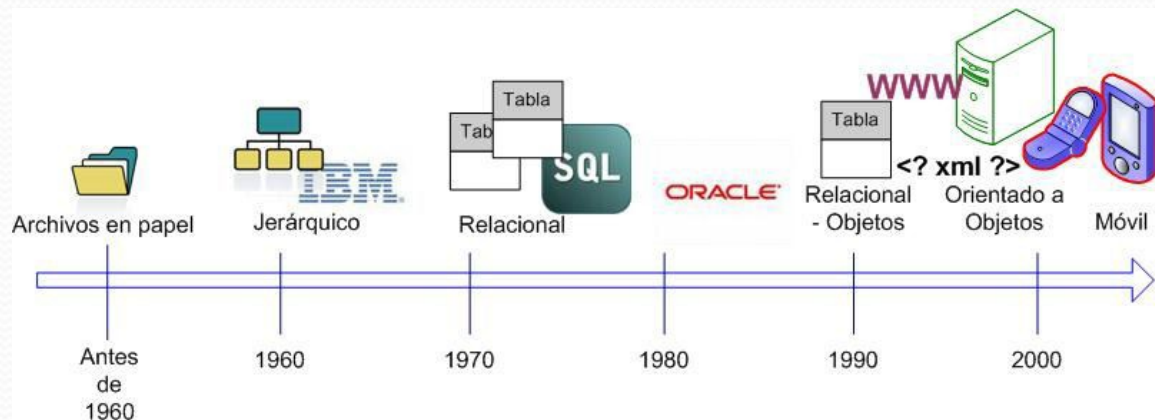


U.T. 1: Introducción al almacenamiento de información.



Contenidos

- ❑ Ficheros. Tipos y formatos.
 - ❑ Bases de datos. Conceptos, usos y tipos.
 - ❑ Sistemas Gestores de Bases de Datos.
-

Ficheros

- ❑ Son estructuras de información que crean los sistemas operativos de los ordenadores para poder almacenar datos.
- ❑ Sirven para organizar la información dentro de los dispositivos de almacenamiento.
- ❑ Se identifican por su nombre y su extensión.
- ❑ El tipo de fichero determina la forma de interpretar la información que contiene.

P.e. imagen -> fichero.bmp (binario)

Clasificación de los ficheros (clásica)

□ Según su contenido

Los datos contenidos en el fichero pueden ser tratados directamente como caracteres alfanuméricos o a estructuras más complejas (imágenes, sonido, vídeo,...)

□ Según su organización

Dicta la forma en la que se han de acceder los datos.

□ Según su utilidad

Es decir, según el uso que se va a hacer de los datos.

Clasificación de los ficheros

- ❑ Según su contenido
 - ❑ Texto
 - ❑ Binario
 - ❑ Según su organización:
 - ❑ secuencial
 - ❑ directa
 - ❑ indexada
 - ❑ Según su utilidad:
 - ❑ maestros
 - ❑ históricos
 - ❑ movimientos
-

Clasificación de los ficheros (actual)

- ☐ Según su **contenido**:

- ☐ Texto
- ☐ Binario

- ☐ Según su **tipo**:

- ☐ imágenes
 - ☐ ejecutables
 - ☐ clips de vídeo
 - ☐ código fuente
 - ☐ etc,...
-

Ficheros de texto

- ❑ Contienen información codificada según un código alfanumérico (ascii, unicode,...)
 - ❑ Suelen tener una extensión que identifica el tipo de texto que contienen:
 - ❑ Ficheros de configuración
 - ❑ Ficheros de código fuente.
 - ❑ Ficheros de páginas web
 - ❑ Formatos enriquecidos
 - ❑ etc...
-

Ficheros binarios

- ❑ Son todos los que no son ficheros de texto plano.
 - ❑ Requieren un formato para ser interpretados.
 - ❑ Tipos de formatos:
 - ❑ De imagen: .jpg, .gif, .tiff, .bmp, .vfw, .png, .pcx, ...
 - ❑ De vídeo: .mpg, .mov, .avi, .qt.
 - ❑ Comprimidos: .zip, .rar, ...
 - ❑ Ejecutables y/o compilados: .exe, .class, .com, .cgi, .jar, ...
 - ❑ Procesadores de texto: .doc, .odt,...
-

Ficheros binarios

- ❑ Ficheros que componen una bases de datos:

- ❑ Oracle: datafiles, tempfiles, logfiles, etc.

- ❑ MySQL: .frm, .myd, .myi.

- ❑ Access: .mdb.

