



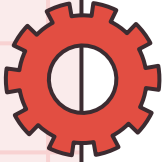
1º DAW

# U.T. 1

## Bases de Datos y Sistemas Gestores de Bases de Datos



Pablo Berciano Posada



# Tabla de contenidos

**01**

**Bases de datos**

**02**

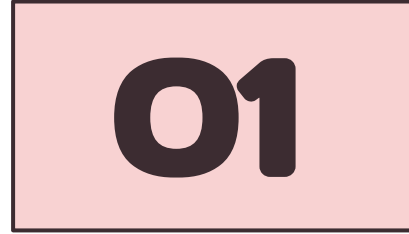
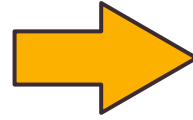
**Tipos de bases de  
datos**

**03**

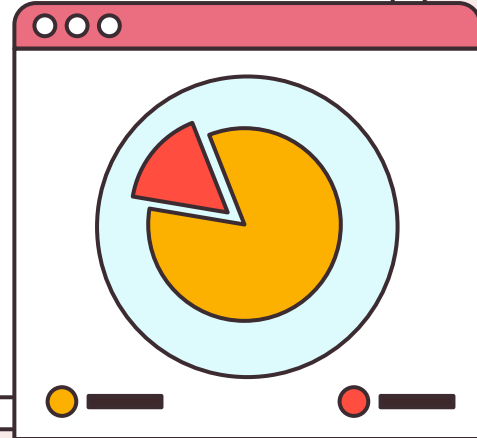
**Sistemas gestores  
de bases de datos**

**04**

**Legislación y  
actualidad**



# Bases de datos





# ¿Qué es una base de datos?



Una **base de datos** es un **conjunto de datos almacenados** entre los que existen relaciones lógicas y ha sido diseñada para satisfacer los requerimientos de información de una empresa u organización. Además, la base de datos no sólo contiene los datos de la organización, también almacena una **descripción** de dichos datos, los conocidos como **metadatos**.

Los sistemas de bases de datos separan la **definición de la estructura** de los datos de los programas de aplicación. Si se añaden nuevas estructuras de datos o se modifican las ya existentes, los programas de aplicación no se ven afectados ya que no dependen directamente de aquello que se ha modificado.

Las bases de datos son **sistemas orientados a los datos**, estos se almacenan **independientemente** de las aplicaciones.



# Características



**Independencia lógica y física de los datos:** se refiere a la capacidad de modificar una definición de un esquema en un nivel de la arquitectura sin que esta modificación afecte al nivel inmediatamente superior.

**Redundancia mínima:** se trata de usar la base de datos como repositorio común de datos para distintas aplicaciones.

**Acceso concurrente por parte de múltiples usuarios:** control de concurrencia mediante técnicas de bloqueo o cerrado de datos accedidos.

**Distribución espacial de los datos:** la independencia lógica y física facilita la posibilidad de sistemas de bases de datos distribuidas.

**Consultas complejas optimizadas:** la optimización de consultas permite la rápida ejecución de las mismas.



# Características



**Integridad de los datos:** son las medidas de seguridad que impiden que se introduzcan datos erróneos. Esto puede suceder tanto por motivos físicos como de operación .

**Seguridad de acceso y auditoría:** se refiere al derecho de acceso a los datos contenidos en la base de datos por parte de personas y organismos. El sistema de auditoría mantiene el control de acceso a la base de datos, con el objeto de saber qué o quién realizó una determinada modificación y en qué momento.

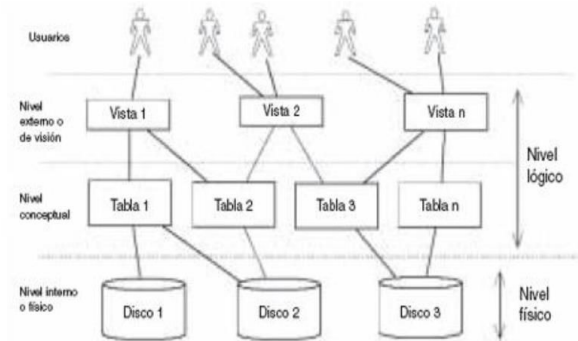
**Respaldo y recuperación:** se refiere a la capacidad de un sistema de base de datos de recuperar su estado en un momento previo a la pérdida de datos.

