

# Resumen-esquema sobre las bases de datos.

## Actividad 4.1

---



### Actividad:

Realizar un resumen-esquema en el que se recojan los siguientes puntos:

1. Definición de Base de Datos.
2. Modelos de Bases de Datos.
3. Bases de datos más utilizadas. Incluir el ranking de 2022 y compararlo con el de años anteriores.
4. Qué es un Sistema Gestor de Bases de Datos.
5. Funciones del Sistema Gestor de Bases de Datos.
6. Dar ejemplo de Sistemas Gestores de Bases de Datos

---

## 1. Definición de bases de datos

Una base de datos es una recopilación organizada de **información** o **datos integrados** y estructurados, que tienen una **redundancia controlada** y una estructura clara que refleja las interrelaciones y restricciones semánticas existentes en el mundo real.

Los datos que componen la base deben mantenerse **independientes** de los usuarios y aplicaciones que los han **compartido** y mantener su descripción y definición.

El conjunto de datos debe tener procesos de actualización y recuperación capaces de conservar su **integridad, seguridad y confidencialidad**.

En este sentido, una biblioteca pública puede considerarse una base de datos compuesta en su mayoría por documentos y textos impresos en papel e indexados para su consulta, sin embargo, hoy en día la mayoría de las bases de datos están en formato digital, donde hay programas, denominados sistemas gestores de bases de datos, de los cuales hablaremos más tarde, que permiten almacenar y acceder a los datos de forma rápida y estructurada.

### Ventajas:

Las bases de datos tienen varias ventajas, estas se pueden dividir y definir como:

- **Control sobre la redundancia de datos**, ya que en los sistemas de bases de datos todos los ficheros están integrados y no se almacenan copias.
- **Consistencia de los datos**, puesto que al eliminar las redundancias reducimos en gran medida el riesgo de inconsistencias.
- **Compartición de datos**, puesto que la base de datos está disponible para todos los usuarios autorizados y las nuevas aplicaciones que se creen pueden reutilizar los datos anteriores.
- **Se simplifica el esfuerzo de programación** y mantenimiento de los programas, ya que gracias a la integración es más fácil respetar los estándares necesarios establecidos a nivel empresarial y nacional.

## 2. Modelos de bases de datos

Los principales modelos de bases de datos se dividen en los siguientes:

### ◇ Jerárquico:

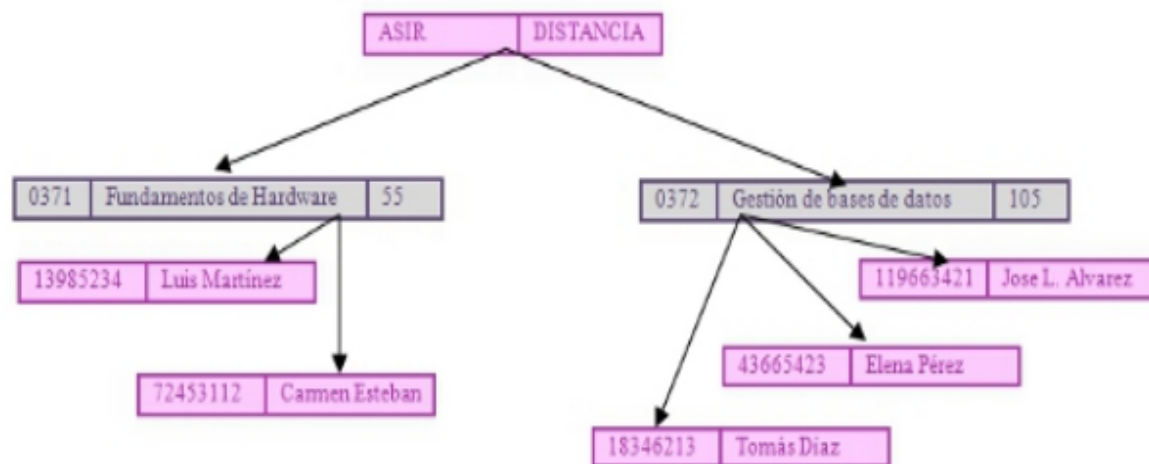
- ▼ Este modelo, también llamado *modelo en árbol*, utiliza su representación para hacer una estructura de tipo árbol invertido. Es decir, una base de datos jerárquica es un conjunto de registros lógicamente organizados en dicha estructura.
- ▼ El nivel superior se percibe como “*el padre*” de todos los registros situados por debajo, de tal forma que cada padre puede tener muchos hijos, pero un hijo nunca tendrá más de un padre.

