

## U.T. 1 – Actividad 1a. Sistemas de almacenamiento de información

Para realizar esta práctica nos serviremos de los contenidos del tema realizados por el profesor y de Internet. Buscaremos respuestas a estas preguntas donde la misma tenga que ver con todo lo explicado anteriormente.

### 1. ¿Qué es un archivo o fichero?

Un archivo o fichero informático es un conjunto de bits que son almacenados en un dispositivo. Un archivo es identificado por un nombre y la descripción de la carpeta o directorio que lo contiene. A los archivos informáticos se les llama así porque son los equivalentes digitales de los archivos escritos en expedientes, tarjetas, libretas, papel o microfichas del entorno de oficina tradicional.

### 2. Tipos de ficheros

Pueden organizarse en dos grandes categorías: ficheros de texto (o ASCII) y ficheros binarios.

- **Ficheros de texto (o ASCII):** Contienen líneas de texto representadas en código ASCII.

- Los ficheros de programas fuente escritos en cualquier lenguaje de programación son ejemplos de ficheros ASCII. No obstante, estos ficheros pueden contener cualquier tipo de información.
- La gran ventaja de este tipo de archivos es que son estándar en todo tipo de sistemas (MS-DOS, Windows, MacOS, Unix, etc.). Son manejados mediante un tipo de programa, estándar en todo sistema, conocido como editor de textos. (Bloc de notas y EDIT en Windows.)
- Debido a su uso estandarizado, la práctica totalidad de herramientas ofimáticas (procesadores de texto, hojas de cálculo, gestores de bases de datos, etc.) tienen la capacidad de manejar este tipo de ficheros.

- **Ficheros binarios:** Son aquellos que no son ficheros ASCII. Ejemplos típicos de ficheros binarios son:

- Ficheros ejecutables: Contienen programas en formato ejecutable, como por ejemplo, los ficheros “.exe” en Windows.
- Ficheros de datos de aplicaciones: Por ejemplo los “.doc” de Word, los “.xls” de Excel, etc. Estos ficheros utilizan un formato que es específico de la aplicación que los genera.
- Ficheros multimedia: Almacenan información multimedia, es decir, imagen (jpeg por ejemplo), sonido (mp3 por ejemplo) o vídeo (mpeg por ejemplo).

### 3. ¿Qué es una Base de Datos?

Una **base de datos** (cuya abreviatura es BD) es una entidad en la cual se pueden almacenar datos de manera estructurada, con la menor redundancia posible. Diferentes programas y diferentes usuarios deben poder utilizar estos datos.

### 4. Diferencias entre un sistema convencional de archivos y un sistema gestor de Base de Datos

#### **Sistema de archivos:**

1. Manejo de archivos en texto plano.
2. Búsqueda de tres tipos: secuencial, aleatorio, indexado.
3. No existen las actualizaciones, se simulan dando de baja el registro dando de alta el nuevo con la modificación.
4. Tienen independencia lógica y física la mayoría de ellos.
5. Cuenta con redundancia no controlada y errores de referencias.
6. Sus componentes son: archivos, registros, campos.

#### **Sistemas de bases de datos:**

1. Existe redundancia controlada e integridad referencia.
2. Implementa mayor seguridad.
3. Existen actualizaciones directas.
4. No tienen independencia ni lógica ni física.
5. Sus componentes son Hw, Sw, BD y Usuarios.

### 5. Operaciones totales sobre archivos

**Creación:** Escritura de todos sus registros. Es la primera operación que sufrirá el archivo de datos. Implica la elección de un entorno descriptivo que permita un ágil, rápido y eficaz tratamiento del archivo.

**Consulta:** Lectura de todos sus registros. Es la operación que permite al usuario acceder al archivo de datos para conocer el contenido de uno, varios o todos los registros.

Actualización. Inserción supresión o modificación de algunos de sus registros.