**Acceso a través de lenguajes de programación estándar:** posibilidad ya mencionada de acceder a los datos de una base de datos mediante lenguajes de programación ajenos al sistema de base de datos propiamente dicho.

# Niveles de abstracción

Uno de los objetivos de las bases de datos es proporcionar una **visión abstracta de los datos**, ocultando ciertos detalles del almacenamiento físico de los mismos. La **independencia de datos** es la capacidad para modificar el esquema de un nivel del sistema sin tener que modificar el esquema de nivel superior. Esto significa:

- **Independencia física de los datos:** aunque cambie el nivel físico, el nivel conceptual no se ve afectado, por ejemplo, si se cambian o añaden discos u otro hardware.
- **Independencia lógica de los datos:** aunque se modifique el esquema conceptual, los esquemas externos y los programas de aplicación no son afectados.





# Niveles de abstracción

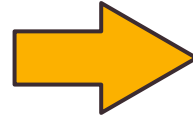


**Nivel interno (estructura física).** Nivel más bajo de abstracción de la información, y por tanto, el más cercano al almacenamiento físico. Permite describir los datos tal como están almacenados en el ordenador. Se describe por el SGBD por medio del esquema interno.

**Nivel conceptual (estructura lógica global).** Es el nivel intermedio, se define por medio de un esquema conceptual. Este nivel contempla la estructura organizacional de los datos, sin preocuparse de la organización física ni sus vías de acceso.

**Nivel externo (estructura lógica del usuario).** Es el nivel más alto de abstracción, y representa la percepción particular de un usuario de la base de datos. El SGBD es el encargado de extraer los datos requeridos por los registros lógicos externos de los registros físicos de la base de datos.





**02**

# **Tipos de bases de datos**





# Clasificación



Según su **flexibilidad de modificación**:

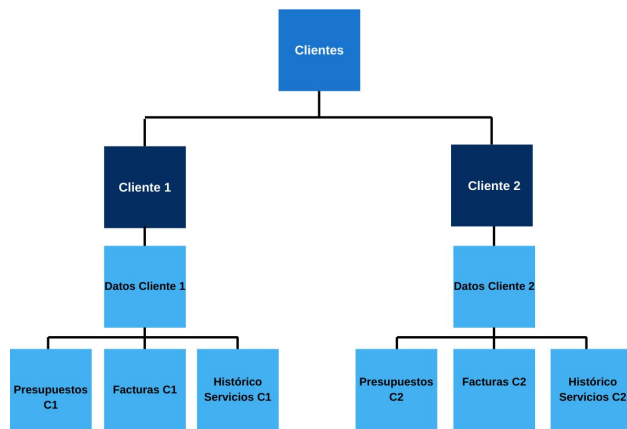
- **Bases de datos dinámicas:** Son aquellas donde los datos pueden actualizarse o incluso modificarse. La mayoría puede ser actualizada en tiempo real.
- **Bases de datos estáticas:** Son bases de datos de consulta cuyos datos no pueden modificarse.

Según su **forma de organización**:

- Bases de datos **jerárquicas**.
- Bases de datos **en red**.
- Bases de datos **relacionales**.
- Bases de datos **NoSQL**.

# Bases de datos jerárquicas

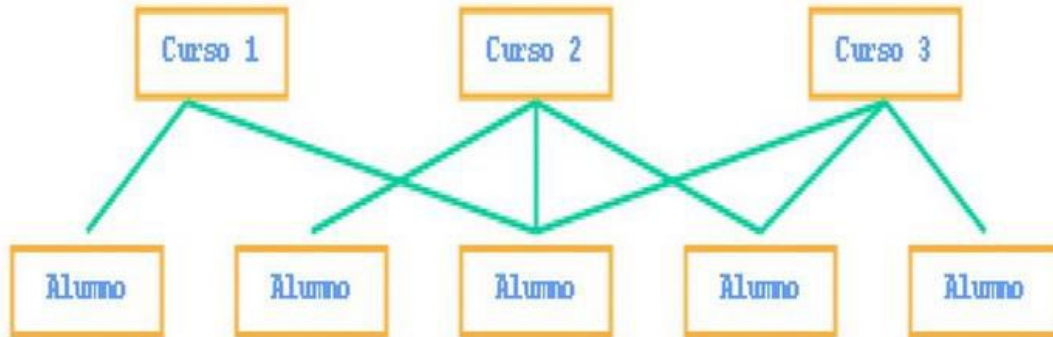
Crean **estructuras jerárquicas en forma de árbol**. Están formadas por una colección de registros con relaciones de uno a uno o de uno a muchos, pero no de muchos a muchos. Son muy **útiles con grandes volúmenes de información** y datos compartidos, permitiendo crear estructuras estables y de gran rendimiento.





