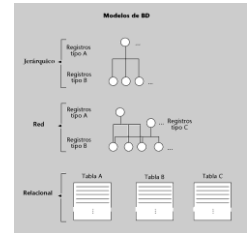


Este tema trata acerca del modelado de datos y por qué los modelos de datos son importantes. Aprenderemos los elementos básicos del modelado de datos, qué y cómo influyen las reglas de negocio en el diseño de bases de datos y cómo pueden clasificarse los modelos de datos por nivel de abstracción.



INDICE

| | |
|---|----------|
| 1. MODELADO DE DATOS Y MODELOS DE DATOS | 1 |
| 2. LA IMPORTANCIA DE MODELOS DE DATOS | 2 |
| 3. ELEMENTOS BÁSICOS DE UN MODELO DE DATOS | 2 |
| 4. REGLAS DE NEGOCIOS | 4 |
| 4.1. DESCUBRIMIENTO DE LAS REGLAS DE NEGOCIO | 4 |
| 4.2. CONVERSIÓN DE REGLAS DE NEGOCIOS EN COMPONENTES DE MODELO DE DATOS. | 4 |
| 5. LA EVOLUCIÓN DE LOS MODELOS DE DATOS | 5 |
| 6. GRADOS DE ABSTRACCIÓN DE DATOS | 7 |
| 7. EJERCICIOS | 9 |

1. Modelado de Datos y Modelos de Datos

El diseño de bases de datos se concentra en la forma en que la estructura de bases de datos se usará para guardar y administrar datos del usuario final. El **modelado de datos**, primer paso para diseñar una base de datos, se refiere al proceso de crear un modelo específico de datos para el dominio de un problema determinado. Un **dominio de problema** es un área claramente definida dentro del ambiente real, con ámbito y fronteras bien delimitados.

Un **modelo de datos** es una representación relativamente sencilla, por lo general gráfica, de estructuras de datos reales más compleja.

Dentro del ambiente de una base de datos, el **modelo** representa estructuras de datos y sus características, relaciones, restricciones, transformaciones y otras construcciones con el propósito de sostener un dominio de problema específico.

El **modelado de datos** es un proceso iterativo y progresivo. Se empieza con una comprensión sencilla del dominio del problema y a medida que aumenta ésta, también se incrementa el nivel de detalle del modelo de datos. Si se hace en forma apropiada, el modelo de datos final es un "plano" que contiene todas las instrucciones para construir una base de datos que va a satisfacer todas las necesidades del usuario final. Este "plano" es

narrativo y gráfico, lo cual significa que contiene descripciones de texto en lenguaje sencillo, no ambiguo y claro, que describe los elementos principales de los datos.

Un modelo de datos listo para implementación debe contener al menos los siguientes componentes:

- Una **descripción de la estructura de datos** que guardará los datos del usuario final.
- Un **conjunto de reglas** que se pueden hacer cumplir para garantizar la integridad de los datos.
- Una **metodología de manipulación de datos** para apoyar las transformaciones de los datos reales.

Por tradición, los diseñadores de datos se apoyaban en el sentido común para desarrollar un buen modelo de datos. Desafortunadamente, el sentido común es un concepto relativo y con frecuencia se desarrolla después de muchas pruebas y errores. Por ejemplo, si cada uno de los estudiantes de un grupo ha de crear un modelo de datos para un almacén de video, es muy probable que cada uno de ellos llegue con un modelo diferente. ¿Cuál sería el correcto? La respuesta más sencilla es "el que satisfaga todas las necesidades del usuario final" y puede haber más de una solución correcta. Por fortuna, los diseñadores de bases de datos utilizan construcciones ya existentes del modelado de datos y de poderosas herramientas para el diseño de bases de datos que reducen considerablemente el potencial de errores en el modelado de bases de datos. En los temas siguientes aprenderemos cómo se usan los modelos de datos para representar datos reales y cómo los diferentes grados de abstracción de datos facilitan el modelado de datos.