**Clasificación:** Reubicación de los registros de tal forma que queden ordenados según determinados criterios. Una operación muy importante en un archivo es la clasificación u ordenación (sort, en inglés). Esta clasificación se realizará de acuerdo con el valor de un campo específico, pudiendo ser ascendente (creciente) o descendente (decreciente): alfabética o numérica.

**Borrado:** Eliminando total del archivo, dejando libre el espacio del soporte que ocupaba. Es la operación inversa a la creación de un archivo (kill, en inglés). Cuando se destruye (anula o borra) un archivo, éste ya no se puede utilizar y, por consiguiente, no se podrá acceder a ninguno de sus registros.

**Reorganización de un archivo:** Las operaciones sobre archivos modifican la estructura inicial o la óptima de un archivo. Los índices, enlaces (punteros), zonas de sinónimos, zonas de desbordamiento, etc., se modifican con el paso del tiempo, lo que hace a la operación de acceso al registro cada vez más lenta. La reorganización suele consistir en la copia de un nuevo archivo a partir del archivo modificado, a fin de obtener una nueva estructura lo más óptima posible.

## 6. Define brevemente el concepto de registro y campo

En informática, un campo es un espacio de almacenamiento para un dato en particular. En las bases de datos, un campo es la mínima unidad de información a la que se puede acceder; un campo o un conjunto de ellos forman un registro, donde pueden existir campos en blanco, siendo este un error del sistema operativo.

## 7. Operaciones sobre registros

Sobre los registros de una base de datos podemos hacer las siguientes operaciones que son:

Update > Actualizar o modificar el registro

Select > Consulta sobre un registro

Create, Insert > Crear o almacenar datos sobre un registro

Delete, Drop > Borrado de uno o varios registros

## 8. Factores de utilización de los archivos

Para utilizar un archivo debemos tener en cuenta:

**Índice de Volatilidad;** Un archivo es volátil cuando tiene un alto porcentaje de adiciones y supresiones debido al ingreso o eliminación de registros respecto al número promedio de registros que haya en el archivo.

**Índice de Actividad;** Un archivo es activo cuando tiene un alto porcentaje de utilidad sea de actualización o consulta en un periodo de tiempo fijo respecto al número promedio de registro que se encuentran en el archivo.

El índice de actividad suele emplearse para saber si un archivo puede explotarse como una organización secuencial o relativa

## 9. Explica qué es un SGBD y enumera sus funciones

Un **sistema gestor de base de datos (SGBD)** es un conjunto de programas que permiten el almacenamiento, modificación y extracción de la información en una base de datos, además de proporcionar herramientas para añadir, borrar, modificar y analizar los datos. Los usuarios pueden acceder a la información usando herramientas específicas de consulta y de generación de informes, o bien mediante aplicaciones al efecto. Estos sistemas también proporcionan métodos para mantener la integridad de los datos, para administrar el acceso de usuarios a los datos y para recuperar la información si el sistema se corrompe. Permiten presentar la información de la base de datos en variados formatos. La mayoría incluyen un generador de informes. También pueden incluir un módulo gráfico que permita presentar la información con gráficos y tablas. Generalmente se accede a los datos mediante lenguajes de consulta, lenguajes de alto nivel que simplifican la tarea de construir las aplicaciones. También simplifican las consultas y la presentación de la información. Un SGBD permite controlar el acceso a los datos, asegurar su integridad, gestionar el acceso concurrente a ellos, recuperar los datos tras un fallo del sistema y hacer copias de seguridad. Las bases de datos y los sistemas para su gestión son esenciales para cualquier área de negocio, y deben ser gestionados con esmero.

## 10. Componentes de un Sistema Gestor de Base de Datos

El **motor de la base de datos** acepta peticiones lógicas de los otros subsistemas del SGBD, las convierte en su equivalente físico y accede a la base de datos y diccionario de datos en el dispositivo de almacenamiento.

El **subsistema de definición de datos** ayuda a crear y mantener el diccionario de datos y define la estructura del fichero que soporta la base de datos.

