

UF2213: Modelos de datos y visión conceptual de una
base de datos

Elaborado por: Salvador Trujillo León

Edición: 5.1

EDITORIAL ELEARNING S.L.

ISBN: 978-84-16360-69-7

No está permitida la reproducción total o parcial de esta obra bajo cualquiera de sus formas gráficas o audiovisuales sin la autorización previa y por escrito de los titulares del depósito legal.

Impreso en España - Printed in Spain

Presentación

Identificación de la Unidad Formativa

Bienvenido a la Unidad Formativa **UF2213: Modelos de datos y visión conceptual de una base de datos**. Esta Unidad formativa pertenece al Módulo Formativo **MF0966_3: Consulta y manipulación de información contenida en gestores de datos**, que forma parte del Certificado de Profesionalidad **IFCD0211: Sistemas de gestión de información**, de la familia de **Informática y Comunicaciones**.

Presentación de los contenidos

La finalidad de esta unidad formativa es enseñar al alumno a interpretar las especificaciones de diseño definidas en los sistemas gestores de datos, utilizados por la organización para el almacenamiento de la información.

Para ello, en primer lugar se analizará el modelo de datos conceptual, la introducción a las bases de datos y el análisis detallado del modelo relacional. También se estudiarán los modelos avanzados de BD, y por último, se profundizará en el análisis detallado de la distribución de BD.

Objetivos de la Unidad Formativa

Al finalizar esta Unidad Formativa aprenderás a:

- Identificar las tipologías de gestores de datos y sus correspondientes modelos de datos según unas especificaciones de diseño dadas.

Índice

UD1. Modelo de datos conceptual

1.1. Conceptos básicos	9
1.1.1. La realidad: los objetos	11
1.1.2. Las concepciones: la información	12
1.1.3. Las representaciones: los datos	13
1.2. Características generales de un modelo	15
1.3. Modelo ER (entity-relationship)	17
1.3.1. Construcciones básicas.....	18
1.3.2. Extensiones.....	36
1.4. Modelo UML	44

UD2. Introducción a las bases de datos

2.1. Concepto y origen de las BD y SGBD.....	67
2.2. Evolución	77
2.3. Objetivos y servicios.....	90
2.4. Modelo lógico BD.....	100
2.4.1. Modelo jerárquico	101
2.4.2. Modelo en red	104

2.4.3. Modelo relacional	107
2.4.4. Modelo relacional extendido	111
2.4.5. Modelo orientado a objetos	113
 UD3. Análisis detallado del modelo relacional	
3.1. Estructura de los datos	129
3.2. Operaciones del modelo	140
3.3. Reglas de integridad	141
3.4. Algebra relacional	153
3.5. Transformación del modelo ER	173
3.6. Limitaciones	193
 UD4. Modelos avanzados de BD	
4.1. BD deductivas	213
4.3. BD geográficas	221
4.4. BD distribuidas	228
4.5. BD analíticas (OLAP)	229
4.6. BD de columnas	230
4.7. BD documentales	232
4.8. BD XML	233
4.9. BD incrustadas (embedded)	236
4.10. Nuevas tendencias	237
 UD5. Análisis detallado de la distribución de BD	
5.1. Formas de distribución	247
5.2. Arquitectura ANSI/X3/SPARC	250
5.3. Transacciones distribuidas	257
5.4. Mecanismos de distribución de datos	264
 Glosario	273
 Soluciones	277

UD1

Modelo de datos conceptual

UF2213: Modelos de datos y visión conceptual de una base de datos

- 1.1. Conceptos básicos
 - 1.1.1. La realidad: los objetos
 - 1.1.2. Las concepciones: la información
 - 1.1.3. Las representaciones: los datos
- 1.2. Características generales de un modelo
- 1.3. Modelo ER (entity-relationship)
 - 1.3.1. Construcciones básicas
 - 1.3.2. Extensiones
- 1.4. Modelo UML

1.1. Conceptos básicos

Una base de datos podemos definirla como una colección de datos relacionados entre sí, donde los usuarios pueden obtener información sobre estos.

El objetivo de un sistema de base de datos es proporcionar a los usuarios una visión abstracta de los datos, para que los usuarios puedan manejarlos y obtenerlos correctamente, sin tener que conocer cómo se almacenan y mantienen estos datos en el sistema.