2. La importancia de Modelos de Datos

Los **modelos de datos** son una herramienta de comunicación, es decir, pueden facilitar la interacción entre el diseñador, el programador de aplicaciones y el usuario final. Un modelo bien diseñado puede incluso promover un mejor entendimiento de la organización para la cual se desarrolló el diseño de la base de datos.

La importancia del modelado de datos no se puede exagerar. Los datos constituyen las unidades de información más elementales que un sistema emplea. Las aplicaciones son creadas para manejar datos y ayudar a transformarlos en información. Pero los datos son vistos en distintas formas por personas diferentes aun trabajando en la misma compañía.

Los programas de las aplicaciones tienen también otra vista de los datos, más relacionada con la ubicación de datos, formateo y necesidades específicas de informes.

3. Elementos Básicos de un Modelo de Datos

Los elementos básicos de todos los modelos de datos son: **entidades, atributos, relaciones y restricciones.**

- Una **entidad** es cualquier (una persona, lugar, cosa o hecho) acerca de la cual se han de coleccionar y guardar datos. Una entidad representa un tipo particular de objeto en el mundo real. Como una enti-

dad representa un tipo particular de objeto, las entidades son "distinguibles", es decir, cada vez que se presenta una de ellas es única y distinta. Por ejemplo, la entidad CLIENTES.

- Un **atributo** es una característica de una entidad. Por ejemplo, una entidad CLIENTE sería descrita por atributos tales *apellidos, nombre, teléfono, dirección y límite de crédito del cliente*. Los atributos son el equivalente de los campos en los sistemas de archivos.
- Una **relación** describe una asociación entre entidades. Por ejemplo, existe una relación entre CLIENTES y AGENTES que se puede describir como: un agente puede atender a numerosos clientes y cada cliente puede ser atendido por un agente. Los modelos de datos usan tres tipos de relaciones: uno a muchos, muchos a muchos y uno a uno. Los diseñadores de bases de datos por lo general usan las notaciones breves **1:N ó 1..***, **N:M ó *..*** y **1:1 ó 1..1**, respectivamente. Los siguientes ejemplos ilustran las distinciones entre las tres.
 - **Relación uno a muchos (1:N o 1..*)**. *Un pintor crea muchas obras, pero cada una de ellas es pintada por sólo un pintor*. Así, el pintor (el "uno") está relacionado con las pinturas (las "muchas"). Por tanto, la relación (PINTOR pinta PINTURA) como 1:N. Los nombres de entidad con frecuencia se ponen en mayúsculas como convención, de modo que se identifiquen con facilidad. *Un cliente puede generar muchas facturas, pero cada factura es generada para un solo cliente*, La relación (CLIENTE genera FACTURA) también sería marcada como 1:N.
 - **Relación muchos a muchos (N:M o *..*)**. *Un empleado puede aprender muchas habilidades en el trabajo y cada habilidad de trabajo puede ser aprendida por muchos empleados*, la relación (EMPLEADO aprende HABILIDAD) como N:M. Del mismo modo, *un estudiante puede matricularse de muchos módulos y en cada módulo pueden estar matriculados muchos estudiantes*, dando así la leyenda de relación M:N para la relación expresada por (ESTUDIANTE están matriculados en MODULOS).
 - **Relación uno a uno (1:1 o 1..1)**. La estructura de administración de una compañía de ventas al por menor puede requerir que *cada una de sus tiendas sea manejada por un solo empleado*. A su vez, *el gerente de cada tienda, que es un empleado, maneja sólo una tienda*. Por tanto, la relación (EMPLEADO maneja TIENDA) se marca como 1: 1. Cada ciclo tiene asignado un profesor-tutor a su vez, un profesor-tutor lo es de un solo ciclo. La relación TUTOR-CICLO se marca como 1:1.

