

República Bolivariana de Venezuela
Ministerio Popular de Educación y Deportes
UNEFA
Cátedra: Base de Datos
Unidad I. Introducción

Dato: Hecho o valor a partir del cual se puede inferir una conclusión.

Los datos son aquello que un programa manipula. Sin datos un programa no funcionaría correctamente. Los programas manipulan datos de manera muy diferente según el *tipo* de dato del que se trate. Y hay varios de estos tipos:

Información: es un conjunto de datos acerca de algún suceso, hecho, fenómeno o situación, que organizados en un contexto determinado tienen su significado, cuyo propósito puede ser el de reducir la incertidumbre o incrementar el conocimiento acerca de algo.

Diferencia entre Datos e información

1. Los Datos a diferencia de la información son utilizados como diversos métodos para comprimir la información a fin de permitir una transmisión o almacenamiento más eficaces.
2. Aunque para el procesador de la computadora hace una distinción vital entre la información entre los programas y los datos, la memoria y muchas otras partes de la computadora no lo hace. Ambos son registradas temporalmente según la instrucción que se le dé. Es como un pedazo de papel no sabe ni le importa lo que se le escriba: un poema de amor, las cuentas del banco o instrucciones para un amigo. Es lo mismo que la memoria de la computadora. Sólo el procesador reconoce la diferencia entre datos e información de cualquier programa. Para la memoria de la computadora, y también para los dispositivos de entrada y salida (E/S) y almacenamiento en disco, un programa es solamente más datos, más información que debe ser almacenada, movida o manipulada.
3. La cantidad de información de un mensaje puede ser entendida como el

número de símbolos posibles que representan el mensaje. "los símbolos que representan el mensaje no son más que datos significativos.

4. En su concepto más elemental, la información es un mensaje con un contenido determinado emitido por una persona hacia otra y, como tal, representa un papel primordial en el proceso de la comunicación, a la vez que posee una evidente función social. A diferencia de los datos, la información tiene significado para quien la recibe, por eso, los seres humanos siempre han tenido la necesidad de cambiar entre sí información que luego transforman en acciones. "La información es, entonces, conocimientos basados en los datos a los cuales, mediante un procesamiento, se les ha dado significado, propósito y utilidad"

Conocimiento: es lo útil que está en nuestras cabezas, son nuestras ideas, nuestra experiencia, es lo que sabemos sobre cómo funcionan las cosas o para hacer algo mejor.

Sistema de Información: Un sistema de información es un conjunto de elementos que interactúan entre sí con el fin de apoyar las actividades de una empresa o negocio.

Sistema de Procesamiento de Datos

Es el conjunto de métodos y procedimientos que permiten obtener información. Estos métodos pueden ser distinguidos entre manuales, electromecánicos y electrónicos, el último de ellos es llamado Procesamiento Electrónico de Datos o popularmente Computación.

Sistema de Procesamiento del Conocimiento

El proceso del Conocimiento se define como un mecanismo de conocimiento dentro de las organizaciones como un proceso de conocimiento, dividiéndolo en tres etapas:

- Generación del Conocimiento.

- Codificación del Conocimiento.
- Transferencia del Conocimiento.

El “descubrimiento” del conocimiento es el proceso por el cual ampliamos la cantidad y calidad de nuestro almacén de conocimiento. Esto se puede llevar a cabo a través de una serie de procesos que incluyen la lectura, escritura, conferencias, trabajo en equipo, sueños diarios o trabajo en un equipo directivo.

El objetivo de la codificación es colocar al conocimiento en alguna forma legible, entendible y organizada, para que pueda ser utilizado por todas las personas que necesiten de él.

Sistema de Archivos

Es un método para el almacenamiento y organización de archivos de computadora y los datos que estos contienen, para hacer más fácil la tarea encontrarlos y accederlos. Los sistemas de archivos son usados en dispositivos de almacenamiento como discos duros y CD-ROM e involucran el mantenimiento de la localización física de los archivos.

Más formalmente, un sistema de archivos es un conjunto de tipo de datos abstractos que son implementados para el almacenamiento, la organización jerárquica, la manipulación, el acceso, el direccionamiento y la recuperación de datos. Los sistemas de archivos comparten mucho en común con la tecnología de las bases de datos.

En general, los sistemas operativos tienen su propio sistema de archivos. En ellos, los sistemas de archivos pueden ser representados de forma textual (ej.: el shell de DOS) o gráficamente (ej.: Explorador de archivos en Windows) utilizando un gestor de archivos.