MODULOS

CODIGO	NOMBRE	HORAS
--------	--------	-------

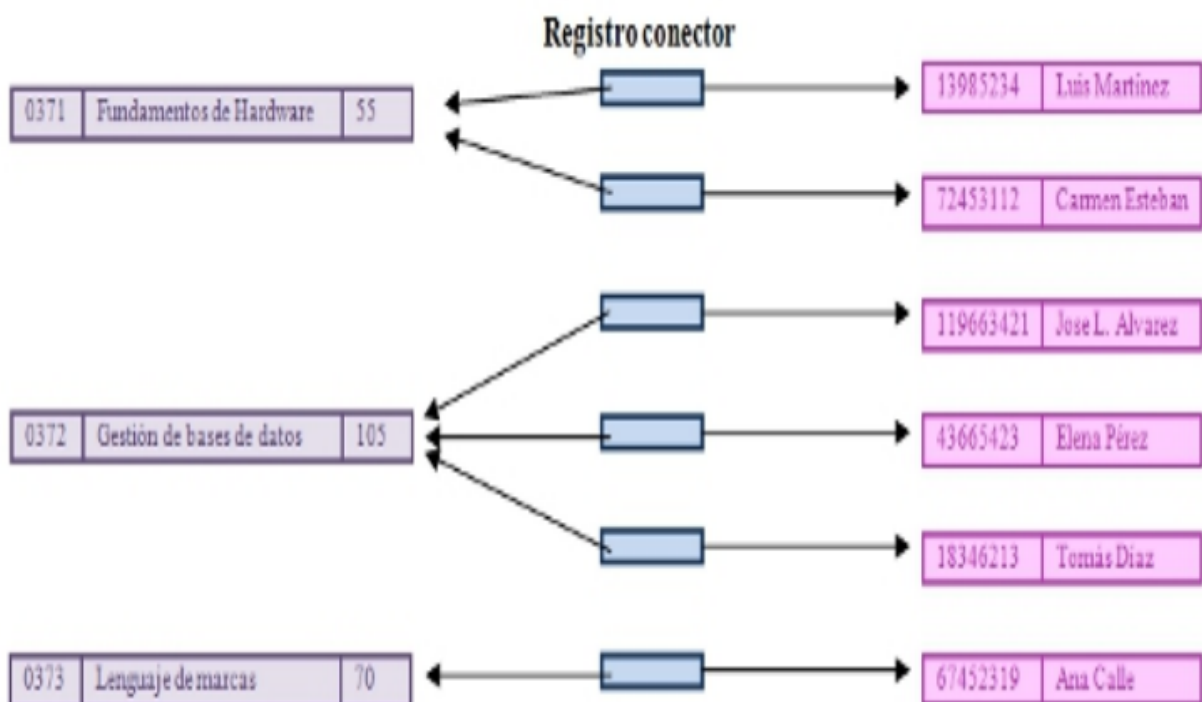
PROFESORES

NRP	NOMBRE
-----	--------



◇ **En red:**

- ▼ El modelo en red es propio de los sistemas comerciales de los años 70, aunque hoy en día siguen estando vigentes si no se piden demasiadas modificaciones.
- ▼ Es más complejo y eficiente que el modelo anterior, imponiendo así un estándar de bases de datos que ayudara a los diseñadores y programadores.

