


COLEXIO VIVAS S.L.	RAMA:	Informática	CICLO:	Desenvolvimento de Aplicacions Multiplataforma		
	MÓDULO:	Bases de datos				CURSO: 1º
	PROTOCOLO:	Apuntes clases	EVAL:	1	FECHA:	2023/2024
	UNIDAD COMPETENCIA:					

Tema 1: Introducción a las Bases de Datos

Índice

1. Introducción.....	1
1.1. Sistemas de archivos	1
1.2. Bases de datos	2
2. Sistema Gestor de Base de Datos (SGBD)	3
3. Modelo de datos.....	4
3.1. Niveles de abstracción en una base de datos	4
3.2. Clasificación de los modelos de datos.....	5
3.3. Instancias y esquemas	6
3.4. Lenguajes de datos.....	6
3.4.1. Lenguaje de definición de datos	6
3.4.2. Lenguaje de manipulación de datos.....	6
3.4.3. Lenguaje de control de datos.....	6
4. Bibliografía	7

	RAMA:	Informática	CICLO:	Desenvolvemento de Aplicacions Multiplataforma			
	MÓDULO:	Bases de datos				CURSO:	1º
	PROTOCOLO:	Apuntes clases		EVAL:	1	FECHA:	2023/2024
	UNIDAD		COMPETENCIA:				

1. Introducción

En una organización uno de los elementos de más importancia es la información que posee junto con los tratamientos que realice con ella usando, para ello, usará un sistema de información¹ (SI).

Un sistema de información es un conjunto de elementos relacionados que permite obtener, procesar, administrar y almacenar información con el fin de cubrir las necesidades u objetivos de una organización.

Un sistema de información posee, entre otros, los siguientes elementos que interactúan entre sí.

1. Datos: representan la información que maneja el sistema de información.
2. Software: es la parte lógica del sistema. Está compuesto por todos los programas que utiliza el sistema de información y que pueden incluir entre otros: el sistema gestor de bases de datos, el sistema operativo, ...
3. Hardware: son todos los elementos físicos que utiliza el sistema de información para procesar y almacenar los datos. Se compone de ordenadores, sistemas de almacenamiento, ...
4. Personas: Son los usuarios que hace uso del sistema de información. Este uso puede ser a distintos niveles: nivel desarrollador, nivel administrador, nivel usuario, ...

1.1. Sistemas de archivos


En los primeros sistemas de información cada aplicación utilizada en una organización solía usar sus propios datos que se almacenaban en ficheros con un formato propio que era dependiente de dicha aplicación. Por ejemplo archivos de texto, de gráficos, hojas de cálculo, ...

Estos sistemas se conocen como sistemas orientados al proceso porque la parte relevante es el proceso de tratamiento que reciben los datos en vez de los propios datos.

Este sistema presenta los siguientes problemas:

- Redundancia e inconsistencia de los datos: la redundancia se produce cuando un dato se encuentra almacenado en lugares diferentes mientras que la inconsistencia se produce si estos datos presentan además valores diferentes. En estos sistemas la probabilidad de que se produzcan inconsistencias es alto. Imaginemos dos aplicaciones diferentes que utilizan sus propios datos para una misma persona y en una de ella modificamos un número de teléfono pero nos olvidamos de modificarlo en la otra ... al pasar el tiempo ¿Cómo sabemos que teléfono es el correcto?.
- Espacio de almacenamiento elevado: la redundancia de los datos, visto en el punto anterior, provoca un aumento del espacio de almacenamiento usado debido a que los mismos datos se encuentran almacenados en más de un lugar al mismo tiempo.