# Bases de datos en red

Las **bases de datos en red** son similares a las jerárquicas pero estas sí permiten **cualquier tipo de relación**, aunque se distingue entre bases de datos en red simple, sin relaciones muchos a muchos, y bases de datos en red complejas. Cualquier sistema es representable en este tipo de base de datos.



# Bases de datos relacionales

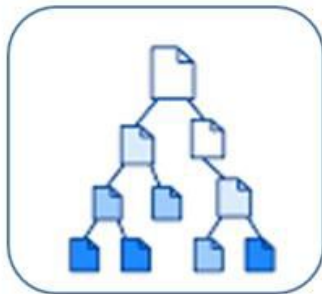
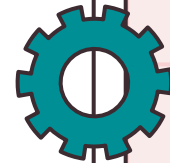
Bases de datos **formadas por tablas**. Si se imponen restricciones a las tablas, se pueden tratar como relaciones matemáticas. El **modelo relacional** es el más utilizado en el diseño y gestión de bases de datos por su facilidad de entendimiento, diseño, modificación, administración y acceso, además de contar con un lenguaje de definición y manipulación de datos potente y sencillo.



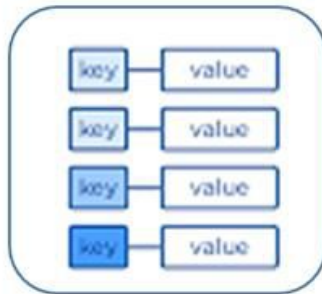


# Bases de datos NoSQL

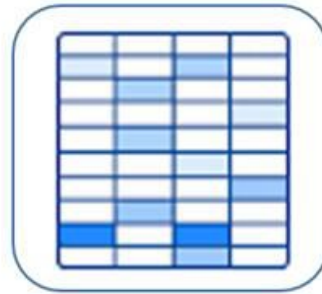
Los sistemas tradicionales son poco eficientes en **aplicaciones de uso intensivo**, con gran indexado de datos, o de streaming audiovisual. Los **sistemas NoSQL o no relacionales** permiten servir gran cantidad de carga de lecturas y escrituras.



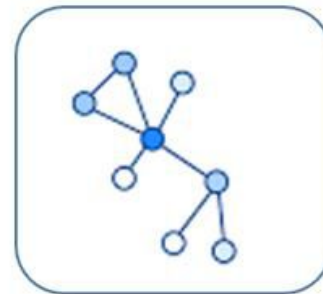
Document  
Store



Key-Value  
Store



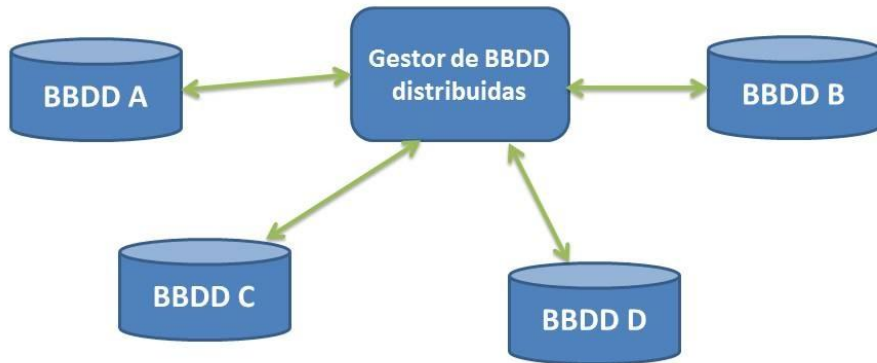
Wide-Column  
Store

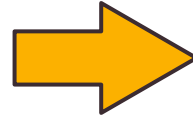


Graph  
Store

# Base de Datos distribuída

Es el caso en el que los clientes acceden a **datos situados en más de un servidor**. La localización de los datos es transparente al usuario. En esta estructura hay un **servidor de aplicaciones** que recibe las peticiones y las traduce a los distintos servidores de datos para obtener los resultados. Hay sistemas **homogéneos**, con el mismo SGBD en los servidores, y **heterogéneos**, con autonomía local de SGBD.