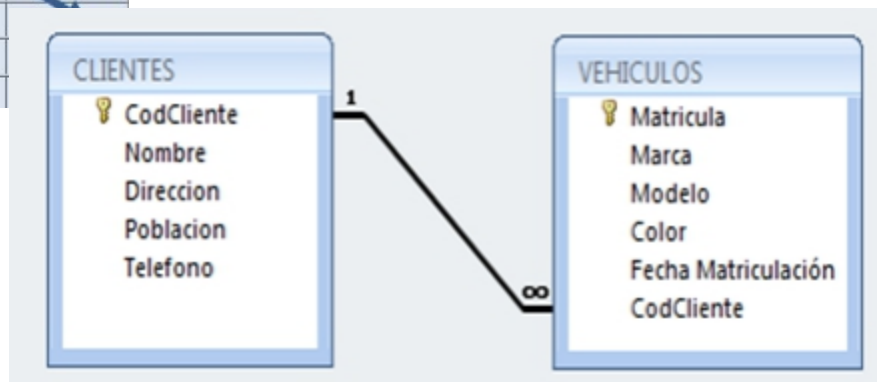


### ◇ Relacional:

- ▼ El modelo relacional fue desarrollado en los años 70 por IBM, aunque en la época los ordenadores no eran capaces de ejecutarlo. Hoy en día, sin embargo, es el modelo más utilizado para modelar y administrar datos dinámicamente.
- ▼ Su principal ventaja es que permite al usuario y diseñador operar usando un entorno gráfico que se percibe como un conjunto de tablas dejando los detalles físicos y complejos al sistema.
- ▼ Sus principales características son:
  - ◆ Representación de datos y relaciones mediante una colección de tablas.
  - ◆ Proporciona una redundancia y una inconsistencia mínima
  - ◆ Independencia de los datos de las aplicaciones y los datos del dispositivo de almacenamiento.

CLIENTES				
CodCliente	Nombre	Dirección	Población	Teléfono
1	Francisco Álvarez	C/ La Mata 9.	Alcántara	925767788
2	Raquel Marcos	C/ La Amapola 7.	Toledo	925998811
3	Carlos Revuelta	C/ La Arboleda 12.	Madrid	919090771
4	José María Sabello	C/ La plaza 72	Madrid	914409071
5	Jorge Peña	C/ Fresno, 23	Guadalajara	949788896
6	Dolores Manzano	C/ Autonomía, 8	Madrid	916767560

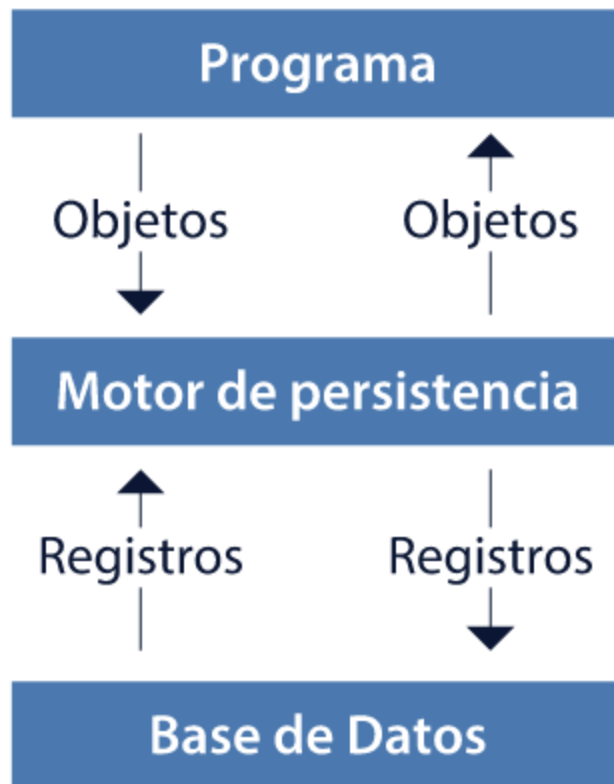
VEHICULOS					
Matricula	Marca	Modelo	Color	Fecha Matriculación	CodCliente
4534 FNG	Ford	Focus	Negro	14/04/2007	2
1203 CLL	Cotroën	C4	Magenta	23/08/2005	5
3367 GHB	Suzuki	Vitara	Cobalto	16/05/2009	
1004 JLG	Kia	Rio	Rojo	02/07/2009	
6709 BFG	Peugeot	206	Gris plata	12/10/2006	