El **subsistema de manipulación** de datos ayuda al usuario a añadir, cambiar y borrar información de la base de datos y la consulta para extraer información. El subsistema de manipulación de datos suele ser la interfaz principal del usuario con la base de datos. Permite al usuario especificar sus requisitos de la información desde un punto de vista lógico.

El **subsistema de generación de aplicaciones** contiene utilidades para ayudar a los usuarios en el desarrollo de aplicaciones. Usualmente proporciona pantallas de entrada de datos, lenguajes de programación e interfaces.

El **subsistema de administración** ayuda a gestionar la base de datos ofreciendo funcionalidades como almacenamiento y recuperación, gestión de la seguridad, optimización de preguntas, control de concurrencia y gestión de cambios.

## 11. Organización de una Base de Datos

Un sistema de base de datos está formado por diferentes módulos y cada uno se encarga de realizar funciones concretas, aunque generalmente algunas de esas funciones las puede realizar el sistema operativo.

Un sistema de base de datos está formado por diferentes módulos y cada uno se encarga de realizar funciones concretas, aunque generalmente algunas de esas funciones las puede realizar el sistema operativo. Las funciones que desarrolla el sistema gestor de la base de datos (aunque no siempre) se pueden incluir en dos módulos:

**Procesador de consultas:** para ello dispone de los siguientes módulos:

- a) Compilador del lenguaje de manejo de datos (DML): que traduce las instrucciones SQL a instrucciones de bajo nivel y transforma las solicitudes del usuario en otras más eficientes, las optimiza.
- b) Pre compilador de lenguaje de manejo de datos (DML): que convierte las instrucciones SQL embebidas en otro lenguaje de programación en instrucciones que entiende el procesador de consultas.
- c) Interpretador del lenguaje de definición de datos (DDL): que interpreta las instrucciones DDL y realiza las anotaciones correspondientes en el catálogo de la base de datos.
- d) Motor de evaluación de consultas: ejecuta las instrucciones de bajo nivel generadas por el compilador del DML.

**Gestión del almacenamiento:** es el interfaz entre los datos almacenados en el disco y las instrucciones de bajo nivel obtenidas por el procesador de consultas, está compuesto por los siguientes módulos:

- a) Gestor de autorización e integridad: comprueba que se satisfagan las reglas de integridad impuestas y las debidas autorizaciones de los usuarios para acceder a los datos.
- b) Gestor de transacciones: se asegura de que la base de datos tenga un estado consistente, aunque existan fallos del sistema y de que las ejecuciones de transacciones consistentes ocurran sin conflictos.
- c) Gestor de archivos: gestiona la reserva de espacio de almacenamiento en disco y las estructuras de datos usadas para representar la información.
- d) Gestor de memoria interna: es el responsable de traer los datos del disco a memoria principal.

## 12. Esquema y sub-esquema de una Base de Datos

Un **esquema** consiste en una definición de todos los tipos de objetos de la aplicación, incluyendo sus atributos, relaciones y restricciones estáticas. Correspondientemente, existirá un repositorio de información, la *base de datos*, que es una instancia del esquema. Un determinado tipo de procesos sólo necesita acceder a un subconjunto predeterminado de entidades definidas en un esquema, por lo que este tipo de procesos puede requerir sólo un subconjunto de las propiedades estáticas del esquema general. A este subconjunto de propiedades estáticas se le denomina **sub-esquema**.

## 13. Modelos de Bases de Datos

Los *Modelos* más comunes de organización de *Bases de Datos* son:

- *Jerárquico*.
- *En Red*.
- *Relacional*.
- *Orientado a Objetos*

## 14. Elementos Software de una Base de Datos

### ➤ Programas (Software)

Un sistema de base de datos incluye dos tipos de programas:

- El *software de propósito general*, para la gestión de la base de datos, comúnmente llamado *Sistema Gestor de Bases de Datos* (*S. G.B.D.*, o también *D. B. M S*, en inglés). El S.G.B.D maneja todas las solicitudes de acceso a la base de datos formuladas por los usuarios y los programas de aplicación.
- El *software de aplicación*, que usa las facilidades del S.G.B.D para manipular la base de datos con el fin de llevar a cabo una función específica en la gestión de la empresa (por ejemplo: la gestión de almacén). Puede estar desarrollado en un lenguaje de programación estándar, tal como COBOL o C, o en un lenguaje propio de los S.G.B.D denominados lenguajes de cuarta generación (4GL).