**03**

# **Sistemas gestores de bases de datos**







# ¿Qué es una SGBD?



El objetivo primordial de un gestor es proporcionar eficiencia y seguridad a la hora de extraer o almacenar información en las BD. Los **sistemas gestores de bases de datos (SGBD)** están diseñados para gestionar grandes bloques de información, que implica tanto la definición de estructuras para almacenamiento como de mecanismos para la gestión de la información.

El **SGBD** es una aplicación que permite a los usuarios definir, crear y mantener la BD y proporciona un acceso controlado a la misma. Un **SGBD** es el **conjunto de programas** que permiten **definir, manipular y utilizar** la información que contienen las bases de datos , realizar todas las tareas de **administración** necesarias para mantenerlas operativas, mantener su integridad, confidencialidad y seguridad. Nunca se accede a una BD directamente, sino que se hace a través del SGBD, es su interfaz entre la BD y el usuario.



# Componentes



Los **SGBD** son **paquetes de software muy complejos** que deben proporcionar una serie de **servicios** que van a permitir almacenar y explotar los datos de forma eficiente. Los **componentes principales** son los siguientes:

- **Lenguajes de los SGBD.**
- **El diccionario de datos**
- **Gestor de almacenamiento**
- **Motor de evaluación de consultas**
- **El administrador de la BD**



# Lenguajes de los SGBD



Todos los SGBD ofrecen **lenguajes e interfaces apropiadas** para cada tipo de usuario: administradores, diseñadores, programadores de aplicaciones y usuarios finales.

Este **lenguaje** es en su mayoría **SQL (Structured Query Language)**, está **estandarizado** por la **ISO**, es decir, todas las bases de datos que soportan SQL, deben tener la misma sintaxis a la hora de aplicar el lenguaje. Se divide en **cuatro sublenguajes**, que permiten al SGBD cumplir con las funcionalidades requeridas:

- Lenguaje de manipulación de datos, DML
- Lenguaje de definición de datos, DDL
- Lenguaje de control de datos, DCL
- Lenguaje de control de transacciones, TCL





# El diccionario de datos



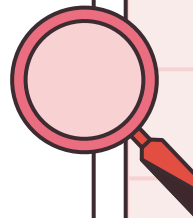
El **diccionario de datos** es el lugar donde se deposita información acerca de todos los datos que forman la BD. Es una guía en la que se describe la BD y los objetos que la forman. En una BD relacional, proporciona información acerca de:

- La **estructura lógica y física** de la BD.
- Las **definiciones de todos los objetos** de la BD: tablas, vistas, índices, disparadores, procedimientos, funciones, etc.
- El **espacio asignado y utilizado** por los objetos.
- Los **valores por defecto** de las columnas de las tablas.
- Información acerca de las **restricciones de integridad**.
- Los **privilegios y roles** otorgados a los **usuarios**.
- **Auditoría de información**, como los acceso a los objetos.