Se identificó cada relación en ambas direcciones; esto es, las relaciones son bidireccionales:

- Un CLIENTE puede generar muchas FACTURAS.
- Cada una de las muchas FACTURAS es generada para un solo CLIENTE.
- Una **restricción** se aplica a los datos. Las restricciones son importantes porque ayudan a asegurar la integridad datos. Las restricciones se expresan normalmente en forma de reglas. Por ejemplo:
 - El salario de un empleado puede tener valores que están entre 6 000 y 350 000.
 - El PDC (promedio de calificaciones) de un estudiante puede estar entre 1.00 y 10.00.
 - Cada módulo es impartido por un profesor y sólo por uno.

¿Cómo identificamos correctamente entidades, atributos, relaciones y restricciones? El primer paso es identificar con claridad las reglas de negocios para el dominio de problema que se modele.

4. Reglas de Negocios

Cuando los diseñadores empiezan a seleccionar o determinar las entidades, atributos y relaciones que se usarán para construir un modelo de datos, pueden empezar por entender qué tipos de datos hay en una organización, cómo se usan los datos y qué marcos de tiempo se usan. Una **regla de negocios** es una descripción breve, precisa y no ambigua de una política, procedimiento o principio dentro de una organización específica.

Las reglas de negocios deben darse por escrito y actualizarse para reflejar cualquier cambio en el ambiente operacional de la organización.

Las reglas de negocios debidamente escritas se usan para definir entidades, atributos, relaciones y restricciones.

Cada vez que se vean enunciados de relación, como "*un agente puede atender a varios clientes y cada cliente puede ser atendido por un solo agente*", se observan reglas de negocios aplicadas.

4.1. Descubrimiento de las reglas de negocio

Las **principales fuentes de reglas de negocios** son los gerentes de compañías, directores de políticas, gerentes de departamento y documentación escrita como los procedimientos y normas de una compañía y los manuales de operaciones.

Una fuente más rápida y directa de reglas de negocios son las entrevistas directas con usuarios finales pero a veces las percepciones difieren, los usuarios finales son a veces una fuente menos confiable.

El proceso de identificar y documentar reglas de negocios es esencial para el diseño de bases de datos, por varias razones:

- Ayudan a estandarizar la visión de datos de la compañía.
- Pueden ser una herramienta de comunicación entre usuarios y diseñadores.
- Permiten que el diseñador entienda la naturaleza, función y propósito de los datos y desarrolle restricciones y reglas apropiadas.

Hay reglas de negocios que no se pueden modelar, por ejemplo, *ningún piloto podrá volar más de 10 horas dentro de cualquier periodo de 24 horas*.

4.2. Conversión de reglas de negocios en componentes de modelo de datos.

Las reglas de negocios preparan el escenario para la correcta identificación de entidades, atributos, relaciones y restricciones. Como regla general, un **sustantivo** en una regla de negocios lo convertirá en **entidad** en el modelo y un **verbo** (activo o pasivo) que asocie sustantivos lo convertirá en una **relación** entre entidades. Por ejemplo, la regla de negocios "*un cliente puede generar muchas facturas*" contiene dos sustantivos (*cliente* y *facturas*) y un verbo (*generar*) que los asocia.

Para identificar en forma correcta el tipo de relación se deben considerar que las relaciones son bidireccionales. Por ejemplo, la regla de negocios "un cliente puede generar muchas facturas" es complementada por la regla "una factura es generada sólo para un cliente". En este caso, la relación es uno a muchos (1:N). El cliente es el lado del "1" y la factura es el lado de "muchos".

Como regla general, para identificar correctamente el tipo de relación, deben hacerse dos preguntas:

- ¿Cuántas instancias de B están relacionadas con una instancia de A?
- ¿Cuántas instancias de A están relacionadas con una instancia de B?

