

A dark blue vertical bar is on the left. A blue arrow points right from it, containing the date.

3-10-2022

Actividades U1

Base de datos

Several thin, curved lines in dark blue and light grey originate from the bottom left and sweep upwards and to the right.

Abel Martínez Peinado
1º DUAL DAW

2. Los ficheros de información.	2
3. Bases de datos.	3
4. Evolución de los modelos.	5
5. Sistemas Gestores de Bases de Datos (SGBD).	5
8. Fragmentación.	6

2. Los ficheros de información.

1. Imagina que vas a desarrollar una aplicación para llevar el control de las ventas de ordenadores de una tienda de informática y en principio no vas a utilizar ningún gestor de BD. ¿Podrías crear dicha aplicación? ¿de qué elemento se dispone en un SO para guardar la información? ¿qué problemas crees que va a plantear este desarrollo? ¿de qué tipos pueden ser los elementos utilizados para guardar la información? Pon un ejemplo de la información que guardarías en cada uno de los tipos.

- Es posible, pero, sería inútil ya que los programas de gestión de base de datos ya lo automatizan.
- Sí podría crear la aplicación haciendo uso de un programa de gestión de base de datos.
- Redundancia e inconsistencia de los datos, dependencia de los datos física-lógica, dificultad para tener acceso a los datos, separación y aislamiento de los datos, dificultad para el acceso concurrente, dependencia de la estructura del archivo con el lenguaje de programación, problemas en la seguridad de los datos y problema de integridad de los datos.

- Tipos de ficheros:

• **Ficheros permanentes:**

- *Ficheros maestros.* (contienen el estado actual de los datos que pueden modificarse desde la aplicación.)
- *Ficheros constantes.* (son aquellos que incluyen datos fijos para la aplicación y estos no suelen ser modificados.)
- *Ficheros históricos.* (contienen datos que fueron considerados como actuales en un periodo o situación anterior.)

• **Ficheros temporales:**

- *Ficheros intermedios.* (almacenan resultados de una aplicación que serán utilizados por otra.)
- *Ficheros de maniobras.* (almacenan datos de una aplicación que no pueden ser mantenidos en memoria principal por falta de espacio.)
- *Ficheros de resultados.* (almacenan datos que van a ser transferidos a un dispositivo de salida)

2. Crea una tabla con las ventajas e inconvenientes que presentan las distintas formas de organizar los archivos o ficheros.

Tipos:	Ventajas	Desventajas
Organización secuencial	<ul style="list-style-type: none"> - Aprovecha al máximo el espacio del almacenamiento. - Todos los lenguajes de programación disponen de instrucciones para trabajar con este tipo de ficheros. - Se pueden grabar en cualquier tipo de soporte, tanto en secuenciales como direccionables. 	<ul style="list-style-type: none"> - Son ficheros mono usuario. - No se pueden Insertar registros entre los que ya están grabados. - Si se necesita actualizar un archivo con organización secuencial se debe crear de nuevo el archivo. - La única forma de acceder a él, es leyendo un registro tras otro desde el principio hasta el final.
Organización secuencial encadenada	<ul style="list-style-type: none"> - Esta organización es flexible. 	<ul style="list-style-type: none"> - Limitación a la consulta secuencial.
Organización secuencial indexada	<ul style="list-style-type: none"> - Permite localizar fácilmente los registros. - Se pueden actualizar los registros en el mismo fichero. - Tiene una zona de desbordamiento para los registros que no caben en la zona 	<ul style="list-style-type: none"> - Ocupa más espacio en el almacenamiento. - Obliga a una versión económica mayor.

Actividades U1

	de registros.	
Organización directa o aleatoria	<ul style="list-style-type: none"> - Posicionamiento inmediato. - Apertura del fichero de manera mixta. - Permite múltiples usuarios. - Permite realizar procesos en tiempo real. 	<ul style="list-style-type: none"> - Tienen un tamaño fijo. - Registros de longitud física. - Esta organización solo es posible en soportes direccionables.

3. Investiga en Internet sobre los sistemas lógicos de almacenamiento. ¿Qué son?, ¿qué características presentan? ¿cuáles son los más utilizados hoy en día?