## 15. Elementos personales de una Base de Datos

**Datos:** Es la parte esencial de la información, es decir, la información que llega a la base de datos.

**Atributos:** Son los diferentes campos que conforman la estructura de una base de datos.

**Campos:** Es la unidad más pequeña de datos.

**Registro:** Es un conjunto de campos o atributos relacionados entre sí.

**Archivo:** Es un conjunto de registros relacionados.

## 16. Haz una lista de las características elementales que debe reunir toda base de datos.

Un SGBD permite el almacenamiento, manipulación y consulta de datos pertenecientes a una base de datos organizada en uno o varios ficheros. En el modelo más extendido (base de datos relacional) la base de datos consiste, de cara al usuario, en un conjunto de tablas entre las que se establecen relaciones. A pesar de sus semejanzas (ambos manejan conjuntos de tablas) existen una serie de diferencias fundamentales entre un SGBD y un programa de hoja de cálculo, la principal es que un SGBD permite:

- El método de almacenamiento y el programa que gestiona los datos (servidor) son independientes del programa desde el que se lanzan las consultas (cliente) .
- En lugar de primarse la visualización de toda la información, el objetivo fundamental es permitir consultas complejas, cuya resolución está optimizada, expresadas mediante un lenguaje formal.
- El almacenamiento de los datos se hace de forma eficiente aunque oculta para el usuario y normalmente tiene, al contrario de lo que ocurre con las hojas de cálculo, poco que ver con la estructura con la que los datos se presentan al usuario.
- El acceso concurrente de múltiples usuarios autorizados a los datos, realizando operaciones de actualización y consulta de los mismos garantizando la ausencia de problemas de seguridad (debidos a accesos no autorizados) o integridad (pérdida de datos por el intento de varios usuarios de acceder al mismo fichero al mismo tiempo).

El programa servidor suele activarse al arrancar el ordenador, podría compararse a un bibliotecario que recibe peticiones (consultas) de diferentes programas clientes de base

de datos, consulta la base de datos y entrega al cliente el resultado de la consulta realizada. Si dos usuarios solicitan al mismo tiempo una modificación de los datos, el programa servidor se encarga de hacerlas ordenadamente para evitar perder datos (lo que ocurriría si ambos usuarios abrieran y modificaran a la vez un fichero con la base de datos).

El diseño de una base de datos implica codificar en formato digital ciertos aspectos del mundo real. Esta codificación implica los mismos 3 pasos que ya se mencionaron en el tema 2, es decir:

- Modelo conceptual
- Modelo lógico
- Modelo digital o implementación física de la base de datos

Hoy en día existen dos grandes modelos, las **bases de datos relacionales** y el **modelo orientado a objetos (modelo OO)**, y un modelo híbrido denominado **modelo Objeto-Relacional (modelo OR)**. En cualquier manual de bases de datos puede encontrarse información acerca de modelos más antiguos.

17. Busca en Internet las referencias de las últimas versiones de los siguientes productos relacionados con la gestión de bases de datos:

a) Microsoft Access

<https://support.office.com/es-es/article/Novedades-de-Access-2016-76454345-f85d-47af-ace1-98a456cb3496>

b) ORACLE

<https://www.oracle.com/es/corporate/features/database-12c/index.html>

c) MySQL

<http://www.mysql.com/products/>

d) PostgreSQL

<http://www.postgresql.org.es/aggregator/sources/1>

e) SQL Server

<https://www.microsoft.com/es-es/server-cloud/products/sql-server/overview.aspx>



18. Busca información sobre todos los sistemas de base de datos que conozcas y haz una tabla con un breve resumen comparativo de las siguientes características.

