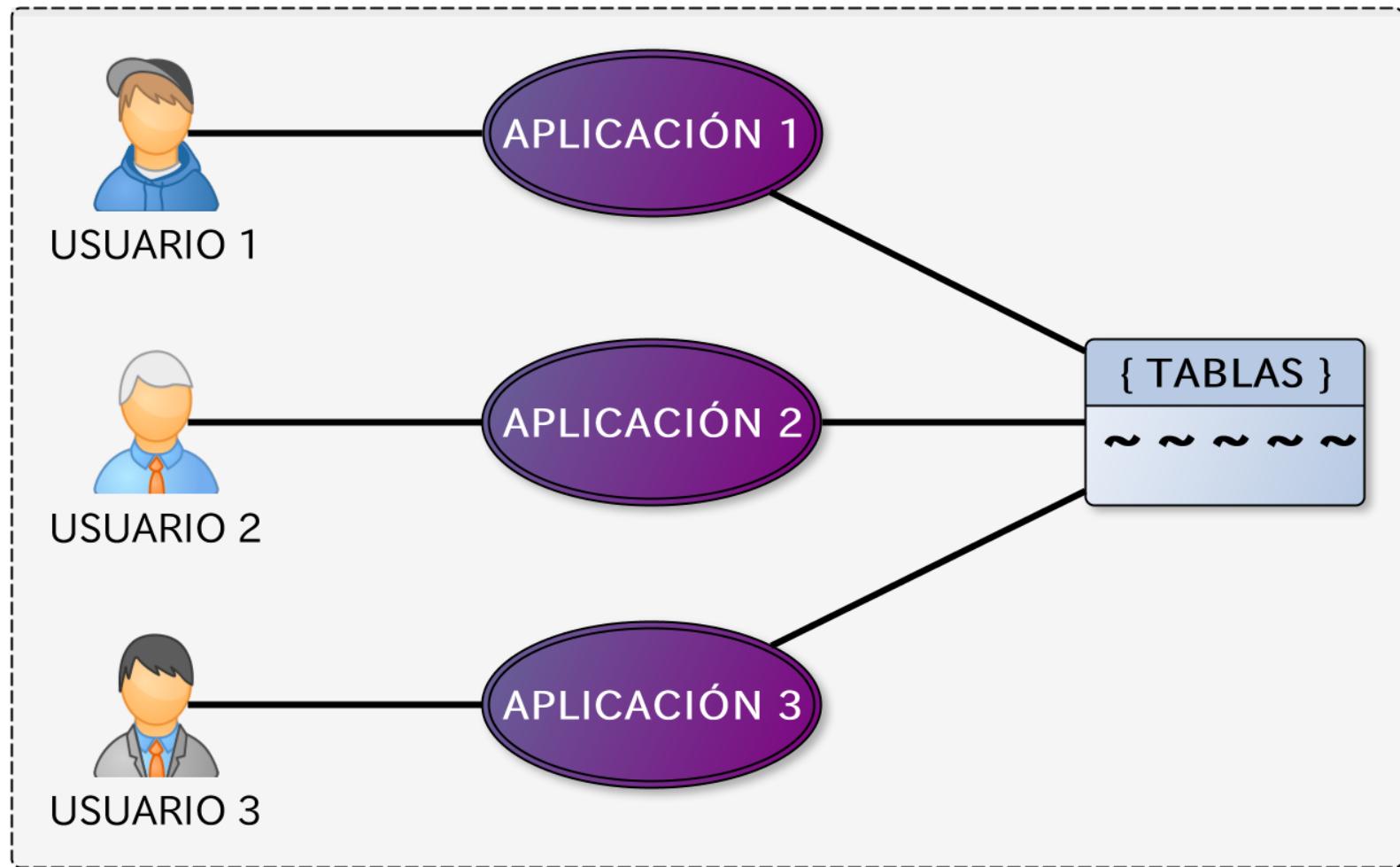
Arquitectura SABD

- En una organización se pueden desarrollar aplicaciones independientes, cada una con su propio repositorio de datos.
- No es conveniente desde un punto de vista integral.
- ¿Por qué?
 - Se presentan problemas de donde están los datos más actualizados y veraces.
 - Es difícil su utilización para la generación de informes para los niveles medios y ejecutivos de la organización.