- Es una estructura de directorios completa, que incluye un directorio raíz y cualquier subdirectorio y archivos por debajo suyo. En esta tabla se agrupan las tareas más sencillas que puede necesitar para mantener sistemas de archivos.

- Hoy en día se utilizan: FAT32, NTFS, ext4, HFS+.

- Características:

- **NTFS:** su principal característica es derechos de acceso y recursos compartidos de los archivos y carpetas pueden definirse al detalle y de manera integral.
- **FAT32:** limita a los archivos a no superar 4GB de almacenamiento.
- **ext4:** las particiones se pueden ampliar y reducir según sea necesario, e incluso durante el procesamiento. También tiene la función extents, que optimiza la gestión de archivos grandes y evita la fragmentación de manera más eficaz que sus predecesores.
- **HFS+:** admite hasta 4000 millones de bloques de archivos o carpetas.

3. Bases de datos.

4. Cita las ventajas de las BD sobre los sistemas de gestión de datos basados en ficheros.

- **Acceso múltiple:** diversos usuarios o aplicaciones podrán acceder a la base de datos.
- **Utilización múltiple:** cada uno de los usuarios o aplicaciones podrán disponer de división particular de la estructura de la base de datos.
- **Flexibilidad:** la forma de acceder a la información puede ser establecida de diferentes maneras, con esto se consiguen tiempos de respuesta muy reducidos.
- **Confidencialidad y seguridad:** el control del acceso a los datos podrá ser establecido para que los unos usuarios o aplicaciones puedan acceder a unos datos o a otros, impidiendo a los usuarios no autorizados la utilización de la base de datos.
- **Protección contra fallos:** en caso de errores en la información, existen mecanismos bien definidos que permiten la recuperación de los datos de forma fiable.
- **Independencia física:** un cambio de soporte físico de los datos, no afectaría a la BD o a las aplicaciones que acceden a ellos.
- **Independencia lógica:** los cambios en la BD no afectan a las aplicaciones que la usan.
- **Redundancia:** los datos se almacenan una única vez. Aunque se podría repetir la operación de forma controlada.
- **Interfaz de alto nivel:** mediante la utilización de lenguajes de alto nivel puede utilizarse la BD de manera sencilla y cómoda.
- **Consulta directa:** existen herramientas para poder acceder a los datos interactivamente.

5. En función de la ubicación de la información, ¿qué tipos de almacenamiento podemos tener? ¿qué ventajas e inconvenientes presenta cada uno de ellos?

- Discos, Cintas magnéticas, RAIDs, Almacenamiento en red y Almacenamiento en la nube.

Actividades U1

Ubicación de la información	Ventajas	Desventajas
Discos	<ul style="list-style-type: none"> - Con estos discos podemos montar NAS. - Trabajan de manera mixta (escritura/lectura). 	<ul style="list-style-type: none"> - Para ampliar este tipo de almacenamiento tienes que comprar más unidades de tipo HDD, SSD o SSD (NVMe). - Estos son sensibles a las vibraciones.
Cintas magnéticas	<ul style="list-style-type: none"> - Hay variedad de este tipo de almacenamiento. - Gran utilidad por su gran capacidad de almacenamiento. - Gran durabilidad y bajo coste. 	<ul style="list-style-type: none"> - Está un poco obsoleto ya que se utilizan otros tipos de almacenamiento. - No tiene tanta capacidad como los dispositivos de almacenamiento de hoy en día. - Este se rige por pistas en la cinta y estas son fácil de romper.
RAID	<ul style="list-style-type: none"> - RAID 0: con este tipo de RAID la información se escribe más rápido ya que los datos se escriben de manera alterna. - RAID 1: este es más seguro ya que copia los datos de forma paralela con la otra unidad de almacenamiento. - RAID 5: con este RAID la información se reparte de forma equivalente entre todas las unidades excepto en la última, que ahí se hace una copia de los datos. Con esto aumenta la velocidad de escritura y la seguridad. 	<ul style="list-style-type: none"> - RAID 0: los datos no se duplican y si se pierde un archivo no hay posibilidad de recuperarlo. - RAID 1: se necesitan más de una unidad de almacenamiento. - RAID 5: se debe tener un mínimo de 3 unidades de almacenamiento.
Almacenamiento en la red	<ul style="list-style-type: none"> - Lo bueno es que esto depende a cuanto velocidad vaya tu conexión a internet, cuanto más velocidad de esta conexión, mayor es la velocidad de escritura. - Te permite acceder a estos datos desde cualquier sitio, mientras que estés conectado a la misma red. - Optimiza el procesamiento de aplicaciones, etc. 	<ul style="list-style-type: none"> - Tienes que estar conectado a la misma red.
Almacenamiento en la nube	<ul style="list-style-type: none"> - Puedes acceder a los datos desde cualquier parte ya que se encuentran en un servidor. - Hay suscripciones a servicios en la nube gratuitos como DropBox, Google Drive, etc. 	<ul style="list-style-type: none"> - Hay suscripciones a servicios en la nube que son caros en caso de que sean de pago. - Tienes que tener constantemente conexión a Internet para acceder a los datos.