---

◇ **Orientadas a objetos:**

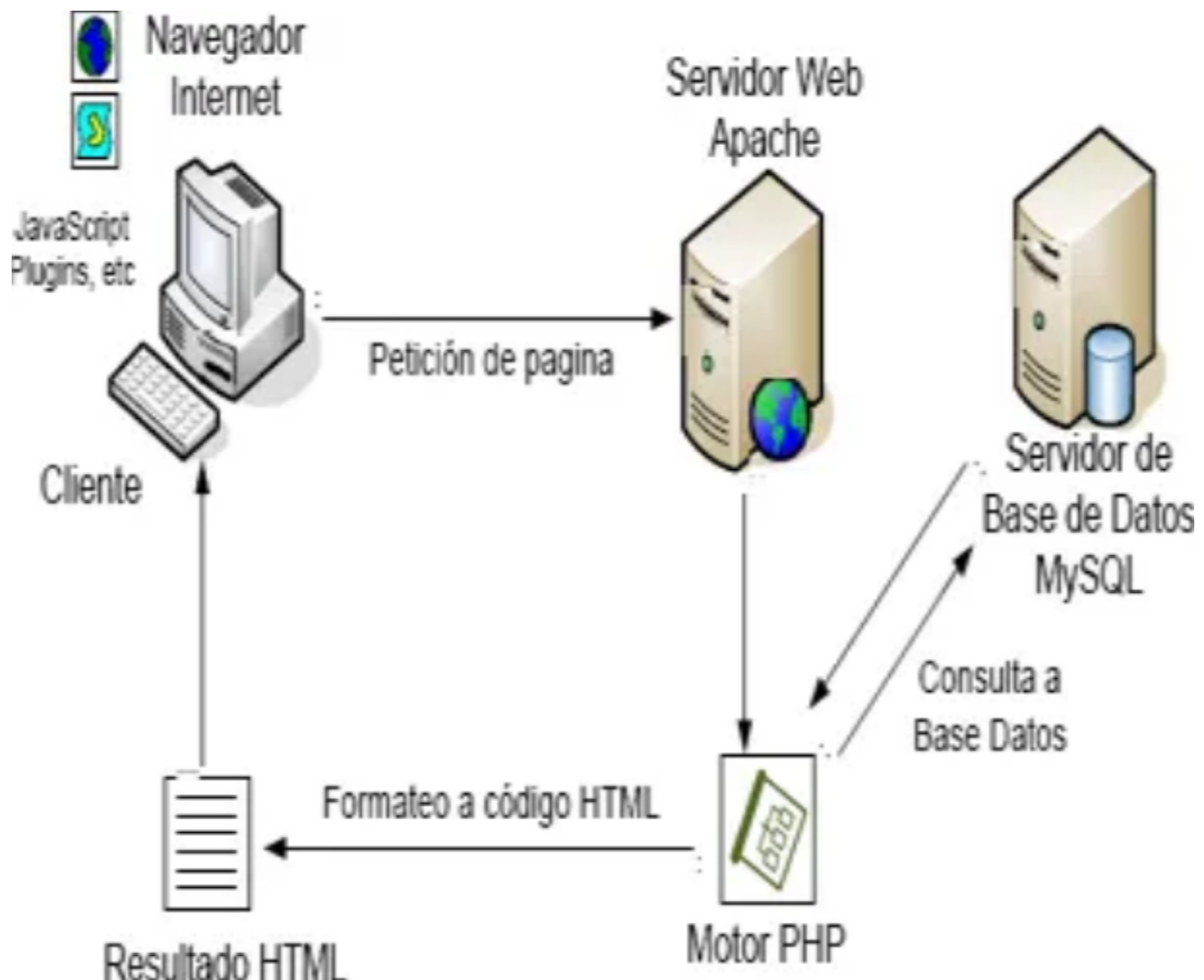
- ▼ Los modelos orientados a objetos surgieron recientemente (años 90) para intentar representar con más fidelidad los problemas del mundo real.
- ▼ Se denominan así ya que su estructura básica es un objeto que recoge tanto datos como relaciones, es decir, supone una forma diferente de definir y utilizar las entidades.
- ▼ Un objeto se describe como un conjunto orientado de hechos que incluye información sobre la relación de dentro del objeto y con otros objetos, junto a todas las operaciones que puedan ser realizadas en él.