Definición de base de datos

- ❑ **Colección** de información perteneciente a un mismo contexto, que está almacenada de forma organizada en ficheros.
 - ❑ A nivel lógico, los datos se almacenan en **tablas** formando vínculos entre sí (relaciones).
 - ❑ Las **relaciones** ayudan a mantener la información ordenada y coherente.
 - ❑ Cada tabla está dispuesta en **filas** (registros) y **columnas** (campos).
-

Conceptos de bases de datos (I)

- ❑ **Dato**: Trozo de información concreta sobre algún concepto o suceso. Pertenecen a un tipo.
 - ❑ **Tipo de Dato**: Naturaleza del campo de información.
 - ❑ **Campo**: Identificador de un conjunto de datos. También se denomina *columna*.
 - ❑ **Registro**: Recolección de datos referentes a un mismo concepto o suceso.
 - ❑ **Campo clave**: Campo especial que identifica de forma única cada registro.
 - ❑ **Tabla**: Conjunto de registros bajo un mismo nombre que representa a dicho conjunto.
-

Conceptos de bases de datos (II)

- ❑ **Consulta:** Instrucción para hacer peticiones a la bd.
 - ❑ Búsqueda simple de un registro
 - ❑ Solicitud de un conjunto de registros en base a un criterio.
- ❑ **Índice:** Almacena los campos clave de una tabla, organizándolos para hacer más fácil las búsquedas.
- ❑ **Vista:** Transformación o combinación de varias tablas para obtener una nueva tabla virtual.
- ❑ **Informe:** Listado estructurado de los campos y registros de una consulta con un formato determinado.
- ❑ **Guiones o scripts:** Conjunto de instrucciones asociadas a tareas de mantenimiento de la base de datos.
- ❑ **Procedimientos:** Tipo especial de script que está almacenado en la bd.

Estructura de una base de datos

- ❑ Una base de datos almacena los datos a través de un esquema.
 - ❑ El **esquema** es la definición de la estructura de la base de datos, conteniendo información sobre:
 - ❑ Tablas
 - ❑ registros
 - ❑ campos
 - ❑ procedimientos
 - ❑ vistas
 - ❑ índices,
 - ❑ etc...
 - ❑ Los gestores de base de datos como MySQL, Oracle y DB2 almacenan el esquema en forma de tablas.
-

Usos de las bases de datos

- ❑ **Administrativas**: Clientes, pedidos, facturas, productos, etc.
 - ❑ **Contables**: pagos, balances de pérdidas y ganancias, patrimonio, etc.
 - ❑ **Motores de búsqueda**: Bases de datos de Google o Altavista.
 - ❑ **Científicas**: Datos climáticos, medioambientales, geológicos..
 - ❑ **Configuraciones**: Registro de windows.
 - ❑ **Bibliotecas**: Tienda online amazon.
 - ❑ **Censos**: Información demográfica de pueblos, ciudades y países.
 - ❑ **Virus**
 - ❑ Otros usos: militares, videojuegos, deportes, etc.
-

Evolución de las bases de datos

- ❑ Las bases de datos han ido cambiando la forma de representar y extraer la información con el avance de la tecnología.
 - ❑ **Década 50:**
 - ❑ Cintas magnéticas
 - ❑ Lectura secuencial y ordenada de los datos.
 - ❑ La base de datos la componen ficheros independientes de movimientos y maestros.
 - ❑ **Década 60:**
 - ❑ Discos magnéticos
 - ❑ Acceso directo a los datos en los ficheros.
 - ❑ Bases de datos jerárquicas y en red con la información estructurada en listas enlazadas y árboles de información.
 - ❑ CODASYL fue el primer modelo de base de datos en red.
-

Evolución de las bases de datos

□ Década 70:

- En 1970 Codd publica la definición del modelo relacional basado en la lógica de predicados y la teoría de conjuntos.
- Nacen las bases de datos relacionales al amparo de la base teórica de Codd.
- Larry Ellison desarrolla el motor de base de datos ORACLE inspirado también en el modelo de Codd.
- Hoy en día, a pesar de las múltiples alternativas, el modelo relacional de Codd es el más utilizado en el diseño de BD.

□ Década 80:

- IBM lanza su motor de bases de datos DB2 y crea SQL (Structured Query Language).
 - SQL es un potente lenguaje de manipulación de datos para BD Relacionales.
-

Evolución de las bases de datos

□ Década 90:

- IBM crea una versión de DB2 capaz de dividir una BD en varios servidores comunicados por líneas de alta velocidad.
- Aparecen así las bases de datos paralelas.

□ Finales de 90:

- IBM y Oracle incorporan a sus bases de datos la capacidad de manipular objetos, creando las BD orientadas a objetos.
 - En lugar de tablas y relaciones, almacenan colecciones de objetos y su comportamiento (instrucciones para su procesamiento).
 - El desarrollo de internet crea nuevos requerimientos para las bases de datos que dan lugar a las bases de datos distribuidas.
 - Las BD Distribuídas multiplican el número de ordenadores que controlan la BD (nodos) intercambiándose información y actualizaciones a través de la red.
-

Evolución de las bases de datos

❑ Actualmente:

- ❑ Aparte de las clásicas bases de datos SQL (RDBMS), aparecen y van tomando fuerza nuevos tipos de bases de datos.
- ❑ Algunas de ellas pueden ser accedidas mediante SQL, pero normalmente no será así, puesto que cada una tendrá una API exclusiva.
- ❑ Cabe destacar que normalmente **no sustituyen a la base de datos clásica SQL**, sino que surgen por otra necesidad. Una necesidad de rendimiento extremo. Si se utilizan de una manera única, o se combinan con una base de datos SQL es una decisión de arquitectura del sistema.
- ❑ Su mayor ventaja es que están preparados para ser muy rápidos. Mucho.