Base de Datos

Se define una base de datos como una serie de datos organizados y relacionados entre sí, los cuales son recolectados y explotados por los sistemas de información de una empresa o negocio en particular.

Características

Entre las principales características de los sistemas de base de datos podemos mencionar:

- Independencia lógica y física de los datos.
- Redundancia mínima.
- Acceso concurrente por parte de múltiples usuarios.
- Integridad de los datos.
- Consultas complejas optimizadas.
- Seguridad de acceso y auditoría.
- Respaldo y recuperación.
- Acceso a través de lenguajes de programación estándar.

Sistema de Gestión de Base de Datos (SGBD)

Los Sistemas de Gestión de Base de Datos (en inglés DataBase Management System) son un tipo de software muy específico, dedicado a servir de interfaz entre la base de datos, el usuario y las aplicaciones que la utilizan. Se compone de un lenguaje de definición de datos, de un lenguaje de manipulación de datos y de un lenguaje de consulta.

Ventajas de las bases de datos

Control sobre la redundancia de datos: Los sistemas de ficheros almacenan varias copias de los mismos datos en ficheros distintos. Esto hace que se desperdicie espacio de almacenamiento, además de provocar la falta de consistencia de datos.

- **Consistencia de datos:** Eliminando o controlando las redundancias de datos se reduce en gran medida el riesgo de que haya inconsistencias. Si un dato está almacenado una sola vez, cualquier actualización se debe realizar sólo una vez, y está disponible para todos los usuarios inmediatamente. Si un dato está duplicado y el sistema conoce esta redundancia, el propio sistema puede encargarse de garantizar que todas las copias se mantienen consistentes.
- **Compartición de datos:** En los sistemas de ficheros, los ficheros

pertenecen a las personas o a los departamentos que los utilizan. Pero en los sistemas de bases de datos, la base de datos pertenece a la empresa y puede ser compartida por todos los usuarios que estén autorizados.

- **Mantenimiento de estándares:** Gracias a la integración es más fácil respetar los estándares necesarios, tanto los establecidos a nivel de la empresa como los nacionales e internacionales. Estos estándares pueden establecerse sobre el formato de los datos para facilitar su intercambio, pueden ser estándares de documentación, procedimientos de actualización y también reglas de acceso.
- **Mejora en la integridad de datos:** La integridad de la base de datos se refiere a la validez y la consistencia de los datos almacenados. Normalmente, la integridad se expresa mediante restricciones o reglas que no se pueden violar. Estas restricciones se pueden aplicar tanto a los datos, como a sus relaciones, y es el SGBD quien se debe encargar de mantenerlas.
- **Mejora en la seguridad:** La seguridad de la base de datos es la protección de la base de datos frente a usuarios no autorizados. Sin unas buenas medidas de seguridad, la integración de datos en los sistemas de bases de datos hace que éstos sean más vulnerables que en los sistemas de ficheros.
- **Mejora en la accesibilidad a los datos:** Muchos SGBD proporcionan lenguajes de consultas o generadores de informes que permiten al usuario hacer cualquier tipo de consulta sobre los datos, sin que sea necesario que un programador escriba una aplicación que realice tal tarea.
- **Mejora en la productividad:** El SGBD proporciona muchas de las funciones estándar que el programador necesita escribir en un sistema de ficheros. A nivel básico, el SGBD proporciona todas las rutinas de manejo de ficheros típicas de los programas de aplicación.
- **Mejora en el mantenimiento:** En los sistemas de ficheros, las descripciones de los datos se encuentran inmersas en los programas de aplicación que los manejan.

- Esto hace que los programas sean dependientes de los datos, de modo que un cambio en su estructura, o un cambio en el modo en que se almacena en disco, requiere cambios importantes en los programas cuyos datos se ven afectados.
- **Aumento de la concurrencia:** En algunos sistemas de ficheros, si hay varios usuarios que pueden acceder simultáneamente a un mismo fichero, es posible que el acceso interfiera entre ellos de modo que se pierda información o se pierda la integridad. La mayoría de los SGBD gestionan el acceso concurrente a la base de datos y garantizan que no ocurran problemas de este tipo.
- **Mejora en los servicios de copias de seguridad:** Muchos sistemas de ficheros dejan que sea el usuario quien proporcione las medidas necesarias para proteger los datos ante fallos en el sistema o en las aplicaciones. Los usuarios tienen que hacer copias de seguridad cada día, y si se produce algún fallo, utilizar estas copias para restaurarlos.