Existen tres niveles de abstracción para simplificar la usabilidad de los usuarios con la base de datos y el sistema:

- **Nivel físico:** es el nivel más bajo de abstracción, en él se especifica la forma en que se va a almacenar los datos en los dispositivos de almacenamiento.
- **Nivel conceptual:** es el nivel medio de abstracción, en él vemos qué datos son almacenados en la base de datos y las relaciones que existen entre estos y su estructura. Este nivel lo realizan los administradores de bases de datos, ya que son los que deben decidir qué información se va a guardar en la base de datos.

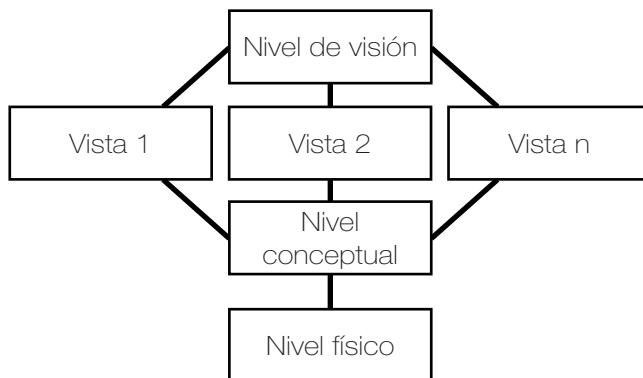
Dentro de este nivel, encontramos los siguientes conceptos:

- **Definición de los datos:** se describen el tipo de datos y la longitud del campo de los elementos de la base de datos. Estos elementos incluyen las entidades y los atributos de las entidades.
- **Relaciones entre datos:** se definen las relaciones entre datos para enlazar los tipos de registros relacionados.

En el nivel conceptual la base de datos aparece como una colección de registros lógicos.

La transformación de registros conceptuales a registros físicos lo lleva a cabo el sistema y es transparente para el usuario.

- **Nivel de visión:** es el nivel más alto de abstracción, el usuario puede visualizar la base de datos en el sistema. El sistema puede proporcionar muchas visiones para la misma base de datos.



Una vez vistos estos conceptos, podemos hablar del modelo de datos conceptual, pero antes debemos conocer qué significa este término.

Un modelo de datos consiste en simbolizar la información del mundo real para que sea más fácil su utilización.

Este modelo debe describir la estructura de los datos, como las relaciones existentes entre estos, o sus restricciones y su significado.



Un **modelo de datos** “Intenta plasmar información del mundo real, que será almacenado y utilizado por un sistema gestor de bases de datos, conocido por sus siglas SGBD”.

Una vez hecho este análisis de los datos, debemos realizar un esquema, para esto tenemos que llevarlo al modelo conceptual. Este esquema no debe tener ninguna correlación con la implementación del esquema, ni debe incluir elemento relacionado con la futura base de datos.

1.1.1. La realidad: los objetos



Un **esquema** “Consiste en describir la información de forma específica, para poder ser implementado en una base de datos”.

Con la aparición de los sistemas gestores de bases de datos, surgen metodologías de análisis de datos.

La idea central del análisis de datos es la representación de una parte del mundo real, esta representación se realizará acorde a una estructura y una organización mediante un sistema gestor de base de datos (SGBD) dado.

Esta representación deberá contener todos los datos necesarios para la propia organización.

El primer paso será examinar todos los objetos de la organización sobre los que queramos mantener información en el sistema, debemos identificarlos, describirlos e incluirlos en el diccionario de datos del sistema gestor de base de datos (SGBD).

Pero para realizar esto debemos conocer qué es un objeto.



Un **objeto** es “Todo lo que puede ser materia en la realidad, es decir, cualquier cosa”.

Estos objetos pueden ser dos tipos:

- **Inmutables**: son objetos reales.
- **Mutables**: son objetos definidos por el usuario.

Estos objetos se pueden estructurar, independientemente del tipo que sean, en listas, en vectores, en colecciones, etc.

Cuando los objetos se agrupan con la misma estructura, se denomina clase, cuando se forma una clase cada objeto se llama instancia.

El objetivo final del análisis de datos será la implementación y almacenamiento de todos los datos en un sistema gestor de base de datos (SGBD).