❑ Más información en

Evolución de las bases de datos

□ Más información en

<http://campusv.uaem.mx/cicos/imagenes/memorias/7mocicos2009/Articulos/Evolucion%20de%20las%20Bases%20de%20Datos.pdf>

<http://histinf.blogs.upv.es/2011/01/04/historia-de-las-bases-de-datos/>

Sistemas Gestores de Bases de Datos

- ❑ Un SGBD es el conjunto de herramientas que facilitan la consulta, uso y actualización de una base de datos.
 - ❑ Ejemplos de SGBD son Oracle 11g, MySQL 5.7, SQL Server 13.0:
 - ❑ Incorporan un conjunto de herramientas software capaces de estructurar en múltiples discos duros los ficheros de datos, permitiendo el acceso a sus datos a partir de herramientas gráficas y potentes lenguajes de programación (PL/SQL, php, java, c++, ...)
-

Funciones de un SGBD

1. Facilitan el almacenamiento, acceso y actualización de los datos de la BD potenciando el rendimiento y ocultando la complejidad de los dispositivos físicos.
 2. Garantizan la integridad de los datos.
 3. Integran un sistema de seguridad en el acceso a los datos.
 4. Proporcionan un diccionario de metadatos que contiene el esquema de la BD.
 5. Permiten el uso de transacciones de manera segura.
 6. Ofrecen herramientas para monitorizar el uso de la base de datos y optimizar su rendimiento.
 7. Permiten la concurrencia sobre los datos de la BD.
-

Funciones de un SGBD

- 8. Proporcionan independencia entre los datos de la BD y las aplicaciones y usuarios que los utilizan, facilitando su migración a otras plataformas.
 - 9. Proporcionan conectividad con el exterior, a través de protocolos como ODBC.
 - 10. Incorporan herramientas para la restauración de la información en caso de desastre.
-

Tipos de SGBD

En base a la capacidad y potencia del propio gestor:

- ❑ **Gestores de Bases de Datos Ofimáticas:**

Manipulan bases de datos pequeñas orientadas a almacenar datos domésticos o de pequeñas empresas.

Ejemplo: Microsoft Access.

Tipos de SGBD

- ❑ **Gestores de Bases de Datos Corporativas:**

Tienen la capacidad de gestionar bases de datos enormes, de grandes empresas, con una carga de datos y transacciones que requieren un servidor de altas prestaciones.

Estos gestores manipulan grandes cantidades de datos de forma rápida y eficiente.

Ejemplo: Gestor de BD de Oracle y DB2.

Tipos de SGBD

- ❑ **Gestores de nivel intermedio:**

Son más asequibles económicamente, aunque capaces de manipular gran cantidad de datos, cumpliendo los estándares de la arquitectura ANSI SPARC.

No proporcionan su propio lenguaje SQL.

Ejemplo: Gestor MySQL

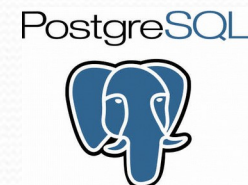
SGBD comerciales

❑ Bases de datos relacionales

- ❑ MySQL (Software Libre) - <http://www.mysql.org>
- ❑ Derby (Software Libre) - <http://db.apache.org/derby>
- ❑ H2 (Software libre) - <http://www.h2database.com/>
- ❑ HSQL (Software libre) - <http://hsqldb.org/>
- ❑ MS SQL Server (Comercial) - <http://www.microsoft.com/sql>
- ❑ PostgreSQL (Software Libre) - <http://www.postgresql.org/>
- ❑ Oracle (Comercial) - <http://www.oracle.com>

RDBMS

ORDBMS



SGBD comerciales

❑ Bases de datos NoSQL

- ❑ El término NoSQL (“no sólo SQL”) define una clase de DBMS que difieren del clásico modelo relacional:
 - ❑ No utilizan estructuras fijas como tablas para el almacenamiento de los datos
 - ❑ No usan el modelo entidad-relación
 - ❑ No suelen permitir operaciones JOIN (para evitar sobrecargas en búsquedas)
 - ❑ Arquitectura distribuida (los datos pueden estar compartidos en varias máquinas mediante mecanismos de tablas Hash distribuidas)
 - ❑ Pueden manejar gran cantidad de datos (“Big Data”): al usar una arquitectura distribuida, en muchos casos mediante tablas Hash.
-

SGBD comerciales

❑ Bases de datos NoSQL

Documento	Clave-Valor	Columna	Grafo
MongoDB	Redis	Cassandra	Neo4J
CouchDB	Membase	BigTable	FlockDB
RavenDB	Voldemort	Hbase (Hadoop)	InfiniteGraph
Terrastore	MemcacheDB	SimpleDB	InfoGrid
	Riak	Cloudera	Virtuoso