- Software libre o propietario.
- Realiza control de transacciones.
- Control de usuarios.
- Requerimientos hardware.
- De servidor o monopuesto.
- Gestión de copia de seguridad.
- Tipo de Almacenamiento físico.

### MySQL:

- Es software dual, puedes usarla en tu proyecto gratuitamente pero tu proyecto tendrá que ser publicado como software libre y si no quieres hacer esto último debes pagar una licencia de uso de MySQL.
- MySQL incluye control de transacciones desde la versión 3.23
- Incluye control de usuarios
- Sistema Operativo: Windows de 32 bits, tal como 9x, Me, NT, 2000, XP, o Windows Server 2003. Se recomienda fuertemente el uso de un sistema operativo Windows basado en NT (NT, 2000, XP, 2003) puesto que éstos permiten ejecutar el servidor MySQL como un servicio. Soporte para protocolo TCP/IP. Suficiente espacio en disco rígido para descomprimir, instalar, y crear las bases de datos de acuerdo a sus requisitos. Generalmente se recomienda un mínimo de 200 megabytes.
- La instalación se ejecuta en servidor en modo servidor
- Incluye varias herramientas para hacer copias de seguridad
- Internamente un mecanismo de almacenamiento (*storage engine*) es el encargado de almacenar en último término los datos de las tablas en dispositivos físicos, para que estos tengan durabilidad. El mecanismo es totalmente clave a la hora de evaluar la rapidez y las funcionalidades que puede ofrecer el SGBD. MySQL tiene la opción de incluir dinámicamente - desde MySQL 5.1- los distintos mecanismos para soportar el almacenamiento de las tablas. Estos mecanismos aplican igualmente a los desarrollos derivados de MySQL como MariaDB, Percona Server, WebScaleSQL y Drizzle.

## SQL-Server:

- Este tipo de software es propietario y que pertenece a Microsoft, pero puede ser usado sin previo pago.
- SQL Server incluye control de transacciones desde la versión del 2008.
- SQL Server incluye control de usuarios.

|  |   |
|--|---|
| Procesador de 1,5 GHz de arquitectura x64 o compatible, doble núcleo | CPU de 2 GHz de arquitectura x64, cuatro núcleos o superior, como los sistemas AMD Opteron o Intel Xeon |
|--|---|

|             |                    |
|-------------|--------------------|
| 4 GB de RAM | 16 GB de RAM o más |
|-------------|--------------------|

|   |   |
|---|---|
| Matriz de discos duros SAS RAID 5 o RAID 10 | Matriz de discos duros SAS RAID 5 o RAID 10 |
|---|---|

- SQL Server se puede ejecutar tanto en servidor como en mono puesto.
- El componente de copia de seguridad y restauración de SQL Server ofrece una protección esencial para los datos críticos almacenados en las bases de datos de SQL Server.

## U.T. 1 – Actividad 1b. Modelización de la información

### 19. ¿Qué es un modelo de datos?

Un modelo de datos es un lenguaje orientado a hablar de una Base de Datos. Típicamente un modelo de datos permite describir: Las estructuras de datos de la base: El tipo de los datos que hay en la base y la forma en que se relacionan.

### 20. Etapas del diseño de una base de datos

**Etapas del diseño conceptual:** en esta etapa se obtiene una estructura de la información de la futura BD independiente de la tecnología que hay que emplear. No se tiene en cuenta todavía qué tipo de base de datos se utilizará –relacional, orientada a objetos, jerárquica, etc...; en consecuencia, tampoco se tiene en cuenta con qué SGBD ni con qué lenguaje concreto se implementará la base de datos. Así pues, la etapa del diseño conceptual nos permite concentrarnos únicamente en la problemática de la estructuración de la información, sin tener que preocuparnos al mismo tiempo de resolver cuestiones tecnológicas.