1.1.2. Las concepciones: la información

Con el fin de facilitar información a los usuarios, hay que realizar diversas tareas con los datos. Pero antes veamos qué es la información.



La **información** es un "Significado que se atribuye a los datos a partir de las reglas convencionales utilizadas para su representación, para una futura interpretación y uso por parte del usuario".

Los datos son los elementos que permiten la transferencia de información entre los usuarios o los distintos sistemas de una empresa.

Esta transferencia de información se conoce como sistemas de información, pero se suele utilizar sus siglas S.I., o I.S., de InformationSystems..

Podemos identificar como componentes de un sistema de información los siguientes elementos:

- **Los datos:** son la información relevante que se almacena y gestiona en el sistema.
- **El hardware:** es el equipamiento físico de un sistema informático, es decir todo aquello que es tangible.
- **El software:** es el equipamiento lógico, estos componentes lógicos incluyen, entre muchos otros, las aplicaciones informáticas que permiten el funcionamiento del sistema, como el propio sistema operativo, que permite al resto de los programas funcionar adecuadamente, facilitando también la interacción entre los componentes físicos y el resto de las aplicaciones, y proporcionando una interfaz con el usuario.

- **Los usuarios:** son los responsables de manejar todos los recursos del sistema de información.

Hay dos tipos de sistemas de información de gestión de datos:

- **Sistemas de información orientados al proceso o sistema clásico de ficheros:** en estos sistemas hay distintas aplicaciones para gestionar diferentes aspectos del sistema de información. Este tipo de sistema tiene muchos inconvenientes, y como única ventaja destacaría que al ser los procesos independientes la modificación de uno no afecta al resto.
- **Sistemas de información orientados al dato o sistema de bases de datos:** tiene grandes ventajas frente al anterior, entre ellas la independencia de los datos, menor duplicidad de los datos, menor espacio de almacenamiento, entre otras. Pero también encontramos inconvenientes en este sistema, ya que su instalación es costosa, se requiere un personal cualificado, su implantación es larga y difícil y se tiene una ausencia de estándares reales.

1.1.3. Las representaciones: los datos

Podemos definir un dato en informática como cualquier elemento informativo que tenga relevancia para un usuario. El sistema de una entidad se suele dividir en varios subsistemas para tratar la información, para manejar esta enorme cantidad de información se emplean los sistemas de información.

Los sistemas actuales manipulan la información a través de las bases de datos y de los sistemas gestores de bases de datos (SGBD), que son elementos indispensable en la actualidad, ya que para obtener cualquier tipo de información se utilizan estos elementos, aunque no nos damos cuenta de que las estamos utilizando.



Los **datos** son una “Representación de hechos, conceptos o instrucciones, realizada de una manera estructurada, apta para su comunicación, interpretación, bien por los usuarios o bien por los sistemas informáticos, manipulados por software específico y representado en diferentes formas”.

Como vemos, los datos son elementos que sirven de base para resolver problemas, pero en sí mismo, un dato tiene poco valor. Los datos permiten la obtención de información cuando están clasificados, almacenados y relacionados entre sí.

Esta información se obtiene a través del procesamiento de datos, que consiste en acumular, agrupar y cruzar datos para transformarlos en información o para obtener otra información.

El procesamiento de datos puede ser:

- **Manual:** cuando la introducción de datos se efectúa manualmente, utilizando la ficha correspondiente.
- **Semiautomático:** cuando la introducción de datos se realiza mediante una máquina de contabilidad en la cual el usuario introduce las fichas, una tras otra, y después de recibir la ficha y los datos iniciales la máquina realiza numerosas operaciones consecutivas ya programadas sin la intervención del usuario.
- **Automático:** cuando la introducción de datos se realiza mediante un sistema informático para que realice todo el conjunto de operaciones, sin la necesidad de la intervención humana entre los procesos.

A día de hoy, recopilar información es fundamental para cualquiera entidad, ya que esta información sirve para que puedan crecer y así adelantar a la competencia, por esto motivo, cada entidad tendrá más de una base de datos, por eso es importante el mantenimiento y crecimiento de estas.

