



ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN	2
A) ¿QUÉ ES UN DATO?	2
B) ¿CUÁNDO UN DATO ES INFORMACIÓN?	3
2. ALMACENAMIENTO DE LA INFORMACIÓN: LOS FICHEROS	3
A) DEFINICIÓN DE FICHERO	3
B) SISTEMAS DE FICHEROS	4
C) INCONVENIENTES DE LOS SISTEMAS DE FICHEROS.	5
3. BASES DE DATOS	6
A) CONCEPTO DE BASE DE DATOS	6
B) VENTAJAS E INCONVENIENTES DE UTILIZAR BD	6
C) OBJETIVOS DE LA ORGANIZACIÓN DE UNA BASE DE DATOS.	8
D) ARQUITECTURA DE UNA BASE DE DATOS: ANSI/SPARC	10
E) MODELOS DE DATOS	14
4. SISTEMAS GESTORES DE BASES DE DATOS	17
A) OBJETIVOS DEL SGBD	18
B) FUNCIONES DE LOS SGBD	18
C) COMPONENTES DE UN SGBD	19



1. Introducción

a) ¿Qué es un dato?

Existen muchas definiciones de “Dato” dependiendo del ámbito en el que nos movamos:

- Desde el punto de vista **humanístico** un dato es una expresión mínima de contenido sobre un tema.

Ejemplos de datos son: la altura de una montaña, la fecha de nacimiento de un personaje histórico, el peso específico de una sustancia, el número de habitantes de un país, etc

- En un contexto **empresarial** un dato se entiende como un registro de transacciones.
- Desde el punto de vista **científico** y **tecnológico** el dato es una representación simbólica (numérica, alfabética, algorítmica etc.), de un atributo o una característica de un elemento de la realidad. Son datos por ejemplo: 3, alto, j=5, rojo...

Como se puede ver en cualquiera de los tres casos y en cualquier otro caso que pensemos, un dato no dice nada sobre el por qué de las cosas, y por sí mismo tiene poca o ninguna relevancia o propósito. Así que podríamos definir **dato** como:

“Un pequeño trozo de información que no tiene significado para los humanos.”



b) ¿Cuándo un dato es Información?

Pero no cualquier conjunto de datos ordenados constituye en sí una **información**. Es decir, no cualquier cosa que nos digan, leamos o escribamos tiene que ser información. Para que la unión de una serie de datos se considere información debe tener al menos las siguientes características:

- **Significado** (semántica): ¿Qué quiere decir? Del significado extraído de una información, cada individuo evalúa las consecuencias posibles y actúa según lo que deduce del significado de la información.
- **Importancia** (relativa al receptor): ¿Trata sobre alguna cuestión importante? La importancia de la información para un receptor, se referirá a en qué grado cambia la actitud o la conducta de los individuos.
- **Vigencia** (en la dimensión espacio-tiempo): ¿Es actual o desfasada? En la práctica la vigencia de una información es difícil de evaluar, ya que en general acceder a una información no permite conocer de inmediato si dicha información tiene o no vigencia.
- **Validez** (relativa al emisor): ¿El emisor es fiable o puede proporcionar información no válida (falsa)?
- **Valor** (activo intangible volátil): ¿Cómo de útil resulta para el destinatario?

2. Almacenamiento de la información: los ficheros

a) Definición de fichero

Un fichero (o archivo) es una colección de información, localizada o almacenada como una unidad en alguna parte del ordenador.

O de otra forma: los ficheros son el conjunto organizado de informaciones del mismo tipo, que pueden utilizarse en un mismo tratamiento; como soporte material de estas informaciones.

Los ficheros son el soporte para la entrada y salida de datos del ordenador y siempre son manejados por programas: Sistemas Operativos, Gestores de ficheros, aplicaciones de propósito general, etc.



b) Sistemas de ficheros

Para comprender mejor el funcionamiento de los sistemas de ficheros vamos a utilizar un símil. El símil más común de los ficheros es el concepto de los archivadores de fichas de cartón.

- Todo el archivador será un **fichero**.
- Cada una de las fichas es lo que le llamaremos **registro**.
- Cada información independiente de una ficha, será un **campo**.

Para facilitar la búsqueda de una ficha determinada lo más normal es tener el fichero clasificado por alguna de las informaciones que aparecen en la ficha y que tienen valores diferentes en los registros. Por ejemplo, en el archivador de empleados podrían estar ordenadas alfabéticamente por los apellidos, de forma que si queremos buscar la ficha de un empleado bastaría con saber sus apellidos y lo buscaríamos en la posición que le corresponda por orden alfabético.

En una empresa, además, los archivadores están clasificados de forma que los trabajadores saben en cada momento justo al archivador que deben ir para buscar una ficha determinada.

Al igual que en una empresa, en informática, un sistema de archivos es un método para el almacenamiento y la organización de los ficheros y de los datos que contienen, para hacer más fácil la tarea de encontrarlos y leerlos. Los sistemas de ficheros se utilizan en dispositivos de almacenamiento, como son los discos duros, los CDs o los DVDs

Es importante saber que cada Sistema Operativo tiene su propio sistema de ficheros.



c) Inconvenientes de los sistemas de ficheros.

Antes de la llegada de los sistemas de gestión de bases de datos (SGBD), las organizaciones normalmente han almacenado la información usando estos sistemas, pero mantener la información en estos sistemas de ficheros tiene una serie de inconvenientes importantes:

- **Redundancia de datos.** Existen datos que pueden repetirse en diferentes lugares o archivos, esto provoca que, teniendo esa duplicidad de datos, el almacenamiento y el costo de acceso sean más altos.
- **Inconsistencia de datos.** Se presentará porque las copias de los mismos datos en diferentes archivos pueden no coincidir, pues si en un archivo se hicieron cambios de los datos, en los otros archivos donde estaban los mismos datos no son modificados automáticamente.
- **Dificultad en el acceso a los datos.** Cuando se requiere de ciertos datos diferentes de archivos diferentes, la obtención, consulta y modificación de los datos no puede hacerse directamente de forma práctica y eficiente. Tendrían que desarrollarse sistemas de recuperación de datos para realizar esa operación específica, o desarrollar un sistema de recuperación de datos para uso general y ajustarlo de acuerdo a las necesidades.
- **Aislamiento de datos.** Debido a que los datos están dispersos en varios archivos, y los archivos pueden estar en diferentes formatos, es difícil escribir nuevos programas de aplicación para recuperar los datos apropiados.
- **Problemas de atomicidad.** En muchas aplicaciones es crucial asegurar que, cuando ocurra un fallo y sea detectado, se restauren los datos a un estado de consistencia que existía antes del fallo. Es difícil asegurar esta propiedad en un sistema de archivos tradicional.
- **Anomalías en el acceso concurrente.** en estos sistemas un entorno en el que permita a múltiples usuarios actualizar los datos de un mismo archivo simultáneamente puede dar lugar a datos inconsistentes o un estado incorrecto.
- **Problemas de seguridad.** No todos los usuarios de un sistema de bases de datos deberían poder acceder a todos los datos. En estos sistemas es difícil garantizar tales restricciones de seguridad.

Estas dificultades, entre otras, han motivado el desarrollo de los sistemas de bases de datos para resolver estos problemas.



3. Bases de datos

a) Concepto de Base de Datos

Las bases de datos se han constituido como una de las herramientas más ampliamente difundidas en la actual sociedad de la información, utilizadas como fuentes secundarias en cuanto recuperación y almacenamiento de información en todos los campos a nivel científico, social, económico, político y cultural.

Las bases de datos son una herramienta de vital importancia para el desarrollo de la actividad profesional; han sido utilizadas especialmente como fuentes de consulta y de producción de conocimiento por investigadores, científicos y académicos de todas las áreas, que han encontrado en éstas, una herramienta importante para el desarrollo del conocimiento.

Una **Base de Datos** se puede definir como una colección de datos estructurados, organizados, interrelacionados entre sí y almacenados sin redundancias, consiguiendo una independencia de los procesos que los usan, de manera que permite a múltiples usuarios hacer utilizaciones varias y simultáneas de los datos.

b) Ventajas e inconvenientes de utilizar BD

Ventajas de utilizar bases de datos

1. **Mejora en la integridad de datos.** La integridad de la base de datos se refiere a la validez y la consistencia de los datos almacenados.
2. **Mejora en la seguridad.** La seguridad de la base de datos es la protección de la base de datos frente a usuarios no autorizados.
3. **Mejora en la accesibilidad a los datos.** Muchas bases de datos proporcionan lenguajes de consultas o generadores de informes que permiten al usuario hacer cualquier tipo de consulta sobre los datos, sin que sea necesario que un programador escriba una aplicación que realice tal tarea.
4. **Mejora en el mantenimiento gracias a la independencia de datos.** En los sistemas de ficheros, las descripciones de los datos se encuentran inmersas en los programas de aplicación que los manejan. Esto hace que los programas sean dependientes de los datos, de modo que un cambio en su estructura, o un cambio en el modo en que se almacena en disco, requiere cambios importantes en los programas cuyos datos se



ven afectados. Sin embargo, las bases de datos separan las descripciones de los datos de las aplicaciones. Esto es lo que se conoce como independencia de datos, gracias a la cual se simplifica el mantenimiento de las aplicaciones que acceden a la base de datos.

5. **Aumento de la concurrencia.** En algunos sistemas de ficheros, si hay varios usuarios que pueden acceder simultáneamente a un mismo fichero, es posible que el acceso interfiera entre ellos de modo que se pierda información o, incluso, que se pierda la integridad. La mayoría de las bases de datos gestionan el acceso concurrente y garantizan que no ocurran problemas de este tipo.
6. **Mejora en los servicios de copias de seguridad y de recuperación ante fallos.** Muchos sistemas de ficheros dejan que sea el usuario quien proporcione las medidas necesarias para proteger los datos ante fallos en el sistema o en las aplicaciones. Los usuarios tienen que hacer copias de seguridad cada día, y si se produce algún fallo, utilizar estas copias para restaurarlos. En este caso, todo el trabajo realizado sobre los datos desde que se hizo la última copia de seguridad se pierde y se tiene que volver a realizar. Sin embargo, los SGBD actuales funcionan de modo que se minimiza la cantidad de trabajo perdido cuando se produce un fallo.

Inconvenientes utilizar bases de datos

1. **Complejidad.** Los SGBD son conjuntos de programas muy complejos con una gran funcionalidad. Es preciso comprender muy bien esta funcionalidad para poder sacar un buen partido de ellos.
2. **Tamaño.** Los SGBD son programas complejos y muy extensos que requieren una gran cantidad de espacio en disco y de memoria para trabajar de forma eficiente.
3. **Coste económico.** El coste de un SGBD varía dependiendo del entorno y de la funcionalidad que ofrece. Por ejemplo, un SGBD para un ordenador personal puede costar 500 euros, mientras que un SGBD para un sistema multiusuario que dé servicio a cientos de usuarios puede costar entre 10.000 y 100.000 euros. Además, hay que pagar una cuota anual de mantenimiento que suele ser un porcentaje del precio del SGBD.
4. **Coste del equipamiento adicional.** Tanto el SGBD, como la propia base de datos, pueden hacer que sea necesario adquirir más espacio de almacenamiento. Además, para alcanzar las prestaciones deseadas, es posible que sea necesario adquirir una máquina más grande o una máquina que se dedique solamente al SGBD. Todo esto hará que la implantación de un sistema de bases de datos sea más cara.
5. **Coste de la conversión.** En algunas ocasiones, el coste del SGBD y el coste del equipo informático que sea necesario adquirir para su buen



funcionamiento, es insignificante comparado al coste de convertir la aplicación actual en un sistema de bases de datos. Este coste incluye el coste de enseñar a la plantilla a utilizar estos sistemas y, probablemente, el coste del personal especializado para ayudar a realizar la conversión y poner en marcha el sistema. Este coste es una de las razones principales por las que algunas empresas y organizaciones se resisten a cambiar su sistema actual de ficheros por un sistema de bases de datos.

6. **Prestaciones.** Un sistema de ficheros está escrito para una aplicación específica, por lo que sus prestaciones suelen ser muy buenas. Sin embargo, los SGBD están escritos para ser más generales y ser útiles en muchas aplicaciones, lo que puede hacer que algunas de ellas no sean tan rápidas como antes.
7. **Vulnerable a los fallos.** El hecho de que todo esté centralizado en el SGBD hace que el sistema sea más vulnerable ante los fallos que puedan producirse

c) Objetivos de la organización de una Base de Datos.

Una base de datos, independientemente del modo en que esté organizada, debe cumplir los siguientes objetivos para que sea considerada como tal:

- **Versatilidad** para representar La información. Los usuarios que acceden a la base de datos usarán éstos de diferentes maneras, dependiendo de la información que precisen. Así, un programador utilizará un conjunto de datos de distinto modo que un usuario que necesite únicamente realizar una consulta.
- **Desempeño.** Las bases de datos han de atender con la rapidez adecuada las peticiones de datos que se realicen, según el uso que se vaya a hacer de ellos.
- **Redundancia mínima.** En los sistemas de procesamiento de datos tradicionales existía un alto nivel de redundancia, lo que hacía que los mismos datos apareciesen repetidos en varios archivos, incrementando el coste de almacenamiento y ocasionando inconsistencia entre los datos repetidos, ya que, en un momento dado, podían tener valores diferentes. Un objetivo de las bases de datos es eliminar la redundancia, siempre y



cuando esto no implique un aumento de su complejidad ni una reducción en su rendimiento.

- **Capacidad de acceso.** Los usuarios de bases de datos continuamente están pidiendo información acerca de los datos almacenados. Si realizan una consulta y necesitan una respuesta rápida, el sistema de base de datos ha de tener capacidad para responder inmediatamente. Esta capacidad depende de la organización física de los datos. Los diseñadores de bases de datos deben lograr una capacidad de acceso rápido. Con una organización física bien diseñada, los accesos a la base de datos serán rápidos.
- **Integridad.** La información de una base de datos puede ser usada por muchos usuarios y de diferentes maneras; por eso, el sistema de base de datos ha de asegurar que los datos que se almacenen sean los que se tienen que almacenar; es decir, que los valores almacenados cumplan ciertas restricciones definidas con anterioridad, y que los fallos producidos en el sistema no destruyan los datos.
- **Seguridad y privacidad.** La información almacenada en la base de datos es de gran valor para la empresa. Esto explica la necesidad de disponer de procedimientos que protejan los datos frente a fallos del sistema, evitando su pérdida y facilitando su recuperación, o frente a usos indebidos o intencionados por parte de personas no autorizadas. La privacidad hace referencia a que los datos sean accesibles para unos usuarios e inaccesibles para quienes no están autorizados.

La seguridad es muy importante en una empresa. Algunos requisitos necesarios para la seguridad de la base de datos son éstos:

- La base de datos ha de estar protegida contra el fuego, el robo, etcétera.
- Los usuarios deben identificarse antes de acceder a la base de datos y sólo tendrán acceso a aquellos datos para los que se les haya dado permiso.
- Los datos han de poder reconstruirse total o parcialmente cuando se produzca un fallo de sistema hardware o software.
- Los accesos indebidos o erróneos por parte de los usuarios serán controlados.
- **Afinación.** Con el tiempo, el volumen de datos que necesita la empresa va aumentando, por lo que la organización física de éstos ha de ser buena, y es posible que haya que ajustarla para atender a las nuevas



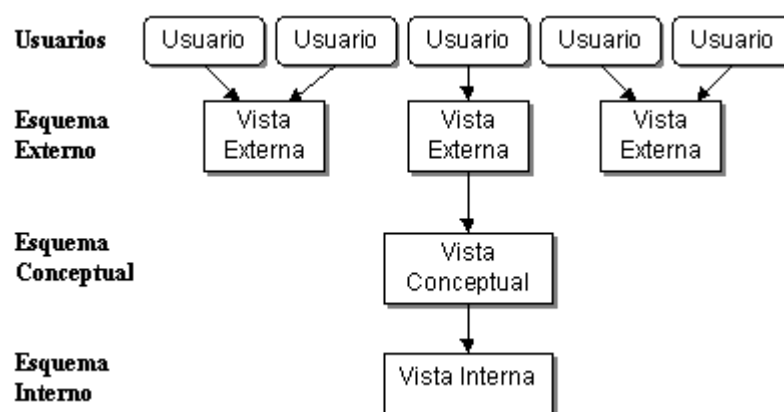
necesidades de almacenamiento. La afinación hace referencia al ajuste de la organización física de los datos con el objeto de mejorar la rapidez para acceder a ellos.

- **Interfaz con el pasado y el futuro.** Con el paso del tiempo, las necesidades de la empresa van cambiando: se precisan nuevos datos, etcétera, por tanto, se hace necesario actualizar el sistema de procesamiento. Cuando se decide instalar un software de base de datos es importante que éste pueda trabajar con los datos ya almacenados. Igualmente, una base de datos ha de estar preparada para los futuros cambios de los datos y de los medios de almacenamiento.

d) Arquitectura de una base de datos: ANSI/SPARC

Uno de los objetivos principales de una BD es proporcionar una visión lo más abstracta posible de la información, es decir, ocultar detalles referentes a la forma en que los datos están organizados y almacenados.

Existen multitud de bases de datos con arquitecturas diferentes en función de sus prestaciones y los sistemas en que se implantan. La arquitectura más estándar y por tanto, la más utilizada, es la realizada por el comité ANSI/SPARC que hace una división en niveles de la BD. Se consideran tres niveles según la perspectiva desde la que se ve la BD: nivel externo, nivel conceptual y nivel interno.



Al organizar la BD en tres niveles de abstracción se pretende:

- Conseguir la separación entre los programas de aplicación y los datos
- Conseguir el manejo de múltiples vistas por parte de los usuarios.



- El uso de un catálogo o diccionario de datos para almacenar el esquema de la base de datos.

El nivel interno es el más bajo de abstracción, dónde se describe la información en función del sistema en que se implante la BD. Después, por encima se encuentra el nivel conceptual, que representa a alto nivel toda la información de la BD independientemente de la máquina dónde se implante la BD. Por último, el nivel externo, gestiona la información desde el punto de vista individual de cada usuario, grupo de usuarios o programadores.

Nivel Externo

Reproduce la visión individual de cada usuario o grupo de usuarios. Se describe mediante el esquema externo (o conjunto de vistas de usuario). Una vista, representa sólo aquella parte de la base de datos que le es importante a un usuario o grupo de usuarios.

Cada vista o esquema:

- Define la parte del esquema conceptual que es de interés.
- Podrá ser compartido por varios usuarios y cada usuario tendrá la posibilidad de acceder a distintas vistas.
- Podrá mezclar campos de distintos registros, omitir campos, cambiar el orden de los campos, añadir campos calculados a partir de otros campos, etc

Nivel Conceptual

Es la representación de toda la información contenida en la base de datos. Son los datos “tal cual son”:

Según ANSI/SPARC:

- Es un nivel intermedio entre el nivel interno y el externo
- Existirá una única vista o esquema conceptual aunque a ella puede haber asociada más de una vista externa
- Se define utilizando el DDL que no tiene en cuenta aspectos de almacenamiento de la información.
- Incluye aspectos como controles de seguridad e integridad



El esquema conceptual se obtiene a partir de los requerimientos de los usuarios potenciales de la BD, sin importar la forma ni el lugar en el que se almacenarán y recuperarán los datos. Contiene los datos elementales, los datos compuestos, las relaciones existentes entre ellos, las reglas que rigen la empresa, etc.

Nivel Interno

El nivel interno es el que trata los aspectos de almacenamiento físico de la información, y recoge la representación de almacenamiento de la información.

El nivel interno se encuentra en el paso anterior a los aspectos físico como pista o cilindro, y es independiente de los dispositivos de almacenamiento, que son tratados por el gestor de ficheros o por el propio dispositivo de almacenamiento.

Correspondencia entre niveles

Hay que destacar que los 3 esquemas no son más que descripciones de los mismos datos pero con distintos niveles de abstracción. Los únicos datos que realmente existen están a nivel físico, almacenados en un dispositivo como puede ser un disco.

En una base de datos basada en la arquitectura ANSI/SPARC, cada grupo de usuarios hace referencia exclusivamente a su propio esquema externo. Por lo tanto, se debe transformar cualquier petición expresada en términos de un esquema externo a una petición expresada en términos del esquema conceptual, y luego a una petición en el esquema interno, que se procesará sobre la BD almacenada. De igual manera, el resultado de la petición tendrá que ser transformada de un formato a otro.

El proceso de transformar peticiones y resultados se denomina **correspondencia**.

Independencia de datos

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 en 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 sólo a la separación entre las aplicaciones y las estructuras físicas de almacenamiento, es más fácil de conseguir que la independencia lógica.

La independencia de datos se consigue porque al modificarse el esquema en algún nivel, el esquema del nivel inmediato superior permanece sin cambios, sólo se modifica la correspondencia entre los dos niveles. No es preciso modificar los programas de aplicación que hacen referencia al esquema del nivel superior.

Por lo tanto, la arquitectura de tres niveles puede facilitar la obtención de la verdadera independencia de datos, tanto física como lógica. Sin embargo, los dos niveles de correspondencia implican un gasto extra durante la ejecución de una consulta o de un programa, lo cual reduce la eficiencia de la base de datos.



e) Modelos de datos

Un **modelo de datos** es una colección de herramientas que servirá para describir los datos, las relaciones, la semántica y las restricciones para la consistencia de las bases de datos.

Teniendo en cuenta que una BD almacena de forma estructurada información relevante para la empresa, la forma en que sea almacenada dará origen a uno u otro tipo de BD, ya que para transformar la información obtenida del mundo real en datos estructurados se habrá empleado un conjunto de abstracciones y reglas existentes previamente. Cada conjunto de reglas y mecanismos de abstracción corresponde a un modelo de datos diferente.

Los modelos de datos se pueden dividir en dos grandes grupos: basados en objetos y basados en registros:

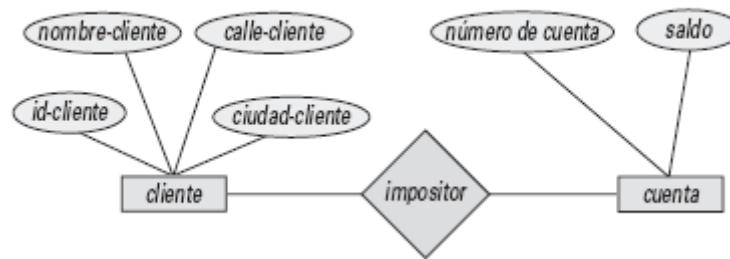
1. Modelos basados en objetos
 - a. Modelo orientado a objetos
 - b. Modelo entidad –relación
 - c. Modelo binario
 - d. Modelo semántico
 - e. Modelo funcional de datos
2. Modelos basados en registros
 - a. Modelo Jerárquico
 - b. Modelo Relacional
 - c. Modelo de Red

Veremos a continuación en detalle algunos de los más importantes:

Modelo entidad-relación

Está basado en la percepción del mundo real que consta de una colección de objetos básicos llamados *entidades*, y de *relaciones* entre estos objetos. Una entidad es una cosa u objeto del mundo real que se distingue de otros objetos. Las entidades se describen en una base de datos mediante un conjunto de **atributos**. Una **relación** es una asociación entre varias entidades. Al conjunto de todas las entidades del mismo tipo se le llama **conjunto de entidades**, y al conjunto de todas las relaciones del mismo tipo se le llama **conjunto de relaciones**.

La estructura lógica de una BD se puede expresar gráficamente mediante un *diagrama E-R*. Por ejemplo:



- **Rectángulos:** Representan conjuntos de entidades.
- **Elipses:** Representan atributos.
- **Rombos:** Representan relaciones entre conjuntos de entidades.
- **Líneas:** Unen los atributos con los conjuntos de entidades y los conjuntos de entidades con las relaciones.

Modelo relacional

Se utilizan tablas para representar los datos y las relaciones entre ellos. Las tablas están compuestas por columnas, y cada columna tiene un nombre único. Por ejemplo:

id-cliente	nombre-cliente	calle-cliente	ciudad-cliente
19.283.746	González	Arenal	La Granja
01.928.374	Gómez	Carretas	Cerceda
67.789.901	López	Mayor	Peguerinos
18.273.609	Abril	Preciados	Valsaín
32.112.312	Santos	Mayor	Peguerinos
33.666.999	Rupérez	Ramblas	León
01.928.374	Gómez	Carretas	Cerceda

Tabla *cliente*

número-cuenta	saldo
C-101	500
C-215	700
C-102	400
C-305	350
C-201	900
C-217	750
C-222	700

id-cliente	número-cuenta
19.283.746	C-101
19.283.746	C-201
01.928.374	C-215
67.789.901	C-102
18.273.609	C-305
32.112.312	C-217
33.666.999	C-222
01.928.374	C-201

Tabla *cuenta*

Tabla de relación *cliente-cuenta*

El modelo relacional está basado en registros. Las columnas de la tabla corresponden a los atributos del tipo de registro. Este modelo es el más ampliamente usado.



Modelo en red

El modelo de datos en red general representa las entidades en forma de nodo de un grafo, y las asociaciones o interrelaciones entre éstas mediante los arcos que unen dichos nodos.

El modelo de red evita esta redundancia en la información, a través de la incorporación de un tipo de registro denominado el conector, que en este caso pueden ser las calificaciones que obtuvieron los alumnos de cada profesor.

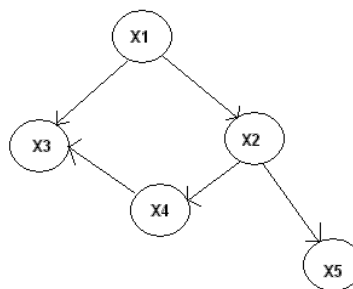
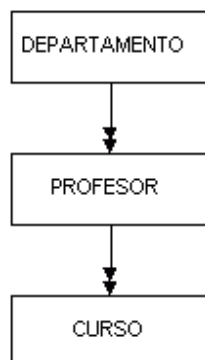


Figura 1. Red Bayesiana

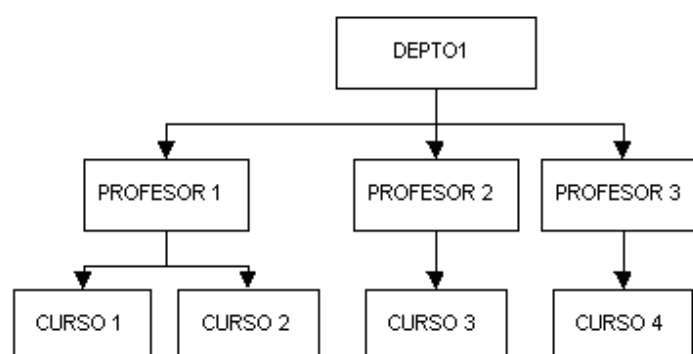
Modelo Jerárquico

En el modelo de red, los datos se representan mediante colecciones de registros y las relaciones entre los datos por medio de enlaces. Esta representación también es válida para el modelo jerárquico. La única diferencia es que el modelo jerárquico los registros se organizan para formar colecciones de árboles en vez de grafos arbitrarios.

Estructura lógica



Ejemplo de base de datos



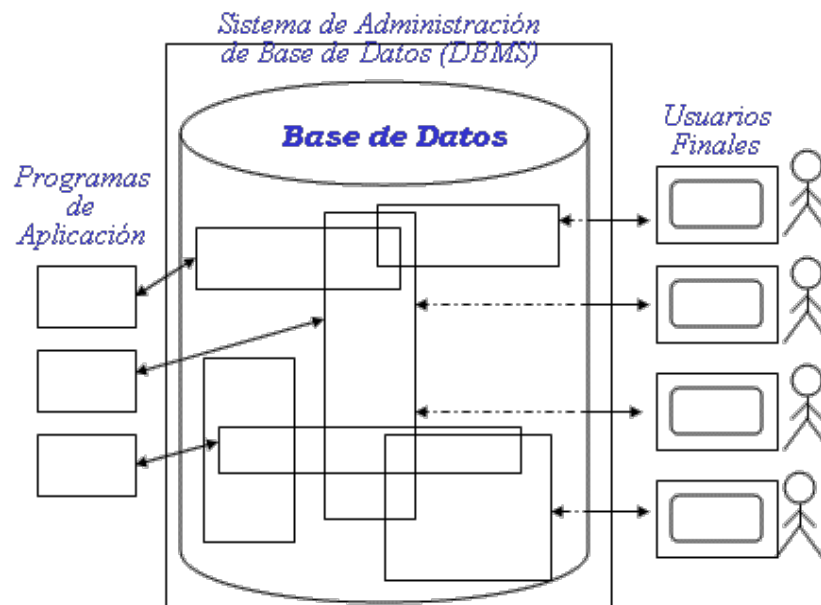


4. Sistemas Gestores de Bases de Datos

SGBD: Colección de programas de aplicación que proporcionan al usuario de la BD los medios necesarios para realizar las siguientes tareas:

- Definición de los datos a los distintos niveles de abstracción.
- Manipulación de los datos de la BD.
- Mantenimiento de la integridad de la BD.
- Control de la privacidad y seguridad de los datos en la BD.

En definitiva, un SGBD aporta al usuario los medios necesarios para el establecimiento de todas las características que una base de datos debe tener.





a) Objetivos del SGBD

Si se tiene en cuenta que en una BD existen una gran variedad de usuario con necesidades diversas y variables a lo largo del tiempo y que estos mismos usuarios podrán trabajar simultáneamente con estos datos, es imprescindible dotar al sistema de la adecuada flexibilidad para que pueda atender las exigencias de todos los usuarios y par que sea capaz de responder a los posibles cambios a un coste no excesivo.

Por tanto un SGBD debe ser capaz de:

- Controlar la redundancia de los datos.
- Atender a múltiples usuarios y a diferentes aplicaciones
- Mantener la independencia tanto física como lógica entro los datos y tratamientos.
- Mantener un único esquema de la base de datos.
- Facilitar la actualización y recuperación de los datos.

b) Funciones de los SGBD

Un SGBD debe cumplir tres funciones fundamentales para ser considerado bueno. A continuación se detallan estas tres funciones:

Función de definición o descripción:

El sistema gestor de bases de datos debe permitir al diseñador de la BD especificar los elementos de datos que la integran, su estructura y las relaciones que existen entre ellos.

Esta función es realizada por el Lenguaje de Definición de Datos (DDL) propio de cada sistema. Este lenguaje debe suministrar los medios necesarios para definir las 3 estructura de datos (externa, conceptual en interna).

A nivel interno, el DDL permitirá indicar el espacio reservado para la base de datos, la longitud de los campos o elementos de datos, su modo de representación, etc.

A nivel externo y conceptual, el DDL proporcionará los instrumentos para la definición de objetos (tablas, registros, etc) así como su identificación, atributos, interrelaciones, accesos, ets.



Función de manipulación:

Una vez creada la base de datos, el SGBD debe aportar al usuario la capacidad de actualizar o recuperar la información.

Para esto, los sistemas gestores de bases de datos cuentan con el Lenguaje de Manipulación de Datos (DML) que aporta las dos funciones de manipulación principales:

- Consulta de los datos
 - Total
 - Selectiva
- Actualización de los datos
 - Inserción
 - Borrado
 - Modificación

Función de control:

Con esta función, el SGBD permite al administrador crear todas las interfaces que necesitan los diferentes usuarios para comunicarse con la BD y proporciona:

- Funciones relacionadas con la seguridad
- Funciones relacionadas con la recuperación.
- Funciones de servicio (cambiar capacidad de ficheros, estadísticas, carga de archivos, etc)

c) Componentes de un SGBD

Para lograr aportar todas las características que se han citado anteriormente, el SGBD cuenta con una serie de componentes que realizan dichas tareas:



Lenguajes

Permiten la definición, manejo y control de los datos de la base de datos. Cada SGBD tiene una estructura y organización particular, pero todos ofrecen al administrador y a los usuarios 3 lenguajes:

- Lenguaje de Descripción de datos (DDL)
- Lenguaje de Manipulación de datos (DML)
- Lenguaje de Control de datos (DCL)

Utilidades

Son aplicaciones que facilitan el trabajo a los usuarios y programadores. Suelen tener un interfaz fácil de entender y se basan en menús que guían a los usuarios para conseguir el objetivo:

- Generador de menús.
- Generador de informes.
- Generador de formularios.

Diccionario de datos

En el diccionario de datos se almacenan todas las definiciones realizadas por el DDL y algunas realizadas por el DML. El diccionario contiene, entre otras cosas:

- Descripción de esquema de la BD
- Restricciones de privacidad y acceso
- Reglas, normas y restricciones de seguridad
- Programas de aplicación.
- Operaciones que realizan los usuarios sobre el sistema.



Administrador de la base de datos (DBA)

Se trata de un componente humano de suma importancia ya que va a tener una serie de responsabilidades en cuanto a la definición, administración, seguridad, privacidad e integridad de la información.

- Las tareas más importantes asignadas al DBA son:
- Definición de los esquemas lógico y físico de la BD.
- Definición de los subesquemas o vistas externas de la BD.
- Control de la privacidad de los datos.
- Mantenimiento de los distintos esquemas y procedimientos.

Usuarios de la base de datos

Además del administrador de la base de datos, existen otros dos tipos de usuarios que siempre aparecen en los SGBD.

- Usuarios terminales: acceden a la base de datos a través de aplicaciones. No tienen especialización
- Usuarios técnicos: desarrollan las aplicaciones que utilizan los usuarios terminales. Son profesionales informáticos.

Sin embargo hay que tener en cuenta, que una única persona puede acceder a la base de datos tanto como usuario Terminal, como usuario técnico, teniendo para ello dos perfiles diferentes en el SGBD.