Modelo no integrado



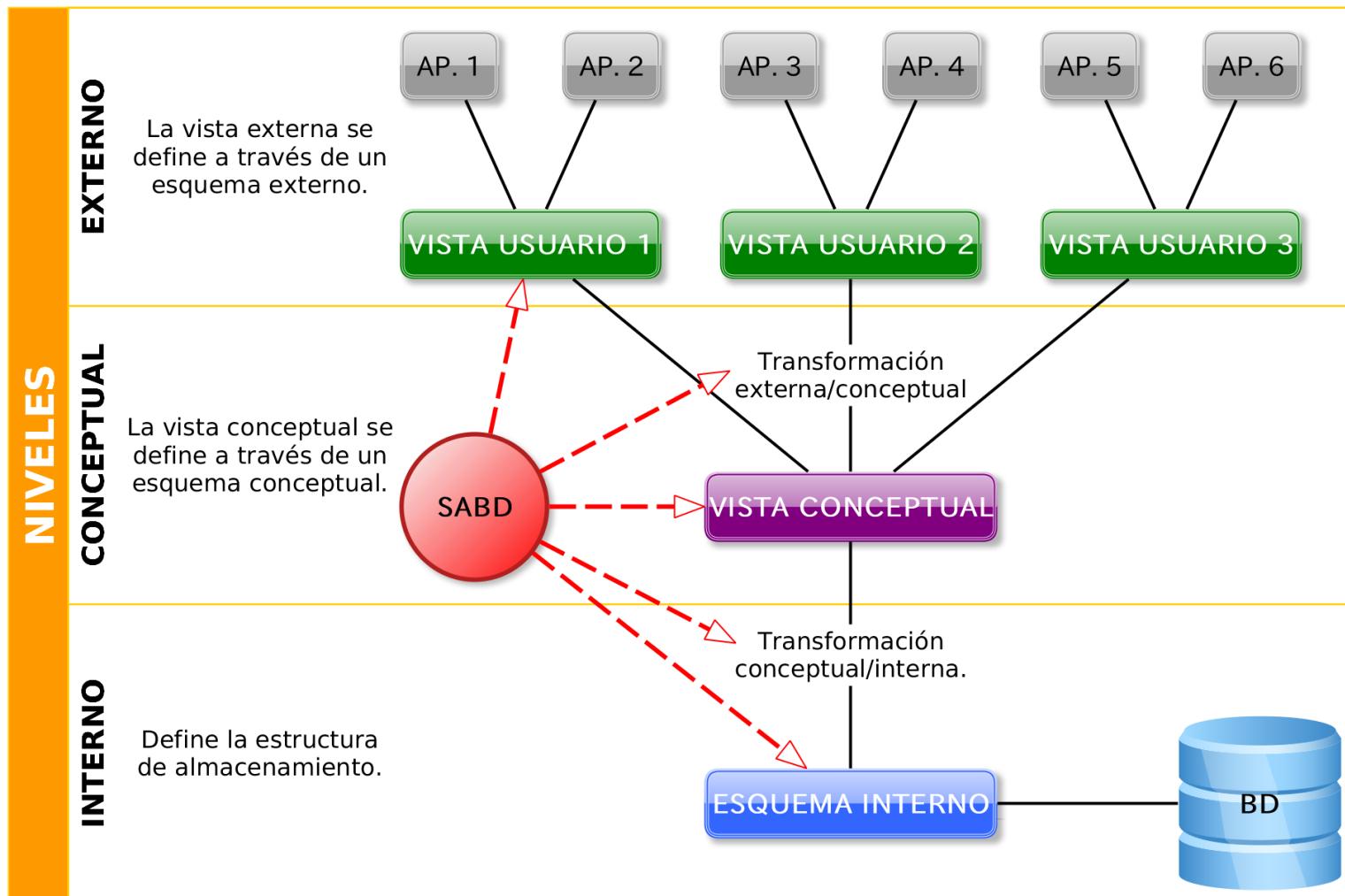
Modelo integrado



Ventajas del modelo integrado

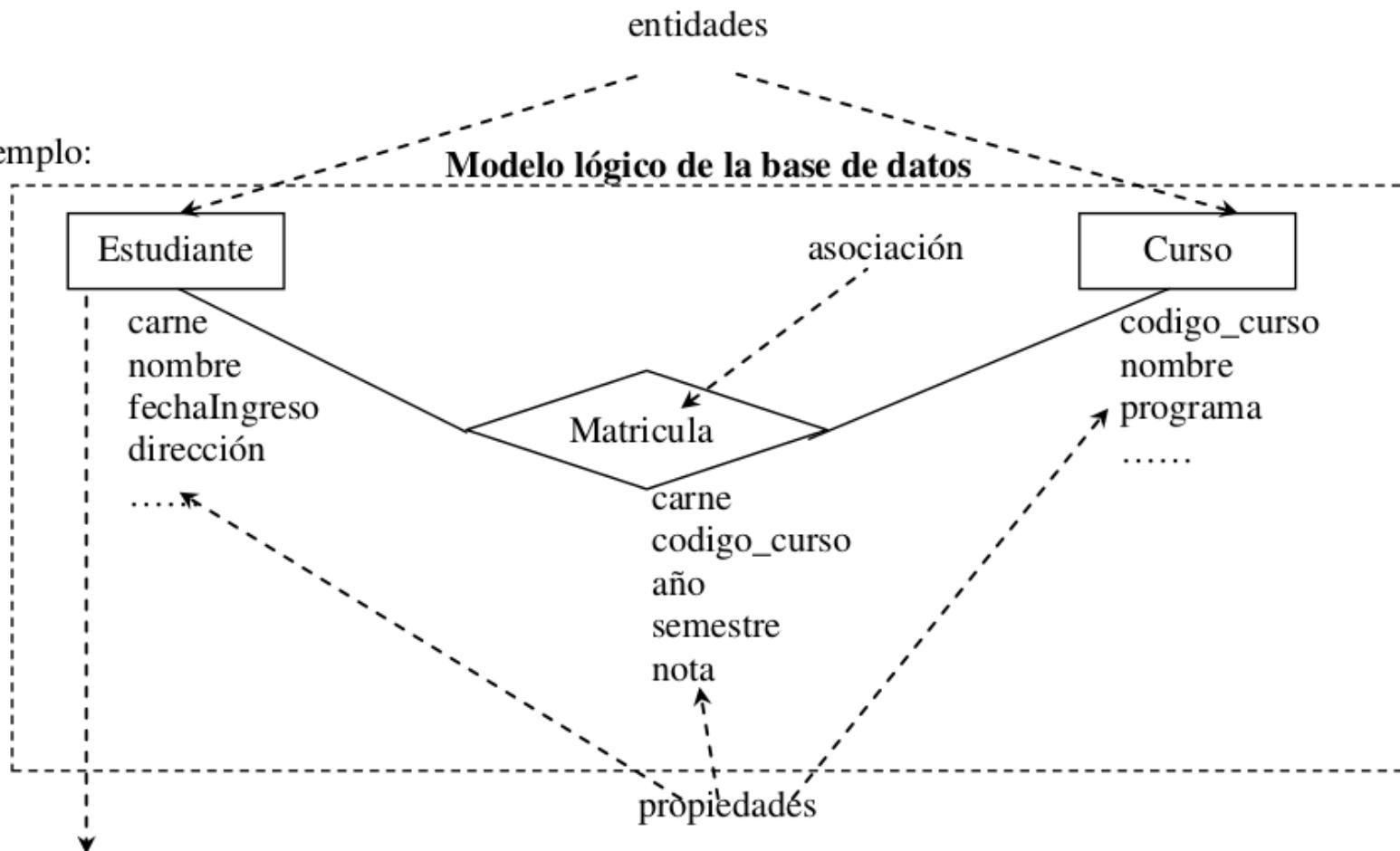
- Reduce la redundancia.
- Facilita el cumplimiento de estándares en cuanto a nombramiento de propiedades.
- Facilita el hacer cumplir la seguridad.