¹ https://es.wikipedia.org/wiki/Sistema_de_información

	RAMA:	Informática	CICLO:	Desenvolvemento de Aplicacions Multiplataforma				
	MÓDULO:	Bases de datos				CURSO:	1º	
	PROTOCOLO:	Apuntes clases		EVAL:	1	FECHA:	2023/2024	
	UNIDAD		COMPETENCIA:					

- Dependencia de la estructura del fichero: la forma de que un dato se almacena en un fichero dependerá del tipo fichero que queramos usar. Esto puede presentar dos problemas:
 - Dificultad para el acceso a los datos: se necesita conocer la estructura de todos los tipos de ficheros que se usen (texto, gráficos, ...) creando el software necesario para acceder a ellos.
 - Dificultad para modificar la estructura de un fichero puesto que también se tendrían que actualizar todos los programas que acceden a esos tipos ficheros para que puedan reconocer los cambios producidos.
- Mantenimiento de la integridad: si un dato posee restricciones de integridad estas han de ser creadas para todos los programas que usen ese dato, lo que consume tiempo y recursos. Por ejemplo para el número de un mes se debe garantizar que sea un número entero entre 1 y 12.
- Acceso concurrente: se tiene que garantizar, en todas las aplicaciones, que no se produzcan problemas de consistencia en los datos ante accesos simultáneos, por distintos usuarios, a los mismos.


1.2. Bases de datos

Las bases de datos surgen como un intento de solucionar los problemas que presentan los sistemas de archivos clásicos (sistemas orientados al proceso). En ellas los datos se encuentran organizados y relacionados en un depósito común centralizado de el cual las aplicaciones obtienen los datos.

Se conocen como sistemas orientados a los datos ya que en este sistema la importancia la tienen los datos en sí mismos y no el tratamiento que se realiza sobre ellos.

Las ventajas de los sistemas orientados a los datos son:

- Los datos están interrelacionados y estructurados siendo esta estructura única y centralizada.
- La independencia de los datos y sus tratamientos.
- Se evitan las redundancias, no controladas, e inconsistencias de los datos.
- Al no existir, de forma general, redundancia el espacio de almacenamiento es menor.
- Permitir el acceso de varios usuarios o aplicaciones de forma simultánea ofreciendo distintas vistas de los datos.
- Garantizar la disponibilidad y la integridad de los datos.
- Permite establecer mecanismos de seguridad que, mediante un usuario y contraseña y unos privilegios, permitan determinar las acciones que un usuario puede realizar en la base de datos.

	RAMA:	Informática	CICLO:	Desenvolvimento de Aplicacions Multiplataforma		
	MÓDULO:	Bases de datos				CURSO: 1º
	PROTOCOLO:	Apuntes clases	EVAL:	1	FECHA:	2023/2024
	UNIDAD COMPETENCIA:					

Siendo los posibles inconvenientes:

- Las ventajas anteriores provocan que la complejidad que las bases de datos sean mayor por lo que se necesitara de personal cualificado.
- Así mismo se requerirá de recursos hardware más potentes.
- Al estar todos los datos en un único depósito central puede producirse problemas de disponibilidad ante fallos de la base de datos (la aplicaciones que hacen uso de la base de datos no pueden acceder a los datos).

Los puntos anteriores provocan que el coste tanto de implantación como de mantenimiento de una base de datos suela ser mayor.

2. Sistema Gestor de Base de Datos (SGBD)

El sistema gestor de bases de datos es un conjunto coordinado de programas y de lenguajes que permite que los usuarios puedan trabajar con la base de datos.

Permite realizar las siguientes tareas:

- Definición: definir la estructura de la base de datos junto con las relaciones que existen entre los datos usando para ello un Lenguaje de Definición de Datos o DDL.
- Manipulación: realizar operaciones de consulta, inserción, actualización y eliminación de datos en la base de datos mediante un Lenguaje de Manipulación de Datos o DML.
- Control: ofrece mecanismos para el mantenimiento de la base de datos: gestión de usuarios y permisos, copias de seguridad, control de concurrencia, diccionario de datos, ... usando para ello el Lenguaje de Control de Datos o DCL.

Podemos clasificar los SGBD más conocidos dentro de las tres categorías siguientes:

Libres	No libres y gratuitos	No libres
➤ MariaDB	➤ Microsoft SQL Server Compact	➤ Oracle
➤ MySQL	➤ Oracle Express Edition	➤ IBM DB2
➤ SQLite		➤ Microsoft SQL Server
➤ PostgreSQL		➤ Microsoft Access
➤ MongoDB		➤ FileMaker
➤ Cassandra		➤ MySQL: Licencia Dual.