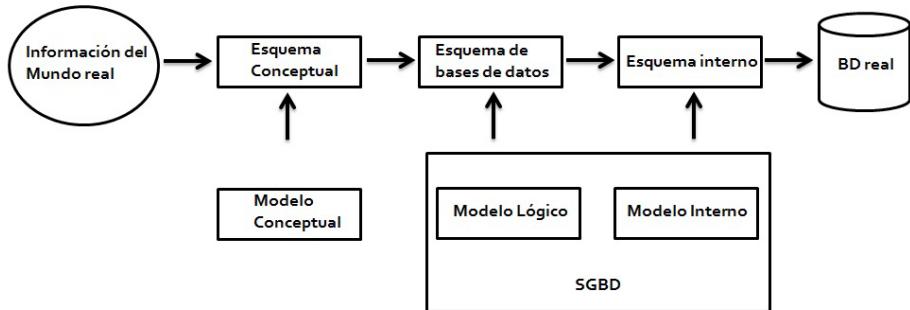
En general, las bases de datos son dinámicas, ya se modifican constantemente, por otro lado están las bases de datos estáticas, que solo recopilan información o documentos históricos.

A continuación, veremos un primer concepto de que es una base de datos.



Una **base de datos** podemos decir que "Es una colección de datos relacionados."

1.2. Características generales de un modelo



Estructura de los modelos de datos.

Información del mundo real	En ella está toda la información que vemos en el mundo real como personas. Es el primer concepto que debemos tomar como punto de inicio.
Esquema conceptual	Simboliza el modelo de datos, independientemente del sistema gestor de bases de datos que se vaya a utilizar.
Esquema de bases de datos	Consiste en la representación de los datos, de una forma más cercana al sistema informático. También es conocido como esquema canónico.
Esquema interno	Consiste en la representación de los datos en un modelo concreto, según el sistema gestor de bases de datos utilizado.
Base de datos real	Consiste en la representación y almacenamiento de los datos, en un dispositivo de almacenamiento físico.

Para poder pasar de un simple esquema a una base de datos real debemos seguir unos pasos.

Estos pasos son los llamados modelos de datos. Estos modelos de datos tienen pasos concretos para poder ir de un esquema a una implantación en una base de datos.

Los dos modelos de datos utilizados son el conceptual y el lógico.

- **El modelo conceptual:** consiste en organizar un esquema teórico de los datos, ya que será necesario para poder pasarlo de la forma real a la forma correcta para un sistema informático. Esta información será almacenada y gestionada por un sistema gestor de bases de datos (SGBD). En este modelo encontramos distintos modos de utilizarlo como el modelo entidad-relación (E/R), el modelo RM/ T y el modelo semántico.
- **El modelo lógico:** consiste en la representación de los datos dirigida a un sistema gestor de bases de datos (SGBD). En este modelo encontramos distintos modos de utilizarlo como el modelo jerárquico, el modelo en red, el modelo relacional, el modelo relacional extendido y el modelo orientado a objetos.

Ahora veremos algunas diferencias entre el modelo lógico y el modelo conceptual.

- El modelo conceptual es ajeno al sistema gestor de bases de datos, en cambio el modelo lógico depende del sistema gestor de bases de datos que se vaya a utilizar.
- El modelo conceptual es más cercano a las personas, por contraposición el modelo lógico es más cercano al sistema informático.

Después de elegir el modelo podemos implementar el diseño, para poder implementarlo, debemos seleccionar un sistema de gestión de bases de datos.

Para poder crear y diseñar una buena base de datos debemos seguir unas series de objetivos, como:

Base de Datos	Debe de ser flexible ante posibles cambios de la información
	El diseño debe ser fácilmente entendible
	Debe ser accesible para que se facilite la entrada a los datos
	Debe optimizarse para el almacenamiento
	Debe implementarse seguridad para los datos
	Debe optimizarse para un alto rendimiento

Como vemos hay varios objetivos que se contraponen, por lo que debemos conseguir una mezcla de estos objetivos, para poder realizar una base de datos correctamente.

Unos de estos objetos que claramente se contraponen son el de flexibilidad y rendimiento.

Por un lado la flexibilidad implica la no utilización directa de un sistema gestor de bases de datos (SGBD), mientras que el rendimiento tiene que cumplir con el estándar del sistema gestor de bases de datos (SGBD).