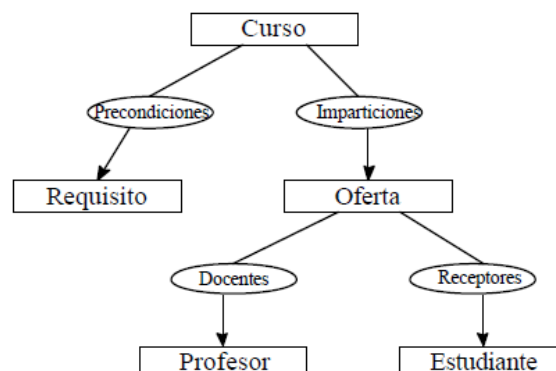
Los **nombres de entidad** deben ser descriptivos de los objetos en el ambiente de negocios y debe emplearse terminología que sea conocida por los usuarios. Un nombre de **atributo** también debe ser descriptivo de los datos representados por ese atributo. Es una buena práctica aplicar un prefijo al nombre de un atributo, con el nombre de la entidad (o una abreviatura del nombre de la entidad) en la que se presente. Por ejemplo, en la entidad CLIENTES, el límite de crédito del cliente puede recibir el nombre de CLI_LIMITE_CREDITO.

5. La Evolución de los Modelos de Datos

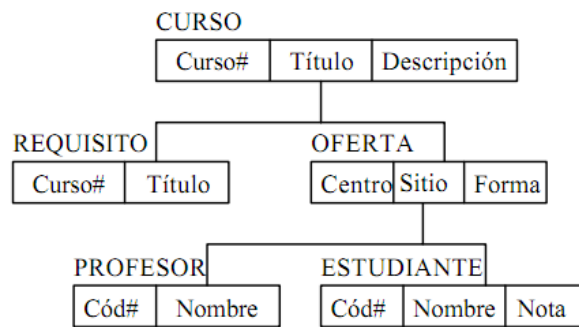
La búsqueda de una mejor administración de datos ha llevado a varios modelos que tratan de resolver los defectos críticos del sistema de archivos. Estos **modelos** representan qué es una base de datos, qué debe hacer, los tipos de estructuras que debe emplear, así como la tecnología que debería usarse para implementar estas estructuras.

Los modelos de datos jerárquico y de red fueron los precursores y ya no se usa, pero algunos conceptos se encuentran en actuales modelos de datos.

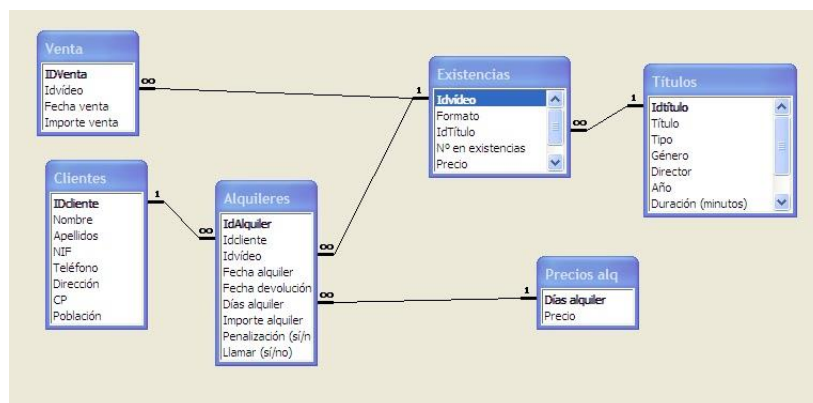
- El modelo de datos **jerárquico** describe un conjunto de relaciones uno a muchos (1:N) entre un padre y sus segmentos hijos.