4. Evolución de los modelos.

6. Teniendo en cuenta los distintos modelos de BD, ¿qué tipo de BD es el más utilizado en la actualidad? ¿qué tipos de BD no se utilizan actualmente? ¿qué tipo de BD NoSQL están apareciendo en la actualidad y en qué tipo de aplicaciones se suelen utilizar?

- El más utilizado actualmente es la BD Relacional.
- Los menos utilizados son las BD Jerárquico y en Red.
- Modelos no SQL son: Modelo orientado a Documentos y orientado a Objetos.
- Para estos modelos se utilizan las aplicaciones: MongoDB, Cassandra, Redis y CouchDB.

7. Visita la URL <https://db-engines.com/en/ranking> e indica el modelo de BD más utilizado en las primeras 20 posiciones. ¿Qué otros modelos de BD aparecen en las primeras 20 posiciones? Investiga en Internet y describe los modelos que aparecen en las primeras 20 posiciones y no han sido tratados en los apuntes.

- El modelo que más se utiliza es el Relacional.
- Además del relacional también se utilizan: clave-valor, search engine, wide column, grafos, multi modelo.
- **Clave-valor:** Solo pueden almacenar pares de claves y valores, así como recuperar valores cuando se conoce una clave.
- **Search Engine:** Son sistemas de gestión de bases de datos NoSQL dedicados a la búsqueda de contenido de datos.
- **Wide Column:** almacena datos en registros con la capacidad de contener un gran número de columnas dinámicas. Dado que los nombres de columna y las claves de registro no son fijos, y dado que un registro puede tener miles de millones de columnas, los almacenes de columnas anchas pueden verse como almacenes bidimensionales de clave-valor.
- **Grafos:** representa los datos en estructuras gráficas como nodos y bordes, que son relaciones entre nodos. Permiten un fácil procesamiento de datos en esa forma y un cálculo simple de propiedades específicas del gráfico, como la cantidad de pasos necesarios para llegar de un nodo a otro.
- **Multi-Modelo:** es un sistema de administración de bases de datos diseñado para admitir múltiples modelos de datos en un solo backend integrado. Por el contrario, la mayoría de los sistemas de gestión de bases de datos están organizados en torno a un único modelo de datos que determina cómo se pueden organizar, almacenar y manipular los datos.

5. Sistemas Gestores de Bases de Datos (SGBD).

8. ¿Qué funciones realiza un SGBD? Describe cada una de ellas.

- **Función de descripción o definición:** permite al diseñador de la BD crear las estructuras apropiadas para integrar adecuadamente los datos. Esta función es la que permite definir las tres estructuras de la BD: estructura interna, estructura conceptual y estructura externa. Se lleva a cabo mediante el DDL.
- **Función de manipulación:** permite a los usuarios de la BD buscar, añadir, suprimir o modificar los datos de la misma, siempre de acuerdo con las especificaciones y las normas de seguridad dictadas por el administrador. Se lleva a cabo mediante DML.
- **Función de control:** permite al administrador de la BD establecer mecanismos de protección de las diferentes visiones de los datos asociadas a cada usuario, proporcionando elementos de creación y modificación de dichos usuarios. Se lleva a cabo mediante el DCL.