1.3. Modelo ER (entity-relationship)



El modelo de entidad-relación es conocido como E/R, estas siglas provienen del inglés entity-relationship y fue creado por Peter Chen en los años 1976 y 1977.

Se trata de un modelo que sirve para crear esquemas conceptuales de bases de datos, y es prácticamente un estándar. También es conocido como El (Entidad / Interrelación).

Inicialmente sólo se incluían los conceptos de entidad, relación y atributos. Despues se añadieron otras propuestas (superclase, subclase, especialización, generalización, herencia, categoría y agregación) que forman el llamado modelo entidad relación extendido, conocido por sus siglas (ERE).

Lo primero que debemos de hacer es elaborar un análisis de los datos, para obtener el diseño conceptual de los datos. Y así poder pasarlo al método más utilizado, que es el de entidad-relación (E/R).

Los puntos para realizar el modelo entidad-relación son los siguientes:

- Encontrar las entidades.
- Encontrar las relaciones.
- Encontrar los atributos y asociarlos a entidades y relaciones.
- Especificar las relaciones.
- Buscar los identificadores.

- Especificar los roles.
- Especificar las restricciones.
- Especificar las cardinalidades.
- Encontrar las claves.
- Dibujar el diagrama entidad-relación.
- Revisar el esquema conceptual.

Para entender bien este modelo, debemos conocer antes estas definiciones, para poder elaborarlo correctamente.

1.3.1. Construcciones básicas

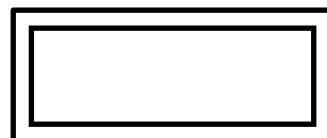
- **Entidad:** no es un dato como tal, pero representa a uno, y en ella se guarda la información del objeto real.

Existen dos tipos de entidades:

- **Fuertes:** son entidades normales que tienen existencia por sí mismas sin depender de otras. Su representación gráfica es de la siguiente forma, un rectángulo.



- **Débiles:** su existencia depende de otras. Las entidades débiles se representan de la siguiente forma, un rectángulo dentro de otro.



Dentro de las entidades débiles, podemos hablar de la existencia de dos:

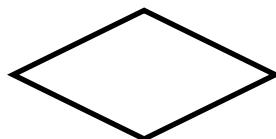
- › **Dependencia por existencia:** la entidad débil puede identificarse sin necesidad de identificar la entidad fuerte, mediante un atributo identificador clave.
- › **Dependencia por identidad:** la entidad débil no puede ser identificada sin la entidad fuerte relacionada.

Para entenderlo mejor veamos un ejemplo:



- **Relación:** describe el vínculo existente entre dos o más entidades, es decir, es el concepto que permite relacionar las dos entidades.

La representación gráfica de una relación se haría de la siguiente forma, un rombo.



Continuando con el ejemplo, la relación que asocia el "tren" con el "vagón" sería "contiene".

La representación gráfica quedaría de la siguiente forma, uniéndose con una línea, que se denomina conector.

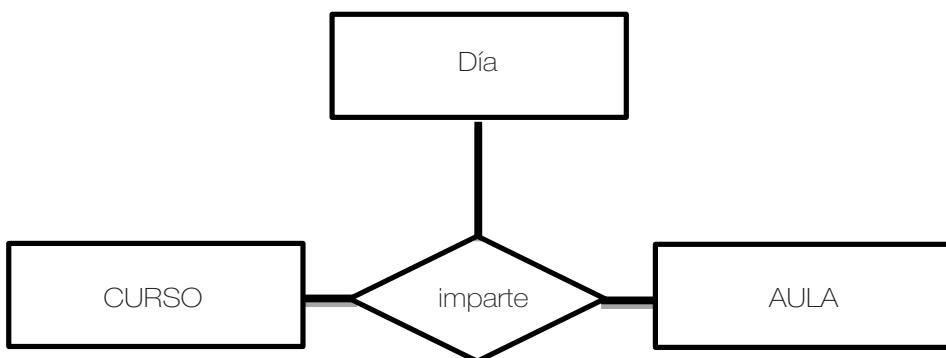


Otros ejemplos de relaciones serían las siguientes:

- Relación binaria:



- Relación ternaria:



- Relación doble:

