

U1 BASES DE DATOS

Sistemas de información

- **Sistemas de información:**

Actualmente las organizaciones manejan multitud de datos, procedentes de diversas fuentes: sus aplicaciones de gestión (compras, ventas, nóminas, gestión de stocks,...), información pública del sector, utilización de sus sistemas por parte de los usuarios, etc.

Todos estos datos conforman un activo muy importante para la organización, y transformar estos datos en información puede suponer un valor añadido muy importante para la misma.

Un sistema de información es un conjunto de elementos orientados al tratamiento y administración de datos e información, es decir, un sistema por medio del cual se recopilan datos, se homogeneizan y se transforman en información relevante.

Un sistema de información es especialmente adecuado en entornos complejos, competitivos o muy cambiantes. En estos entornos los cambios son continuos, y la rapidez en la toma de decisiones y la implementación de los cambios es crítica, por lo que los gestores necesitan acceder de forma rápida y eficaz a las fuentes de información que les permitan tomar las decisiones estratégicas más adecuadas.

- **Objetivos: El objetivo de un sistema de información será proveer a la organización de herramientas que le permitan:**

- Automatizar los procesos operativos (cobros, pagos, gestión de entradas y salidas, etc.).
- Proporcionar información oportuna y exacta que sirva de apoyo para la toma de decisiones.
- Obtener ventajas competitivas derivadas de su implantación y uso.

Funciones: Las principales funciones de un sistema de información son las siguientes:

- **Recolección de información:** recoger la información, representarla en un formato adecuado y almacenarla.
- **Tratamiento de la información:** fundamentalmente, aplicar tres operaciones: integración, realización de cálculos y transferir información entre diversas fuentes. La información recogida debe ser clara, precisa, coherente, oportuna y completa.
- **Difusión de la información:** proporcionar a cada usuario la información que necesite. Será necesario identificar qué, cómo, cuándo y a quién se distribuirá la

información.

- **Organización de los sistemas de información.**

Los sistemas de información se organizan en varios niveles, cada uno de los cuales se encarga de una función determinada:

- **Nivel operativo:** sistemas operacionales utilizados en el día a día de la organización (sistema de ventas, sistema de nóminas,...), que son la principal fuente de información relevante para la organización ya que proporcionan información acerca de la propia empresa. Cada departamento o estructura de la organización puede tener su propio sistema operacional, y los usuarios de estos sistemas suelen ser los propios trabajadores de la empresa y los directores operativos.

- **Nivel de conocimiento:** a este nivel corresponden los sistemas orientados a apoyar las operaciones diarias de control. Estos sistemas permiten **explotar la información generada por los sistemas de nivel operativo**. Pertenecen a este nivel los sistemas de inteligencia de negocio (*Business Intelligence*), de minería de datos (*Data Mining*) y herramientas de elaboración de informes (*Reporting*). Los usuarios de estos sistemas tienen un perfil de analistas con un alto conocimiento de algún área específica de la organización.

Business Intelligence se enfoca en proporcionar informes y visualizaciones para ayudar en la toma de decisiones.

Data Mining se enfoca en descubrir patrones de comportamiento y tendencias.

- **Nivel administrativo:** Los usuarios de estos sistemas son los mandos intermedios de la organización, encargados de definir las líneas a seguir en el medio plazo por su unidad organizativa. La información que se maneja en este nivel es más comparativa que descriptiva, y es tanto interna (normalmente procedente del data warehouse) como externa, ya que además de mostrar información relevante es importante contextualizar esa información para poder definir objetivos realistas y medibles. Las herramientas principales a nivel táctico son las herramientas de elaboración de informes y especialmente los cuadros de mando: es una hoja que muestra los datos que tiene una empresa en su base de datos de una forma estratégica e interactiva para que las personas puedan investigar los datos de manera self service, o autoservicio. Están orientados a la toma de decisiones **en un área concreta de la organización** a partir del seguimiento de una serie de indicadores. (PKI)

- **Nivel estratégico:** Los usuarios de estos sistemas son la alta dirección, que trabajan con una visión a medio-largo plazo, definiendo las líneas maestras que debe seguir la organización. La herramienta principal de la dirección en este ámbito son los cuadros de mando integrales, que son cuadros de mando “resumidos” que incorporan información de **todas las unidades de la organización** de forma *precisa y resumida, para facilitar la toma de decisiones estratégicas*.



La información se distribuye por la organización de forma vertical (entre los distintos niveles de la jerarquía) y de forma horizontal (en un mismo nivel entre los distintos departamentos).

Un sistema de información tiene como principal objetivo aportar valor a la información, por lo que **el sistema de almacenamiento de la información** será un componente principal del sistema de información.

Es necesario almacenar la información de forma que se garantice el acceso a la información necesaria en el momento preciso.

Existen varias maneras de almacenar la información:

- Sistemas de ficheros. Un fichero (también denominado archivo) es un conjunto ordenado de datos que tienen entre sí una relación lógica y están almacenados en un soporte de información permanente.

La información se almacena de forma estructurada con el fin de manipular los datos de manera individual. **En los sistemas de ficheros, las aplicaciones acceden**

directamente a los ficheros, por lo que deben conocer su estructura, contenido, y por tanto estos son muy dependientes de las aplicaciones que los utiliza.

Ficheros de texto: ficheros planos codificados como texto. Distintos formatos, lo más usual: ascii y unicode. No solo extensión txt, sino también ini, html, php, java, sql, ... Los ficheros XML son un tipo especial de ficheros de texto que tienen definida una estructura. Los servicios web de internet se comunican con las aplicaciones por medio del intercambio de ficheros XML.

Ficheros binarios: todos los que no son de texto. Se codifican por medio de códigos hexadecimales. Requieren un formato para ser interpretados, si cambias el formato no se podrán abrir correctamente. Cada formato tiene una extensión diferente: imagen (jpg, gif, tiff, ...), video (mpg, avi, mpv), compresión (zip, rar, gz,...), ejecutables (exe, com,...), ofimática (doc, odt, xls,...).

Las BD guardan su información en los ficheros binarios, porque la información almacenada en ellos tiene una estructura lógica y organizada para poder acceder rápidamente.

- Bases de datos: Las Bases de Datos es un tipo de sistema de información que no depende de la aplicación que los utiliza. Las bases de datos también guardan los datos en ficheros pero aquí las aplicaciones no se conectan directamente a los ficheros, hay un software intermedio que se encargará de gestionar las conexiones y peticiones de las aplicaciones, acceder a los ficheros a buscar la información, y enviársela a quien la solicite.

Este software intermedio es lo que se denomina el Sistema Gestor de Bases de Datos (SGBD).

Una base de datos por tanto permite almacenar datos de forma independiente de las aplicaciones que acceden a ellos, y permite realizar sobre ellos operaciones de consulta y manipulación (inserción, modificación y eliminación de datos en las tablas)

- Definiciones:

** Una **base de datos** o **banco de datos** es un conjunto de datos que pertenecen al mismo contexto almacenados sistemáticamente para su posterior uso.*

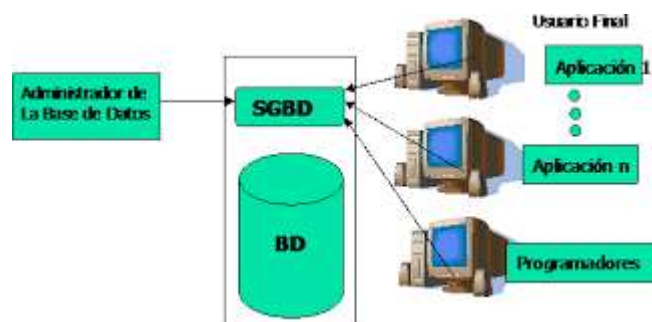
** Una **base de datos** es una colección o depósito de datos, donde estos están lógicamente relacionados entre si, tienen una definición y descripción comunes y están estructurados de una forma particular. Una base de datos es también un modelo del mundo real y como tal debe servir para una gama de usos y aplicaciones.*

Esta definición especifica que la base de datos debe representar fielmente la realidad y sus condiciones.

La base de datos no solo contiene datos, sino también procedimientos personalizados, es decir, no solo guarda los datos, sino también reglas de negocio. Para una empresa es una solución de negocio a escala tecnológica para gestionar la globalidad del sistema de información de una empresa.

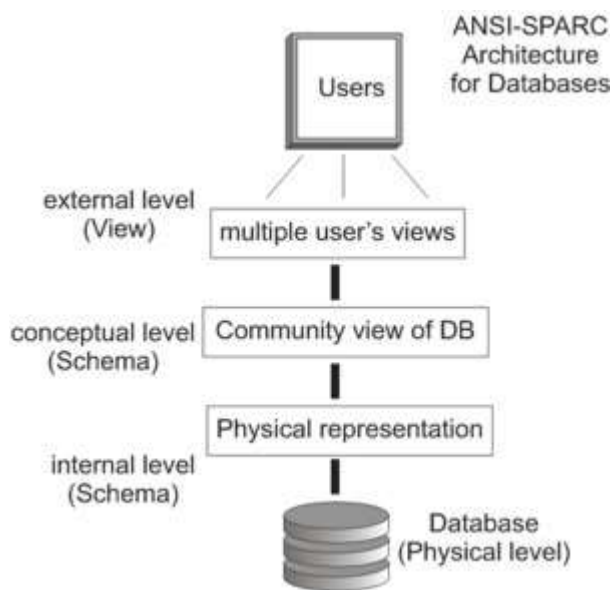
- *Una **base de datos** es una colección o depósito de datos integrado con redundancia controlada y con una estructura que refleje las interrelaciones y restricciones existentes en el mundo real. Los datos que han de ser compartidos por diferentes usuarios y aplicaciones deben mantenerse independientes de estas y su definición y descripción única para cada tipo de datos, han de estar almacenadas junto con los mismos. Los procedimientos de actualización y recuperación, comunes y bien determinados, habrán de ser capaces de conservar la seguridad (integridad, confidencialidad y disponibilidad) del conjunto de los datos.*

Gráfico de una Base de Datos



- Arquitectura de una base de datos: Nivel Interno (físico), Nivel conceptual (lógico) y Nivel externo (vista)

Una de las principales funciones del SGBD es proveer una capa de abstracción entre la información y las aplicaciones que permita gestionar los datos de forma independiente a los programas que acceden a ellos. Para conseguir esta abstracción, se utiliza una arquitectura en 3 niveles, definida en 1975 por la ANSI-SPARC (American National Standard Institute – Standards Planning and Requirements Committee) para conseguir la separación entre aplicaciones y datos, el manejo de múltiples vistas y el uso de catálogos para el almacenamiento de los esquemas de las BD.



Nivel físico: Archivos con toda la información de la BD. Determina cómo se va a guardar la información. El dbms funciona con el SO de la computadora para almacenar los archivos de datos de manera automática.

Nivel lógico: Son las estructuras abstractas de datos, integradas en el nivel físico. Le denominamos **esquema conceptual o modelo lógico**. Determina cómo se va a organizar la información (Primera capa de abstracción)

Nivel externo: (Segunda capa de abstracción) Vistas de usuario. Determina cómo se va a mostrar la información a los usuarios. Aquí los usuarios plantean consultas contra la base de datos. Pueden ser momentáneas o planeadas (vistas).

Los tres niveles no son más que descripciones de los mismos datos pero en distintos niveles de abstracción. Los únicos datos existentes están a nivel físico, almacenados en un dispositivo, como puede ser un disco.

Esta arquitectura favorece 'la independencia de los datos', esto es la capacidad para modificar el esquema en un nivel, sin tener que modificar el esquema del nivel inmediato superior.

Definimos dos tipos de independencia de datos:

- Independencia lógica de los datos.

Son los cambios que se pueden realizar en la capa lógica sin tener que cambiar nada en la capa externa (usuarios y procesos existentes).

- Independencia física de los datos.

Son los cambios que se pueden realizar en la capa física sin tener que cambiar nada en la capa lógica. La separación entre la capa física y lógica de la BD aporta la independencia física de los datos.

Se pueden mover, dividir, combinar objetos, renombrar objetos, agregar, etc. sin tener que modificar la capa lógica. Pero no se pueden eliminar fácilmente.

Es más fácil de conseguir que la independencia lógica.

- Componentes de un sistema de base de datos

Un sistema de base de datos está formado por los siguientes componentes:

- **Datos.** La principal característica de estos sistemas es que la información debe estar integrada (sin redundancias) y podrá ser compartida por muchos usuarios.
- **Equipo** (Hardware). Conjunto de dispositivos físicos utilizados para almacenar y procesar los datos: Ordenadores, Unidades de almacenamiento.
- **Programas** (Software). Un sistema de base de datos incluye dos tipos de programas:
 - El software de propósito general, para la gestión de la base de datos, llamado normalmente Sistema Gestor de Bases de Datos.
 - El software de aplicación.
- **Personal.** En un sistema de base de datos intervienen un número importante de usuarios, que podemos clasificar en tres grupos:

- Informáticos:

*Diseñadores: Realizan el diseño lógico de la BD, debiendo identificar los datos, las relaciones entre ellos y las restricciones sobre ambos. El diseñador debe tener un profundo conocimiento de los datos de la empresa y de sus reglas de negocio.