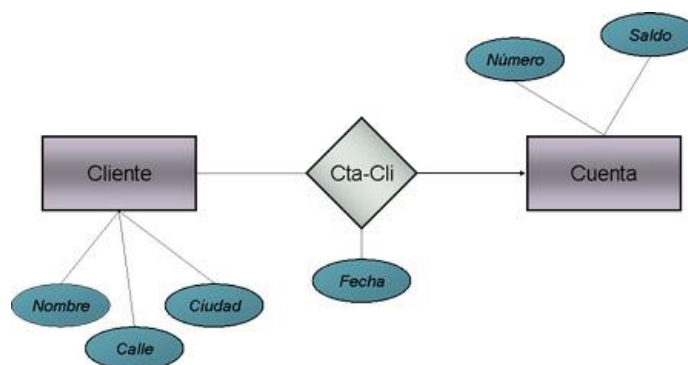
- El modelo de datos **de red** usa conjuntos para representar relaciones 1:N entre tipos de registros.



- El modelo de datos **relacional** es el estándar actual de implementación de bases de datos. En este modelo, el usuario final percibe los datos como almacenados en tablas. Las tablas están relacionadas entre sí por medio de valores comunes en atributos comunes. El modelo entidad-relación (ER) es una popular herramienta para modelar datos que complementa al modelo relacional. El modelo ER permite a diseñadores de BD presentar visualmente diferentes vistas de los dato.

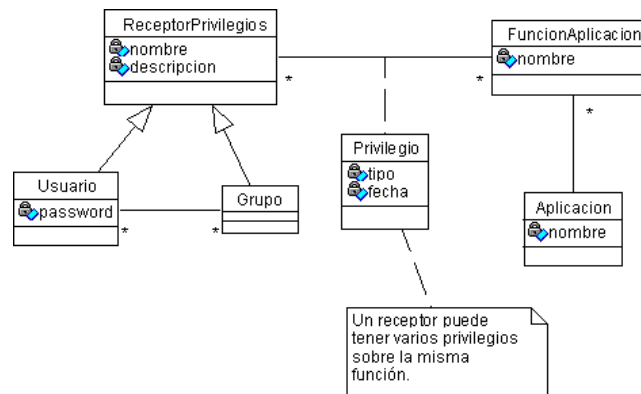


Modelo Relacional



Modelo Entidad-Relación

- El modelo de datos **orientado a objetos** (OODM) usa objetos como estructura básica. Un objeto se asemeja a una entidad en la que incluye los datos que lo definen.



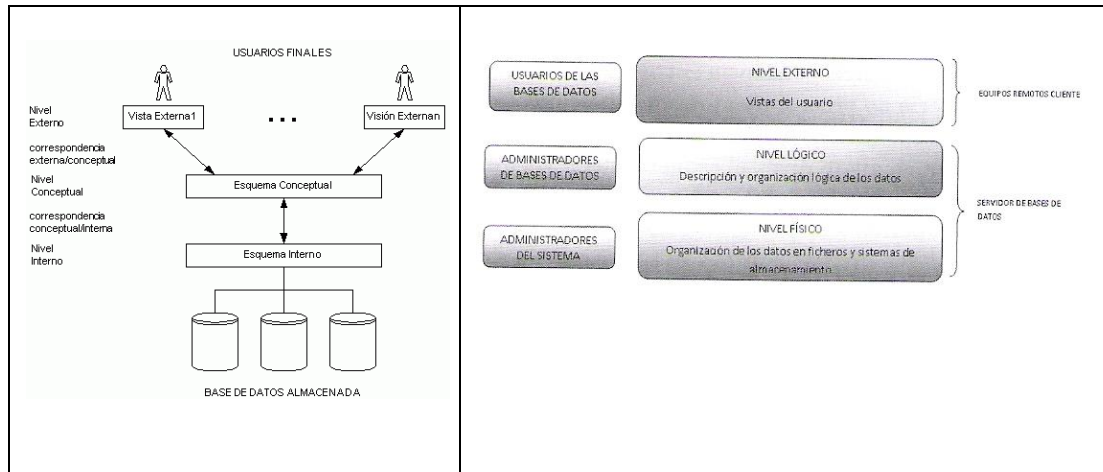
- El modelo de datos **objeto/relacional**, es una extensión de la base de datos relacional tradicional, a la cual se le proporcionan características de la programación orientada a objetos (POO).
- El modelo de datos **XML**, para administrar datos no estructurados dentro del formato XML¹.