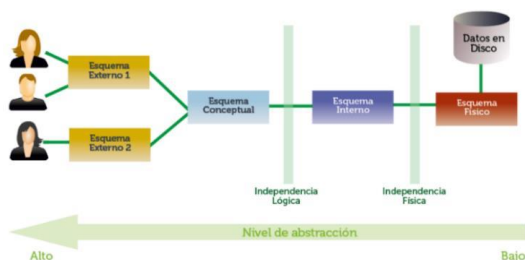
9. ¿Cuáles son los componentes de un SGBD? ¿qué función realiza cada uno de ellos? Describe y comenta la imagen del punto 5.2.

- **Lenguajes de la BD:** a través de estos lenguajes se pueden especificar los datos que componen la BD, su estructura, reglas de integridad, control de acceso, características físicas vistas externas de los usuarios. Los lenguajes son: DDL, DML y DCL.
- **El diccionario de datos:** se trata de la información útil para los programadores de las aplicaciones. Es el lugar donde se deposita la información sobre la totalidad de los datos que forman.
- **El gestor de la BD:** Este componente es el encargado de proporcionar una interfaz entre los datos almacenados y los programas de aplicación que los manejan. Es un intermediario entre el usuario y los datos. Es el encargado de garantizar la privacidad, seguridad e integridad de los datos, controlando los accesos concurrentes e interactuando con el sistema operativo.
- **Usuarios de la BD:** Existen muchos tipos de ellos.
 - **El administrador de la BD,** es la persona o conjunto de personas encargadas de la función de administración de la BD.
 - **Los usuarios de la BD,** son diferentes usuarios de la BD con diferentes necesidades sobre los datos, así como diferentes accesos y privilegios.
- **Herramientas de la BD:** son un conjunto de aplicaciones que permiten a los administradores la gestión de la BD, de los usuarios y permisos, generadores de formularios, informes, interfaces gráficas, generadores de aplicaciones, etc.

10. ¿Cuántos modelos conceptuales se pueden tener de una BD? ¿y externos? ¿qué diferencia hay entre ellos?

- **Modelos conceptuales:** 1.
- **Modelos externos:** todos los que quieras.
- **Diferencias:** en el nivel externo se describen las diferentes vistas que los usuarios percibirán de la BD y a nivel conceptual se describe la estructura completa de la BD a través de un esquema conceptual que detalla las entidades, atributos, relaciones, operaciones de los usuarios y restricciones.

11. Dada la imagen siguiente, describe los pasos que sigue el SGBD para interpretar una solicitud de datos de un usuario determinado. ¿Qué significa el nivel de abstracción?



- 1º El usuario solicita los datos creando una consulta.
- 2º El SGBD verifica y acepta el esquema externo para ese usuario.
- 3º Transforma la solicitud al esquema conceptual.
- 4º verifica y acepta el esquema conceptual.
- 5º Selecciona la o las tablas implicadas en la consulta y ejecuta la consulta.
- 6º Transforma el esquema interno conceptual, y del conceptual al externo.
- 7º Finalmente, el usuario ve los datos solicitados.

8. Fragmentación.

12. ¿Qué permiten las BD distribuidos? Enumera sus ventajas e inconvenientes.

- Permiten integrar la información de varias fuentes y la evolución de las tecnologías de comunicaciones, es decir puede acceder a los datos desde cualquier parte de la red.

Ventajas	Inconvenientes
El acceso y procesamiento de los datos es más rápido ya que varios nodos comparten carga de trabajo.	La probabilidad de violaciones de seguridad es creciente si no se toman las precauciones debidas.
Desde una ubicación puede accederse a información alojada en diferentes lugares.	Existe una complejidad añadida que es necesaria para garantizar la coordinación apropiada entre los

Actividades U1

	nodos.
Los costes son inferiores a los de las BD centralizados.	La inversión inicial es menor, pero el mantenimiento y control puede resultar costoso.
Existe cierta tolerancia a fallos. Mediante la replicación, si un nodo deja de funcionar el sistema completo no deja de funcionar.	Dado que los datos pueden estar replicados, el control de concurrencia y los mecanismos de recuperación son mucho más complejos que en un sistema centralizado.
El enfoque distribuido de las BD se adapta más naturalmente a la estructura de las organizaciones. Permiten la incorporación de nodos de forma flexible.	El intercambio de mensajes y el cómputo adicional necesario para conseguir la coordinación entre los distintos nodos constituyen una forma de sobrecarga que no surge en los sistemas centralizados.
Aunque los nodos están interconectados, tienen independencia local.	Dada la complejidad del procesamiento entre nodos es difícil asegurar la corrección de los algoritmos, el funcionamiento correcto durante un fallo o la recuperación.