COLEXIO VIVAS S.L.	RAMA:	Informática	CICLO:	Desenvolvimento de Aplicacions Multiplataforma		
	MÓDULO:	Bases de datos				CURSO: 1º
	PROTOCOLO:	Apuntes clases	EVAL:	1	FECHA:	2023/2024
	UNIDAD COMPETENCIA:					

En los siguientes enlaces se muestra diversos rankings con los SGBD más usados:

- Ranking en bd-engines.com:
<https://db-engines.com/en/ranking>
- Rankings en stackoverflow.com generados a partir de encuestas a sus usuarios en distintos años:
 - <https://insights.stackoverflow.com/survey/2018/#technology--databases>
 - <https://insights.stackoverflow.com/survey/2019/#technology--databases>
 - <https://insights.stackoverflow.com/survey/2020/#technology-databases>
 - <https://insights.stackoverflow.com/survey/2021/#most-popular-technologies-database>
 - <https://survey.stackoverflow.co/2022/#section-most-popular-technologies-databases>
 - <https://survey.stackoverflow.co/2023/#most-popular-technologies-database>

3. Modelo de datos

El modelo de datos es un conjunto de herramientas conceptuales que permite describir los diferentes elementos de un problema que se quiere modelar: datos, sus relaciones y las restricciones de consistencia que deben cumplir.

3.1. Niveles de abstracción en una base de datos²

En el año 1975 la organización ANSI-SPARC definió una arquitectura en tres niveles de abstracción para diseño de base de datos. El objetivo de esta arquitectura es conseguir que estos tres niveles sean independientes entre sí permitiendo la modificación del esquema de un nivel sin que los demás niveles se vean afectados. Se busca ocultar la complejidad de la base de datos a los usuarios.

Los tres niveles son:

- Nivel externo: representa como los usuarios perciben la base de datos. Cada usuario tendrá una visión propia, dependiendo de los privilegios que posea, existiendo tantas vistas como usuarios. Es la más cercana al usuario.
- Nivel conceptual: describe de forma única la estructura de la base de datos: datos y la relaciones entre ellos. Es independiente de los usuarios.
- Nivel interno o físico: especifica, en detalle, como se almacenan los datos de la base de datos en el disco usando para ello el sistema operativo. El administrador es el único usuario que tiene acceso a este nivel. Para el resto de usuarios el cómo se almacenan los datos en disco no es relevante ya que solo les debe interesar que datos están almacenados y como se relacionan entre sí. Este es el único nivel que existe realmente.



² https://es.wikipedia.org/wiki/Arquitectura_ANSI-SPARC

<div>COLEXIO</div> <div>VIVAS S.L.</div>	RAMA:	Informática	CICLO:	Desenvolvemento de Aplicacions Multiplataforma				
	MÓDULO:	Bases de datos				CURSO:	1º	
	PROTOCOLO:	Apuntes clases		EVAL:	1	FECHA:	2023/2024	
	UNIDAD		COMPETENCIA:					

Estos tres niveles buscan crear las siguientes independencias:

- Independencia Física: es la capacidad de poder modificar el esquema interno sin tener que modificar el esquema externo. Es decir cambios a nivel interno de cómo se almacena los datos en disco sin tener que modificar los programas que acceden a esos datos.
- Independencia Lógica: es la capacidad de modificar el esquema conceptual sin tener que modificar el esquema externo, es decir, se modifica la estructura de la base de y/o las relaciones que tienen entre ellos, sin modificar los programas que acceden a estos datos. La independencia lógica es más difícil de conseguir que la física.


Por lo que se puede definir el modelo de datos como un conjunto de conceptos, reglas y convenciones que permiten describir, en distintos niveles de abstracción, la estructura (el esquema) de la BD³.

3.2. Clasificación de los modelos de datos

Los modelos de datos pueden clasificarse en:

Físicos	Describe los datos a un nivel muy bajo de detalle. No es muy usado.		
Lógicos	Basados en objetos. Niveles conceptual y externo.	<u>Modelo Entidad-Relación</u>	Modelo que permite representar el mundo real mediante entidades y sus relaciones. <u>Tema 2</u> .
	Representa como captamos el mundo real.	Modelo orientado a objetos	La información se representa mediante objetos tal y como se realiza en la programación orientada a objetos.
	Basados en registros. Niveles conceptual y físico. Especifica la estructura conceptual de la BD en registros en los cuales se almacena los datos.	Modelo jerárquico	Las relaciones entre registros forman una estructura de árbol invertido. Cada registro solo puede estar relacionado con un registro padre. En desuso.
		Modelo en red	Los registros se relacionan mediante enlaces direccionales pudiendo tener cada uno más de un hijo y de un padre. Codasyl. En desuso.
		<u>Modelo relacional</u>	Se basa en el algebra relacional y la teoría de conjuntos. Los datos se almacenan en relaciones (tablas). Fue propuesto por Codd. <u>Tema 3</u> .