◇ **Orientadas a internet:**

▼ Son bases de datos que se comunican fácilmente con Internet. Sus principales características son:

- ♦ Acceso flexible a Internet.
- ♦ Fácil conexión a distintas estructuras de datos.
- ♦ Diseño conceptual sencillo.
- ♦ Dispone de herramientas de diseño, consulta, desarrollo de aplicaciones y una interfaz gráfica que facilite el trabajo.



---

### 3. Bases de datos más utilizadas en 2022 y comparación con años anteriores.

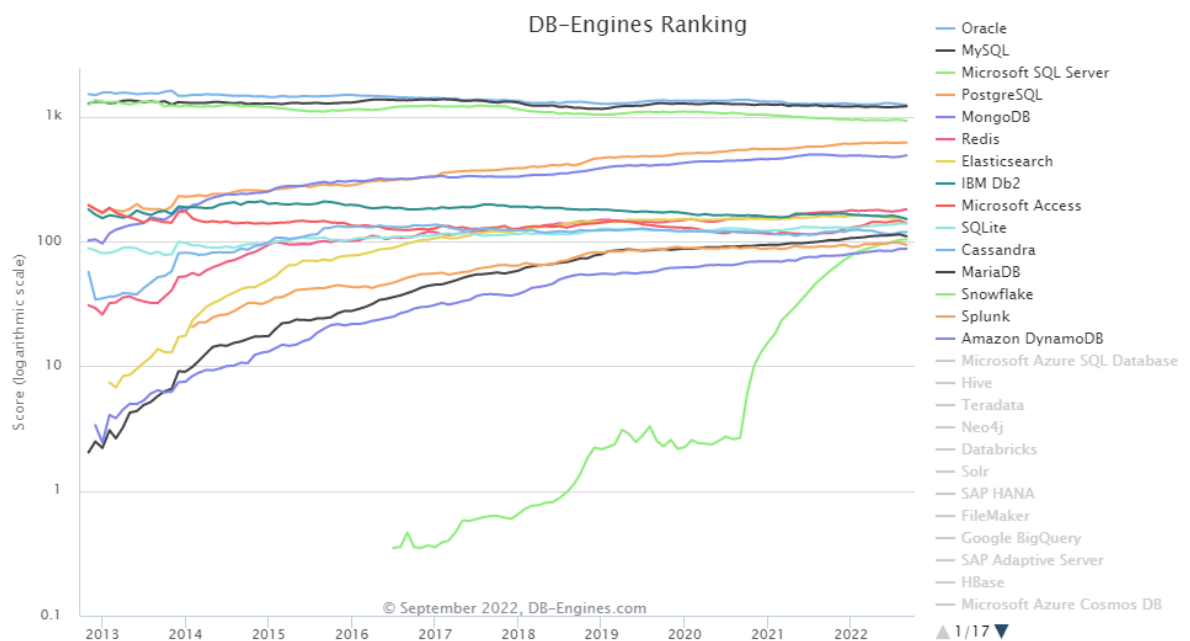
Las bases de datos más utilizadas según [DB-Engine ranking](#), son las siguientes:

Posición	Base de datos	Rank
1	Oracle	1268.84
2	MySQL	1154.27
3	Microsoft SQL Server	1040.26
4	PostgreSQL	466.11
5	MongoDB	387.18
6	IBM Db2	179.85
7	Redis	149.01
8	Elasticsearch	143.44
9	Microsoft Access	141.62
10	SQLite	126.8
11	Cassandra	122.98
12	Splunk	81.43
13	MariaDB	78.82
14	Teradata	76.19
15	Hive	69.91
16	Solr	61.48
17	HBase	60.39
18	FileMaker	57.15
19	SAP HANA	56.64
20	Amazon DynamoDB	55.09
21	SAP Adaptive Server	55.04
22	Neo4j	46.8
23	Couchbase	34.59
24	Memcached	29.54
25	Microsoft Azure SQL Database	27.2



El ranking está calculado en base a los siguientes factores:

- ▼ Número de búsquedas en Google, Bing y otros buscadores.
- ▼ Frecuencia de búsquedas en Google Trends.
- ▼ Frecuencia de discusiones en foros especializados en IT.
- ▼ Número de ofertas de trabajo y perfiles de trabajadores.
- ▼ Menciones en twitter y otras redes sociales.