13. Considera la siguiente relación, ¿qué fragmentos se obtendrían al realizar una fragmentación horizontal por ESCUELA?

FH1 [ESCUELA=" CIME"]

Jno	Nombre	Nota	Escuela
J1	LUIS YÁNEZ	8	CIME
J2	ERIKA QUIROZ	8	CIME

FH2 [ESCUELA=" EISIC"]

Jno	Nombre	Nota	Escuela
J3	DANIEL MURILLO	9	EISIC
J4	Mª JOSÉ MENDEZ	10	EISIC
J5	ADONIS PABÓN	9	EISIC

14. Considera las siguientes relaciones, ¿qué fragmentos principales y derivados se obtendrían al realizar una fragmentación horizontal por ESCUELA?

FH1 [ID#, ESCUELA=" EISIC"] Principal

ID	MATERIA	ESCUELA	CRÉDITOS	NIVEL
1	Análisis Matemático	EISIC	4	1
2	Sistemas Operativos	EISIC	6	3
4	Tecnología Eléctrica	EISIC	4	2

FH2 [ID#, ESCUELA=" CIME"] Principal

ID	MATERIA	ESCUELA	CRÉDITOS	NIVEL
3	Programación II	CIME	6	2
5	Técnicas de aprendizaje	CIME	4	1
6	Dibujo Mecánico	CIME	6	3

Actividades U1

FH3 [Jno#, ESCUELA=" EISIC"] Derivado

Jno	NOMBRE	MATERIA	NOTA
J1	LUIS YÁNEZ	Sistemas Operativos	8
J4	Mª JOSE MENDEZ	Análisis Matemático	10

FH4 [Jno#, ESCUELA=" CIME"] Derivado

Jno	NOMBRE	MATERIA	NOTA
J2	ERIKA QUIROZ	Dibujo Mecánico	8
J3	DANIEL MURILLO	Técnicas de aprendizaje	9
J5	ADONIS PABÓN	Programación II	9

15. Considera la relación del ejercicio 13. Realiza una fragmentación vertical por NOTA y por NOMBRE/ESCUELA.

FV1= [Jno#, NOTA]

Jno	NOTA
J1	8
J2	8
J3	9
J4	10
J5	9

FV2= [Jno#, NOMBRE, ESCUELA]

Jno	NOMBRE	ESCUELA
J1	LUIS YÁNEZ	CIME
J2	ERIKA QUIROZ	CIME
J3	DANIEL MURILLO	EISIC
J4	M.ª JOSE MENDEZ	EISIC
J5	ADONIS PABON	EISIC

16. Realiza una fragmentación mixta HV. La fragmentación horizontal se debe realizar por ESCUELA y la vertical por ESCUELA/NOTA_INGRESO/NOMBRE y NOMBRE/BECA.

FH1 [Escuela=" EUI"]

FV1 [DNI#, Escuela, Nota ingreso, Nombre]

FV2 [DNI#, Nombre, Beca]

DNI#	Escuela	Nota ingreso	Nombre	DNI#	Nombre	Beca
87633483	EUI	5.6	Concha Queta	87633483	Concha Queta	No
99855743	EUI	7.2	Josechu Letón	99855743	Josechu Letón	Si
05399075	EUI	5.0	Bill Gates	05399075	Bill Gates	No
44543324	EUI	7.5	Maite Clado	44543324	Maite Clado	Si

FH2 [Escuela="EUIT"]

FV3 [DNI#, Escuela, Nota ingreso, Nombre]

FV4 [DNI#, Nombre, Beca]

DNI#	Escuela	Nota ingreso	Nombre	DNI#	Nombre	Beca
------	---------	--------------	--------	------	--------	------

Actividades U1

33887293	EUIT	6.1	Oscar Romato	33887293	Oscar Romato	Si
44343234	EUIT	8.0	Pepe Pótamo	44343234	Pepe Pótamo	No
66553234	EUIT	6.6	Ernetto Mate	66553234	Ernetto Mate	No