6. Grados de Abstracción de Datos

A principios de los 1970, el grupo de trabajo **SPARC** del comité ANSI (comúnmente ANSI/SPARC), definió tres niveles de abstracción de datos que son:

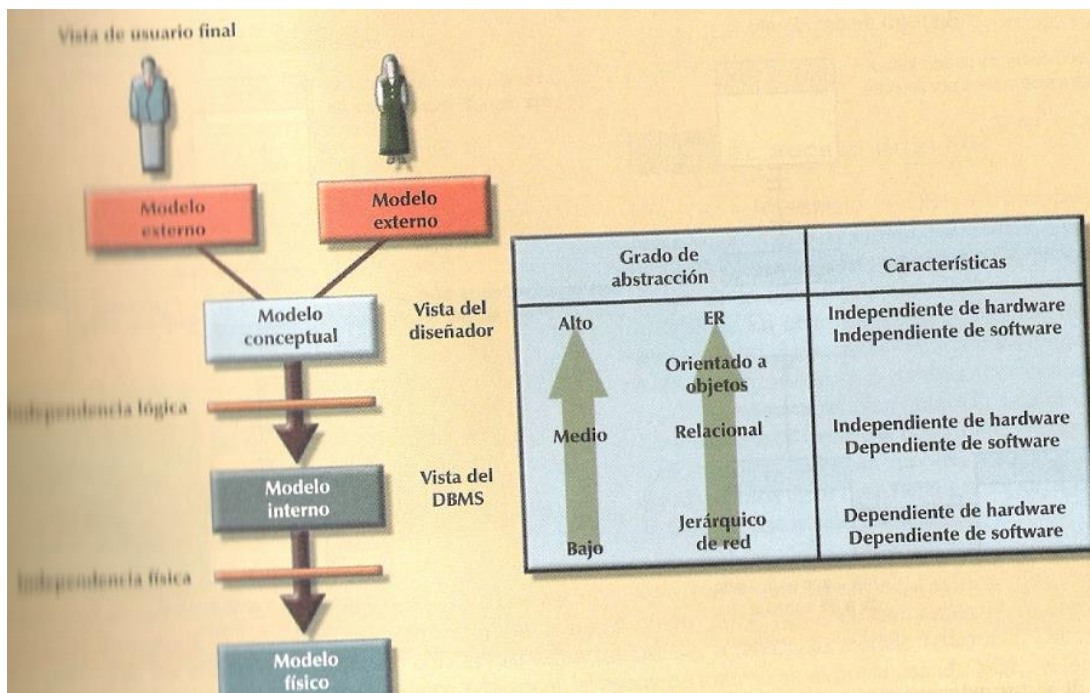
- Modelo externo**, es la vista que tienen los usuarios finales del ambiente de datos. Cada vista externa describe la parte de la base de datos que interesa a un grupo de usuarios determinados y oculta a ese grupo el resto de la base de datos.
- Modelo conceptual**, se describe la estructura de toda la base de datos para una comunidad de usuarios mediante un esquema conceptual. Este esquema oculta los detalles de las estructuras de almacenamiento y se concentra en describir entidades, atributos, relaciones de los usuarios y restricciones. El modelo **Entidad/Relación** y **Modelo Relacional** sigue siendo el modelo más popular.
- Modelo físico**, se describe la forma en la que los datos se guardan en medios de almacenamiento como discos o cintas. estructura física de la base de datos mediante un esquema interno. Este esquema se especifica mediante un modelo físico y describe todos los detalles para el almacenamiento de la base de datos, así como los métodos de acceso.