³ de Miguel, Piattini y Marcos (1999)

	RAMA:	Informática	CICLO:	Desenvolvemento de Aplicacions Multiplataforma				
	MÓDULO:	Bases de datos				CURSO:	1º	
	PROTOCOLO:	Apuntes clases	EVAL:	1	FECHA:	2023/2024		
	UNIDAD		COMPETENCIA:					

3.3. Instancias y esquemas

Dentro de una base de datos se pueden definir los siguientes términos:

- Esquema: describe la estructura de la base de datos: los datos, sus relaciones, ...
- Instancia: la forman todos los datos, que un momento dado, tiene la base de datos. Cada vez que varían los datos almacenados en la base de datos se modifica la instancia generando una nueva. Se suele comparar con una fotografía de la base de datos en un momento dado en la que se capturan todos los datos de la misma.

El esquema de la base de datos es siempre el mismo lo que varía es la instancia al modificarse los datos que contiene la base de datos.

3.4. Lenguajes de datos

Son lenguajes proporcionados por el sistema gestor de bases de datos. Actualmente, en las bases de datos relacionales, todos estos lenguajes suelen estar integrados dentro del lenguaje SQL⁴.

Podemos encontrar los tres siguientes tipos de lenguajes:

3.4.1. Lenguaje de definición de datos⁵

El lenguaje de definición de datos (Data Definition Language o DDL) permite definir la estructura de la base de datos y los procedimientos y funciones que permiten consultarlos.

3.4.2. Lenguaje de manipulación de datos⁶

El lenguaje de Manipulación de Datos (Data Manipulation Language o DML) permite, a los usuarios, la consulta y manipulación de los datos contenidos en la base de datos.

Las operaciones que se pueden realizar son:

- Recuperación de la información.
- Inserción de nueva información.
- Actualización de la información ya existente en la base de datos.
- Borrado de información.

3.4.3. Lenguaje de control de datos⁷


El Lenguaje de Control de Datos (Data Control Language o DCL) permite al administrador, mediante una serie de privilegios, determinar tanto las acciones que un usuario puede realizar como los datos a los que puede acceder.

⁴ <https://es.wikipedia.org/wiki/SQL>

⁵ https://es.wikipedia.org/wiki/Lenguaje_de_definición_de_datos

⁶ https://es.wikipedia.org/wiki/Lenguaje_de_manipulación_de_datos

⁷ https://es.wikipedia.org/wiki/Lenguaje_de_control_de_datos

	RAMA:	Informática	CICLO:	Desenvolvimento de Aplicacions Multiplataforma		
	MÓDULO:	Bases de datos				CURSO: 1º
	PROTOCOLO:	Apuntes clases	EVAL:	1	FECHA:	2023/2024
	UNIDAD COMPETENCIA:					

4. Bibliografía

1. Sistema de información
2. Uso de base de datos:
 1. [Ranking de SGBD en bd-engines](#)
 2. [Ranking 2018 de SGBD en starkoverflow](#)
 3. [Ranking 2019 de SGBD en starkoverflow](#)
 4. [Ranking 2020 de SGBD en starkoverflow](#)
 5. [Ranking 2021 de SGBD en starkoverflow](#)
 6. [Ranking 2022 de SGBD en starkoverflow](#)
 7. [Ranking 2023 de SGBD en starkoverflow](#)
3. [Arquitectura en tres niveles](#)
4. [SQL](#)
5. Lenguajes de datos:
 1. [DDL - Lenguaje de definición de datos](#)
 2. [DML - Lenguaje de manipulación de datos](#)
 3. [DCL - Lenguaje de control de datos](#)