Arquitectura de 3 esquemas



Modelo lógico

Ejemplo:



Modelo lógico

- Cada entidad y asociación entre entidades del modelo lógico se representa como una tabla, donde en la primera fila se colocan las propiedades y en las filas siguientes los valores de esas propiedades.

Modelo lógico

- Entidad Estudiante:

▼

carne	nombre	fechaIngreso
20060001	José ...	02-02-2006
		
		

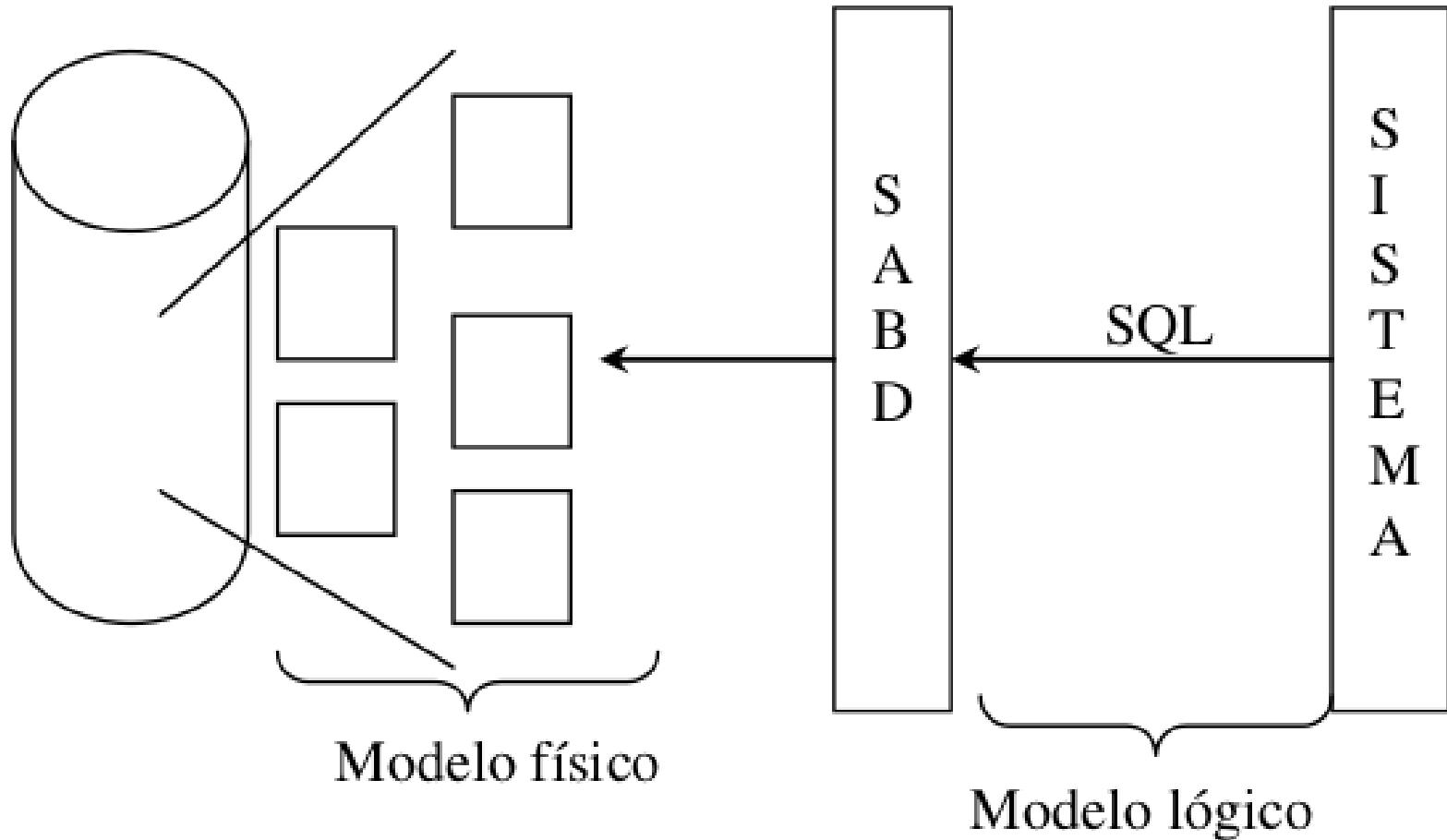
|

		
20070998	Pedro	02-02-2007

Modelo lógico

- Para la especificación de un modelo lógico es necesario realizar el análisis de requerimientos del área que abarcará el sistema, y dar respuesta a las siguientes preguntas:
 - ¿Qué datos deben ser capturados?
 - ¿Cómo deben ser procesados los datos?
 - ¿Cómo deben ser reportados los datos?
- Las preguntas se expanden a una serie de preguntas más detalladas que incluyen: formatos de pantalla para captura y consulta de datos, procesos aplicados a los datos, bitácora de auditoria, período de retención, volúmenes de almacenamiento, ciclos de procesamiento, formateo de reportes, entre otros.

Modelo físico - Modelo lógico



Funciones del administrador de una base de datos

- Definir los esquemas conceptual e interno (hacer el script).
- Definir las restricciones de seguridad e integridad.
- Supervisar el rendimiento y hacer propuestas para responder a nuevos requerimientos.

Lenguajes de base de datos – Data Definition Language (DDL)

- Utilizados por los DBA's y diseñadores de la BD para definir los esquemas conceptual e interno.
- El SABD tendrá un compilador DDL para procesar las sentencias para identificar las descripciones de las estructuras del esquema y almacenar la descripción del mismo en el catálogo del SABD.
- En SABD que hacen diferencia entre los niveles conceptual y el interno, se utilizado el DDL solo para el conceptual.

Ejemplo DDL

```
CREATE TABLE estudiante
(
    carne NUMBER(10,0),
    nombre VARCHAR(50),
    fechIngreso      DATE
)
```

Lenguajes de base de datos – Data Manipulation Language (DML)

- El SABD proporciona a los usuarios medios para manipular la data por medio de un lenguaje u operaciones como inserción, borrado, recuperación y modificación.