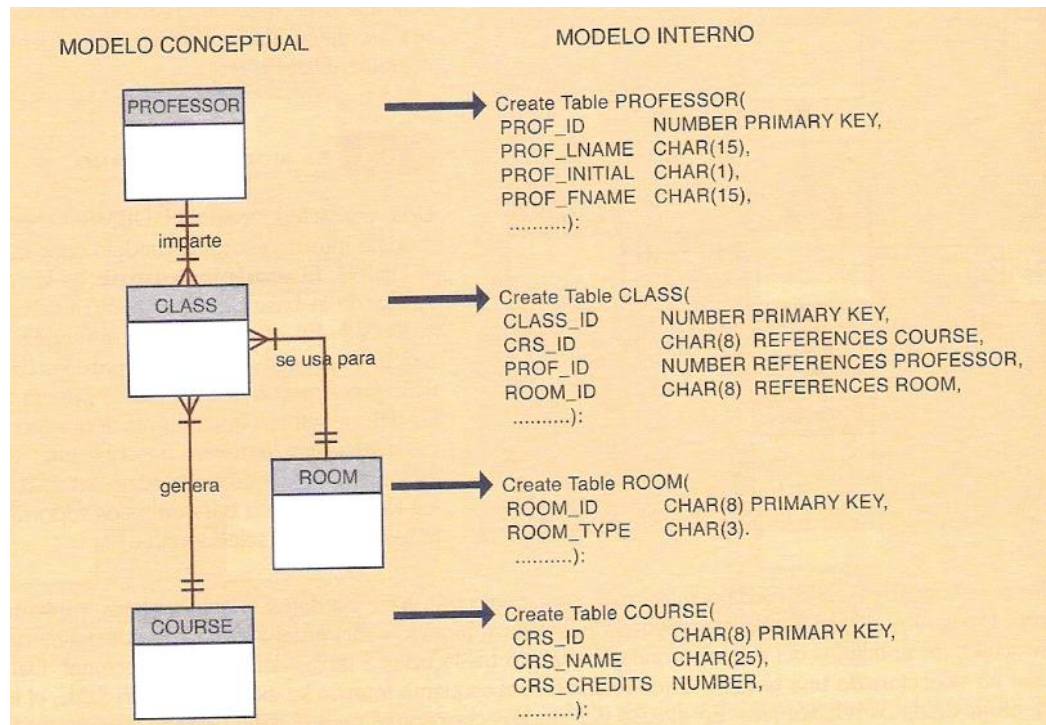
¹ Lenguaje de Marcado Extensible



La arquitectura de tres niveles es útil para explicar el concepto de **independencia de datos**, que podemos definir como **la capacidad para modificar el esquema un nivel del sistema sin tener que modificar el esquema del nivel inmediato superior**. Se pueden definir dos tipos de independencia de datos:

- La **independencia lógica** es la capacidad de modificar el esquema conceptual sin tener que alterar los esquemas externos ni los programas de aplicación. Se puede modificar el esquema conceptual para ampliar la base de datos o para reducirla. Si, por ejemplo, se reduce la base de datos eliminando una entidad, los esquemas externos que no se refieran a ella no deberán verse afectados.
- La **independencia física** es la *capacidad de modificar el esquema interno sin tener que alterar el esquema conceptual (o los externos)*. Por ejemplo, puede ser necesario reorganizar ciertos ficheros físicos con el fin de mejorar el rendimiento de las operaciones de consulta o de actualización de datos. Dado que la independencia física se refiere solo a la separación entre las aplicaciones y las estructuras físicas de almacenamiento, es más fácil de conseguir que la independencia lógica.





7. Ejercicios

Ejercicio 1. Usando el contenido de las figuras siguientes:

- Identifica el Nombre de las entidades.
- Identifica el Atributos de cada entidad.
- Identificar cada tipo de relación aplicando la terminología (1:1, 1:N ó N:M).
- Escribir todas las reglas de negocios.