Aquí podemos comprar la popularidad (basada en el mismo ranking) de los años anteriores. Las mas populares como Oracle o MySQL continúan siendo bastante estables, mientras que se puede ver a otras nuevas como Snowflake subiendo muchísimo en popularidad.

---

#### 4. ¿Qué es un sistema gestor de bases de datos?

Un sistema gestor de base de datos es un paquete de software complejo que debe proporcionar unos servicios que permitan almacenar y explotar los datos de una forma eficiente. La mayoría incorporan lenguajes de cuarta generación, denominados *herramientas de desarrollo*.

Estos sistemas tienen varios lenguajes que se dividen en:

- **LDD o DDL** (*Lenguaje de definición de datos*)
  - El lenguaje de definición de datos se utiliza para definir el esquema conceptual e interno de la base de datos, es decir, los objetos, estructuras de almacenamiento y las vistas de los usuarios, y es empleado por los diseñadores y administradores de esta.
- **LMD o DML** (*Lenguaje de manipulación de datos*)
  - El lenguaje de manipulación de datos se usa para consultar y actualizar los datos de la base de datos, y es empleado por los usuarios que necesitan consultar, insertar y/o modificar o borrar los datos. Suele estar en lenguaje de alto nivel.
- **LCD o DCL** (*Lenguaje de control de datos*)
  - El lenguaje de control de datos se usa para controlar los temas referentes a seguridad y administración de las diferentes bases de datos.

Los SGBD utilizan también una serie de mecanismos para garantizar un acceso correcto, seguro y eficiente de los datos. Esto se hace a través de un componente software que se encarga de diferentes funciones. Estas funciones se dividen en:

- Garantizar el acceso a los datos solo por el personal autorizado
- Disposición de herramientas para planificar y realizar copias de seguridad
- Realización de los procedimientos necesarios para recuperar los datos tras un fallo
- Mecanismos para implantar restricciones de integridad
- Control del acceso concurrente de varios usuarios sin que se pierda la consistencia

---

Los usuarios que tendrán los SGBD se componen por cuatro grupos principales:

- **Programadores:**
  - Los programadores son los responsables de la creación de las aplicaciones que ataquen los datos.
- **Usuarios expertos:**
  - Emplean las diferentes utilidades de la base de datos y el DML para acceder y realizar sus propios procesos sobre los objetos.
- **Usuarios ocasionales:**
  - Utilizan programas de aplicación para acceder a la base de datos, aunque solo pueden usar los objetos de los cuales tienen acceso.
- **Diseñadores-administradores:**
  - Los diseñadores planifican y desarrollan las bases de datos en sí, definen su esquema lógico y físico, priorizando optimizar el almacenamiento y generando la documentación para los programadores. Una vez creadas, ejercen de administradores