*Administradores: se encargan del nivel físico relacionado con el almacenamiento.

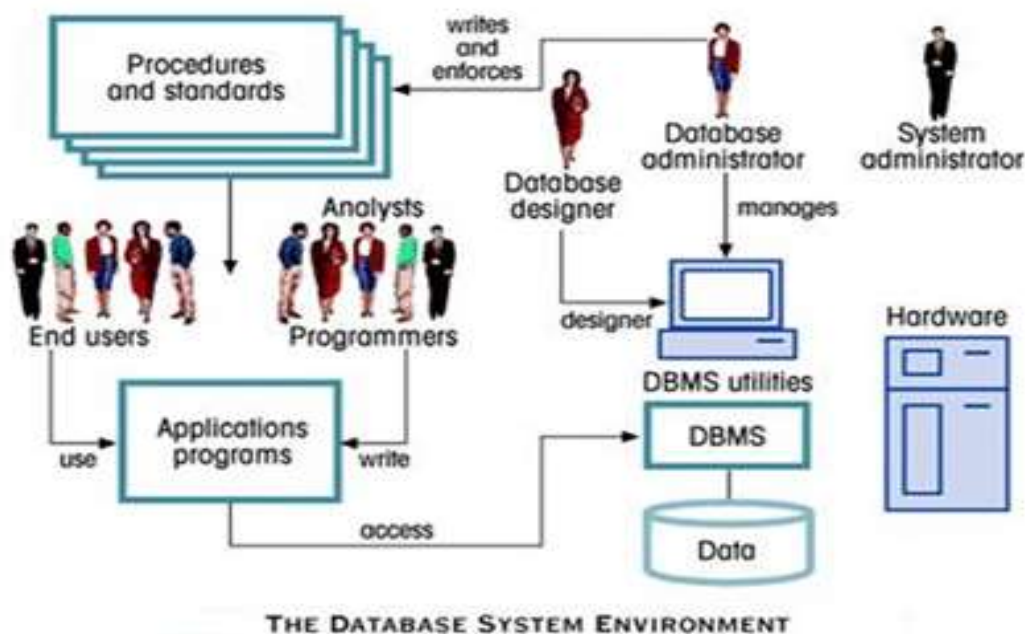
- Administradores del DBMS. Se encargan de instalación y configuración del sistema gestor y, en general, de un buen funcionamiento de este.

- Administradores de la BD. Se encargan del diseño físico, implementación y mantenimiento de la BD.

*Programadores: Realizan las aplicaciones que interaccionan con la BD y objetos de la esta como rutinas y disparadores.

- No informáticos

*Usuarios finales: Clientes de la BD que hacen uso de ella sin conocer su funcionamiento ni organización.



- Tipos de BD según el modelo de datos

Un modelo de BD es una arquitectura organizativa que impone un esquema de acceso a datos. El modelo será un conjunto de conceptos, reglas y convenciones que nos permiten describir los datos.

Los modelos sirven para describir la estructura de una base de datos, es decir los tipos de datos, las relaciones entre ellos y las restricciones que deben cumplir.

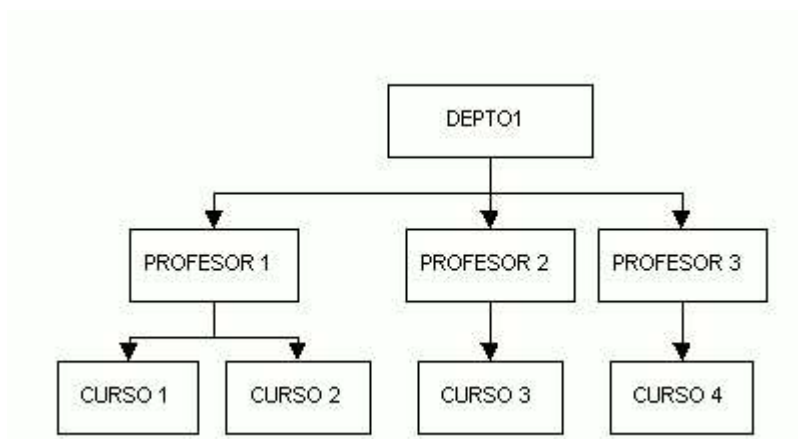
Existen varios tipos de modelos de BD, según el modelo usado tenemos los siguientes tipos de BD:

- **Jerárquicas:** Sustituyeron a los sistemas de ficheros con los registros ordenados

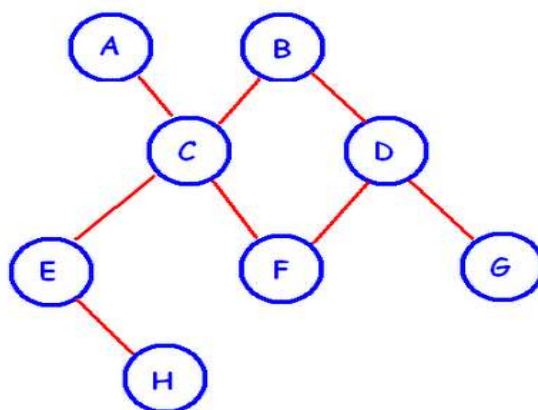
en una jerarquía similar a un organigrama.

Utiliza la estructura lógica de árbol para el almacenamiento de la información. Cada registro principal lleva un apuntador con la dirección o direcciones de los registros relacionados.

Cada elemento primario puede tener varios secundarios, pero uno secundario solo puede tener uno primario. Se eliminan de los nodos secundarios las claves de los primarios, pues ya estos llevan punteros apuntándoles.



En Red: Es la misma filosofía que la anterior pero aquí un elemento secundario puede tener muchos primarios. Es tal la complejidad de las búsquedas y su coste de mantenimiento que dejó de utilizarse enseguida.



En los dos, las búsquedas no normales, son complicadas.

Relacionales: Se basa en que cualquier ruta preconcebida a través de la estructura de datos es demasiado restrictiva. Los usuarios no pueden pensar en cada ruta posible antes de diseñar la base. Este modelo permite relacionar los registros según se requiera y las consultas funcionan como conjuntos de datos. Los datos se presentan en forma de 'tablas' y permiten su combinación formando 'vistas'. Las tablas solo pueden tener un tipo de registro.

Es elegante, sencillo y fácil, proporcionando toda la información posible.

Una base de datos orientada a objetos es una base de datos que incorpora todos los conceptos importantes del modelo de objetos: Encapsulación, Herencia y Polimorfismo.

La base de datos está constituida por objetos, que pueden ser de muy diversos tipos, y por operaciones definidas por los usuarios sobre ellos.

BD XML la principal característica que brindan estas bases de datos es la capacidad de obtener los resultados de las consultas en formato XML.

- Funciones de una Base de Datos

Las principales funciones de una BD como sistema de información son las siguientes:

- Recolección de información: recoger la información, representarla en un formato adecuado y almacenarla.

- Tratamiento de la información: fundamentalmente, aplicar tres operaciones: integración, realización de cálculos y transferir información entre diversas fuentes. La información recogida debe ser clara, precisa, coherente, oportuna y completa.

- Difusión de la información: proporcionar a cada usuario la información que necesite. Será necesario identificar qué, cómo, cuándo y a quién se distribuirá la información.

- Tipos de BD según su uso.

- Documentales
- Geográficas GIS
- Minería de datos (OLAP)
- BD transaccionales. (OLTP)

6. Sistemas de ficheros vs Sistemas de BDs

Ventajas Ficheros

- Menor espacio de almacenamiento
- Acceso más sencillo
- Portabilidad de algunos formatos

Ventajas BDs

- Control sobre la redundancia de datos

- **Control de consistencia**
- **Almacenan datos, metadatos y relaciones entre los datos**
- **Abstracción física y lógica**
- **Mayor disponibilidad**
- **Mayor seguridad**
- **Acceso concurrente**
- **Facilidad de acceso para usuarios**

Desventajas Ficheros

- **Dificultad de mantenimiento**
- **Ausencia de estándares**
- **Gestión manual de duplicados**
- **Dependencia de los programas**
- **Problemas de atomicidad en accesos concurrentes**

Desventajas BDs

- **Coste de creación y mantenimiento**
- **Mayor complejidad**
- **Dependencia del fabricante (licencias costosas)**
- **Frecuentes actualizaciones de versiones**