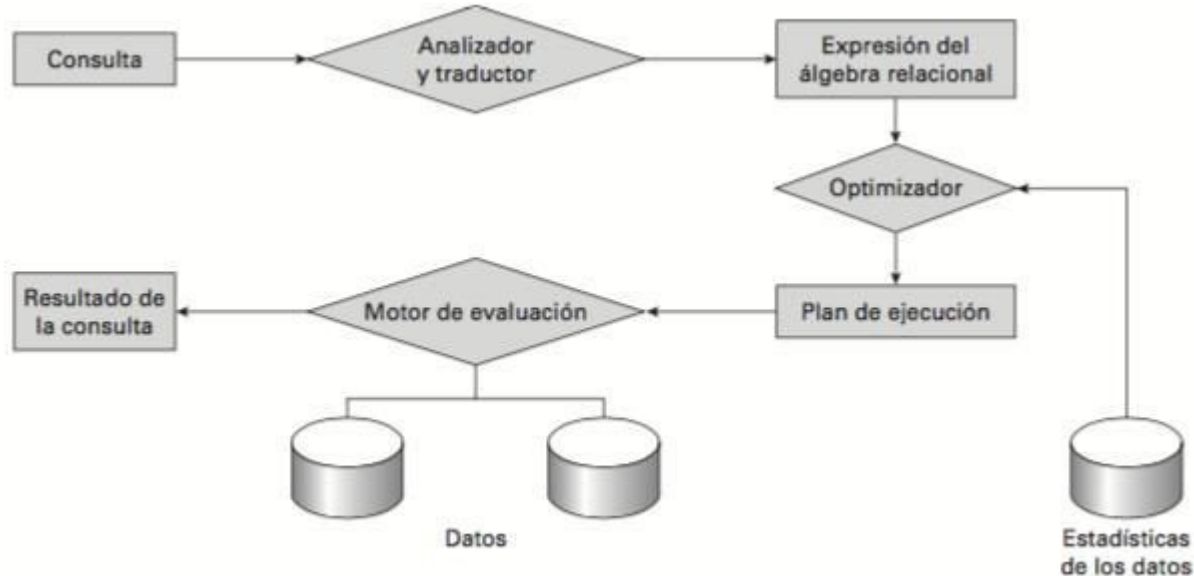
# Gestor de almacenamiento

El **gestor de almacenamiento** es el módulo que proporciona la **interfaz** entre los datos almacenados y los programas de aplicación y consultas solicitadas al sistema. Tiene varias partes:

- **Gestor de transacciones:** mantiene la consistencia ante fallos del sistema.
  - **Gestor de archivos:** gestiona la reserva de memoria a la hora de representar los datos almacenados en el disco.
  - **Gestor de memoria intermedio:** encargado de cargar en memoria los datos del disco y decidir qué datos se llevan a la caché.
- 

# Motor de evaluación de consultas

Es el encargado de ejecutar las diferentes instrucciones dadas al SGBD.





# El administrador de la BD



En los SGBD actuales existen **diferentes categorías de usuarios**. Estas categorías se caracterizan porque cada una de ellas tiene una serie de **privilegios o permisos** sobre los objetos que forman la BD. En los sistemas Oracle las categorías más importantes son:

- Los usuarios de la **categoría DBA (DataBase Administrator)**, cuya función es precisamente administrar la base y que tienen el nivel más alto de privilegios. En la administración de una BD siempre hay que procurar que haya el menor número posible de administradores.
- Los usuarios de la **categoría RESOURCE** pueden crear sus propios objetos y tienen acceso a los objetos para los que se les ha concedido permiso.
- Los usuarios de la **categoría CONNECT** solamente pueden utilizar aquellos objetos para los que se les ha concedido permiso de acceso.



# El administrador de la BD



El objetivo principal de un **DBA** es **garantizar que la BD cumple con los fines previstos** por la organización, lo que incluye una serie de **tareas básicas** como: instalar SGBD en el sistema informático, crear las BBDD que se vayan a gestionar, crear y mantener el esquema de la BD, crear y mantener las cuentas de usuario de la BD, arrancar y parar SGBD, y cargar las BBDD con las que se ha de trabajar.

Además, realiza **tareas de explotación** como: vigilar el trabajo diario colaborando en la información y resolución de dudas de los usuarios de la BD, controlar los acceso, tasas de uso y cargas en el servidor, efectuar copias de seguridad periódicas de la BD, restaurar la BD después de un incidente material, estudiar las auditorías del sistema para detectar anomalías e intentos de violación de la seguridad, y ajustar y optimizar la BD mediante el ajuste de sus parámetros.





# Opciones de funcionamiento



**SGBD de una capa:** instalados en una máquina desde la que se conectan los propios usuarios y administradores.

**SGBD de dos capas:** sigue el modelo tipo cliente/servidor. La base de datos y el sistema gestor se alojan en un servidor al cual se conectan los usuarios desde máquinas clientes. Un software de comunicaciones se encarga de permitir el acceso a través de la red. Hay dos posibilidades: servidor único o multi-servidor, aunque para el usuario no hay diferencia.