Ejemplo DML

- SELECT * FROM estudiante;
- UPDATE estudiante
SET nombre = 'Tobias Flores Mora'
WHERE carnet = 200101;

SQL

- Lenguaje estructurado de consultas.
- Permite la definición de estructuras por medio de un lenguaje de definición de datos (DDL) y la realización de operaciones por medio de un lenguaje de manipulación de datos (DML).

SQL

- Las instrucciones SQL son procesadas por un optimizador de consultas que es parte del SABD, que es quién decide como realizarlas sobre las estructuras en las que se encuentra almacenada físicamente la base de datos.

Diccionario de datos

- Contiene datos (metadatos) acerca de datos.
- Definiciones de objetos.
- Definiciones de restricciones de integridad.
- Estadísticas de uso.
- Usuarios.
- Procedimientos almacenados.
- Restricciones de seguridad.
- Esta integrado dentro de la base datos.

Diccionario de datos

- Herramienta fundamental para creación, diseño y mantenimiento de la base de datos.
- Repositorio que brinda información de los datos que se encuentran almacenados en la base de datos.
- Es a su vez una base de datos.

Diccionario de datos

- Nombres de las relaciones y atributos.
- Llaves primarias, alternas, externas de las relaciones
- Reglas de integridad.
- Para cada atributo debe tener:
 - Tipo de dato.
 - Formato.
 - Rango.
 - Significado.
 - Valores permitidos.
 - Unicidad.
 - Soporte de nulos.

Objetivos del diccionario

- Permitir la administración de la documentación de los datos.
- Entender los datos y lo que significan (de cara al usuario).
- Controlar los recursos de datos.

Beneficios del diccionario

- Lograr homogeneidad en la representación de los datos.
- Lograr concordancia en las definiciones de los datos.
- Generar automáticamente la descripción de los datos.
- Reducir el costo del desarrollo del sistema.
- Cumplir con los estándares de los datos.

Beneficios del diccionario

- Eliminar redundancia de los datos.
- Mejorar el control de cambios.
- Reducir el esfuerzo de análisis de datos.
- Ayudar a la comunicación.
- Ayudar en desarrollo paralelo.
- Mejorar la seguridad.

Gracias - ¿Preguntas?