---

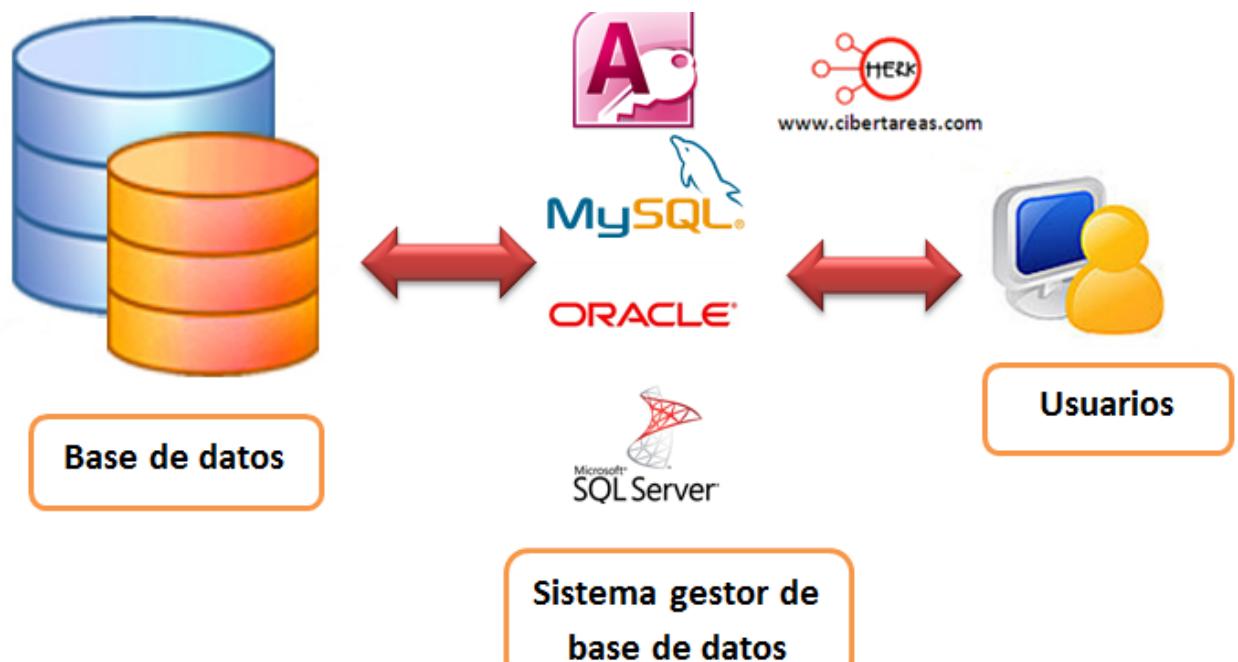
## 5. Funciones de un sistema gestor.

Un sistema gestor realiza funciones que tienen que garantizar la integridad y consistencia de los datos.

Sus principales funciones son:

- **Administración del diccionario de datos:**
  - Es donde se guardan las definiciones de los datos y sus relaciones.
- **Administración del almacenamiento de datos:**
  - Esta función permite la creación de las estructuras necesarias para el almacenamiento de los datos.
- **Transformación y presentación de datos:**
  - Transformación de los datos para introducirlos en las estructuras necesarias para guardarlos.
- **Administración de la seguridad:**
  - Determina a qué datos pueden tener acceso los usuarios y que operaciones pueden realizar.
- **Control de acceso de usuarios múltiples:**
  - Crea algoritmos complejos para no comprometer la seguridad cuando acceden múltiples usuarios.
- **Administración de tareas de respaldo y recuperación:**
  - Proporciona utilidades que permiten realizar ciertos procedimientos en caso de fallos en la red o en el suministro eléctrico.

- **Administración de la integridad de los datos:**
  - Controla que se cumplan las reglas de integridad y de las relaciones.
- **Lenguajes de acceso a la base de datos e interfaces de programación de aplicaciones:**
  - Permite acceder a los datos utilizando un lenguaje de consulta.
- **Interfaces de comunicación de bases de datos:**
  - Permiten que la base de datos acepte solicitudes de usuarios conectados a través de internet.



---

## 6. Ejemplos de sistemas gestores.

- **Oracle:** es un sistema de gestión de base de datos objeto-relacional. Es uno de los sistemas de bases más completos, y destaca su estabilidad, escalabilidad y soporte multiplataforma.
- **MySQL:** con más de seis millones de instalaciones, es un sistema de gestión de bases de datos relacional, multihilo y multiusuario, se ofrece bajo GNU GPL para cualquier uso compatible con esta licencia, aunque también ofrece productos privados.
- **SQL Server:** es producido por Microsoft basado en el modelo relacional. Usa T-SQL y ANSI SQL como lenguajes para las consultas, una de sus características principales es que incluye un entorno gráfico de administración, también permite trabajar en modo cliente-servidor, soporta procedimientos almacenados y transacciones.
- **PostgreSQL:** es un sistema de gestión de bases de datos objeto-relacional. A diferencia de los otros, su código fuente está disponible libremente. *PostgreSQL* usa un modelo cliente/servidor y usa un multiproceso en vez de multihilo (*como usan otros*) para garantizar la estabilidad del sistema.