**SGBD de tres o más capas:** en este caso entre el cliente y el servidor hay al menos una capa intermedia. Esa capa se encargan de recoger las peticiones de los clientes y luego de comunicarse con el servidor para recibir la respuesta y enviarla al cliente. El caso típico es que la capa intermedia sea un servidor web.



# SGBD Comerciales



Los **SGBD relacionales de código cerrado** más conocidos son:

- **Oracle.** Propiedad de Oracle Corporation. Es el SGBD más veterano y más influyente ya que la mayoría de mejoras al SQL original se desarrollaron para este SGBD. Presume de su gran estabilidad y escalabilidad, un control avanzado de transacciones y de sus lenguajes internos de manejo, especialmente famoso es su lenguaje procedimental PL/SQL.
- **DB2.** Propiedad de IBM, es una de las bases de datos comerciales más populares. Desarrollada para Windows, UNIX y Linux. Implementa XML de manera nativa y amplias facilidades de migración de datos, así como uso de transacciones avanzadas.
- **SQL Server.** De Microsoft, dispone de una gran escalabilidad, estabilidad, uso de transacciones, entorno gráfico avanzado y de éxito entre los programadores de la plataforma .NET.

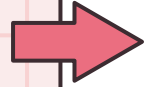


# SGBD Comerciales

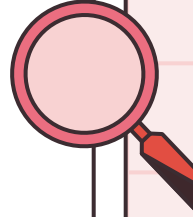


Los **SGBD relacionales de código abierto** más conocidos son:

- **MySQL.** Ha sido considerada como la principal SGBD de la comunidad de programadores de código abierto y de hecho en Internet sigue siendo la principal base de datos asociada a una aplicación web. Mantiene su licencia de tipo GPL, pero posee una segunda licencia cerrada para opciones de trabajo más avanzadas. Es muy popular por su histórica asociación con PHP, por su buena estabilidad y gran escalabilidad.
- **PostgreSQL.** está considerado como el SGBD de código abierto más potente y, sobre todo, más fidedigno con los estándares. Posee uso de transacciones avanzadas, lenguaje procedimental, gran estabilidad y escalabilidad.

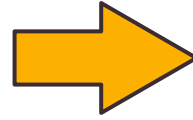


# Políticas de fragmentación de la información



Estas políticas se refieren a **cómo se dividen y almacenan los datos** en fragmentos o particiones dentro de la base de datos. Son esenciales para mejorar el **rendimiento, escalabilidad y disponibilidad** de la base de datos. Las más comunes son:

- **Fragmentación horizontal:** Se divide una tabla por registros completos entre múltiples servidores. Agiliza las consultas.
- **Fragmentación vertical:** Se divide una tabla por columnas. Minimiza la duplicación de datos.



**04**

# **Legislación y actualidad**





# Protección de datos



Con respecto al tratamiento de los datos, hay que tener especial consideración con aquellos que sean de clientes y terceros. Nuestras aplicaciones, y sus desarrolladores por extensión, son responsables de proteger los datos recogidos.

Existen varias **leyes** que aplican a los datos que podemos almacenar en una base de datos:

- **RGPD:** en el marco europeo, se aplicó para toda la UE desde mayo del 2018, establece derechos de los usuarios, responsabilidades de las organizaciones que almacenan los datos, y posibles sanciones por incumplimiento, entre otras cosas.
- **LOPD:** es la ley que especifica para España la ley europea RGPD.



# Big Data

Cuando trabajamos con un **volumen de datos extremadamente grandes y complejos** que superan la capacidad de procesamiento tradicional, estamos hablando de **Big Data**. Estos datos pueden provenir de redes sociales, sensores, u otras fuentes. En ocasiones, el Big Data implica análisis en tiempo real o al menos en ventanas de tiempo variables. El objetivo del Big Data es **extraer información valiosa de los datos masivos y brutos**.



# Business Intelligence

El **Business Intelligence** se refiere a la recopilación, análisis y presentación de datos para respaldar la toma de decisiones empresariales.

En muchos casos implica el uso de **herramientas de visualización de datos** agregados, con mapas, tablas, gráficos, etc. que facilitan entender las estadísticas extraídas.