**Etapas del diseño lógico:** en esta etapa se parte del resultado del diseño conceptual, que se transforma de forma que se adapte a la tecnología que se debe emplear. Más concretamente, es preciso que se ajuste al modelo del SGBD con el que se desea implementar la base de datos. Por ejemplo, si se trata de un SGBD relacional, esta etapa obtendrá un conjunto de relaciones con sus atributos, claves primarias y claves foráneas.

**Etapas del diseño físico:** en esta etapa se transforma la estructura obtenida en la etapa del diseño lógico, con el objetivo de conseguir una mayor eficiencia; además, se completa con aspectos de implementación física que dependerán del SGBD. Por ejemplo, si se trata de una base de datos relacional, la transformación de la estructura puede consistir en lo siguiente: tener almacenada alguna relación que sea la combinación de varias relaciones que se han obtenido en la etapa del diseño lógico, partir una relación en varias, añadir algún atributo calculable a una relación, etc. Los aspectos de implementación física que hay que completar consisten normalmente en la elección de estructuras físicas de implementación de las relaciones, la selección del tamaño de las memorias intermedias (buffers) o de las páginas, etc.

## 21. Tipos de abstracciones

La **abstracción de datos** es una técnica o metodología que permite diseñar estructuras de datos. La abstracción consiste en representar bajo ciertos lineamientos de formato las características esenciales de una estructura de datos. Este proceso de diseño evita los detalles específicos de implementación de los datos.

Un sistema de base de datos mostrará la información que un usuario requiera, la forma de mostrarla dependerá del usuario mismo, escondiendo la complejidad (como la forma de tratar los datos, el cómo están almacenados, detalles que el usuario no necesita conocer) a los usuarios para simplificar su interacción con el sistema.

Este proceso se realiza a través de varios niveles de abstracción:

- **Nivel Físico:** Describe *cómo* se almacenan realmente los datos. Se describen en detalle las estructuras de datos complejas de bajo nivel.
- **Nivel Lógico:** Es el siguiente nivel más alto de abstracción. Describe *qué* datos se almacenan en la base de datos y qué relaciones existen entre esos datos. Los administradores de bases de datos, que deben decidir la información que se mantiene en la base de datos, usan el nivel lógico de abstracción.
- **Nivel de Vistas:** El nivel más alto de abstracción describe sólo parte de la base de datos completa. Los usuarios que utilizan este nivel de abstracción tienen una gran simplificación en su interacción con el sistema, ya que ven un conjunto de programas de aplicación (vistas) que esconden los detalles de los tipos de datos. Además de esconder detalles de nivel lógico también proporcionan un mecanismo de seguridad para evitar que los usuarios accedan a ciertas partes, prohibidas para ellos, de la base de datos.

## 22. Los modelos lógicos

Son orientados a las operaciones más que a la descripción de una realidad. Usualmente están implementados en algún Manejador de Base de Datos. El ejemplo más típico es el Modelo Relacional, que cuenta con la particularidad de contar también con buenas características conceptuales (Normalización de bases de datos).

## 23. El modelo jerárquico

Una **base de datos** jerárquica es un tipo de sistema de gestión de **bases de datos** que, como su nombre indica, almacena la información en una estructura jerárquica que enlaza los registros en forma de estructura de árbol (similar a un árbol visto al revés), en donde un nodo padre de información puede tener varios nodos.

## 24. El modelo en red

Una **base de datos de red** es una base de datos conformada por una colección o set de registros, los cuales están conectados entre sí por medio de enlaces en una red. El registro es similar al de una entidad como las empleadas en el modelo relacional.

## 25. El modelo relacional

El **modelo relacional**, para el modelado y la gestión de bases de datos, es un modelo de datos basado en la lógica de predicados y en la teoría de conjuntos.

- Provee herramientas que garantizan evitar la duplicidad de registros.
- Garantiza la integridad referencial, así, al eliminar un registro elimina todos los registros relacionados dependientes.
- Favorece la normalización por ser más comprensible y aplicable.