En este caso, todo el trabajo realizado sobre los datos desde que se hizo la última copia de seguridad se pierde y se tiene que volver a realizar. Sin embargo, los SGBD actuales funcionan de modo que se minimiza la cantidad de trabajo perdido cuando se produce un fallo.

Desventajas de las bases de datos

- **Complejidad:** Los SGBD son conjuntos de programas que pueden llegar a ser complejos con una gran funcionalidad. Es preciso comprender muy bien esta funcionalidad para poder realizar un buen uso de ellos.
- **Coste del equipamiento adicional:** Tanto el SGBD, como la propia base de datos, pueden hacer que sea necesario adquirir más espacio de almacenamiento. Además, para alcanzar las prestaciones deseadas, es posible que sea necesario adquirir una máquina más grande o una máquina que se dedique solamente al SGBD. Todo esto hará que la implantación de un sistema de bases de datos sea más cara.
- **Vulnerable a los fallos:** El hecho de que todo esté centralizado en el

SGBD hace que el sistema sea más vulnerable ante los fallos que puedan producirse. Es por ello que deben tenerse copias de seguridad (Backup).

Manejadores de Archivo

Cuando abre un archivo al menos que especifique una aplicación en particular o un visualizador para ese archivo este normalmente se abre automáticamente en una aplicación que sea la apropiada para ese archivo. Por ejemplo, un archivo de gráficos normalmente se abrirá automáticamente en una aplicación que maneje gráficos.

La aplicación que está configurada para abrir automáticamente un tipo de archivo particular es conocida como el manejar de archivo para este tipo de archivo. Los manejadores de archivos también son llamados tipos MIME o tipos de archivo.

Usted puede elegir que aplicación o visualizador abrirá automáticamente un tipo de archivo en particular. También puede configurar un manejador de archivo nuevo.

Manejador de Bases de Datos

El sistema manejador de bases de datos es la porción más importante del software de un sistema de base de datos. Un DBMS es una colección de numerosas rutinas de software interrelacionadas, cada una de las cuales es responsable de alguna tarea específica.

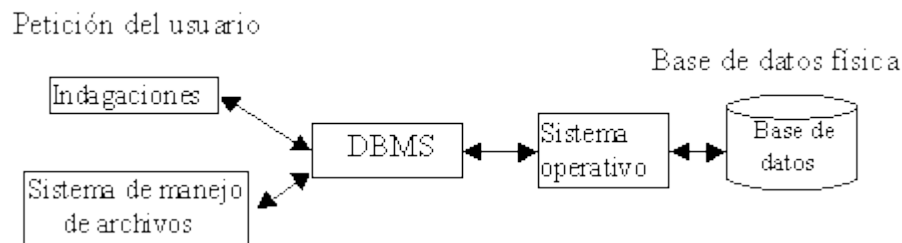
Las funciones principales de un DBMS son:

- **Crear y organizar la Base de datos.**
- **Establecer y mantener las trayectorias de acceso a la base de datos de tal forma que los datos puedan ser accedados rápidamente.**
- **Manejar los datos de acuerdo a las peticiones de los usuarios.**
- **Registrar el uso de las bases de datos.**
- **Interacción con el manejador de archivos.** Esto a través de las sentencias en DML al comando del sistema de archivos. Así el Manejador de base de datos es el responsable del verdadero almacenamiento de los

datos.

- **Respaldo y recuperación.** Consiste en contar con mecanismos implantados que permitan la recuperación fácilmente de los datos en caso de ocurrir fallas en el sistema de base de datos.
- **Control de concurrencia.** Consiste en controlar la interacción entre los usuarios concurrentes para no afectar la inconsistencia de los datos.
- **Seguridad e integridad.** Consiste en contar con mecanismos que permitan el control de la consistencia de los datos evitando que estos se vean perjudicados por cambios no autorizados o previstos.

El DBMS es conocido también como Gestor de Base de datos.



La figura muestra el DBMS como interface entre la base de datos física y las peticiones del usuario. El DBMS interpreta las peticiones de entrada/salida del usuario y las manda al sistema operativo para la transferencia de datos entre la unidad de memoria secundaria y la memoria principal.

En sí, un sistema manejador de base de datos es el corazón de la base de datos ya que se encarga del control total de los posibles aspectos que la puedan afectar.

Sistemas de bases de datos vs sistemas de archivos

El camino hacia las bases de datos ha sido largo y en el trayecto se han desarrollado un gran número de técnicas que forman los cimientos de las bd y de otras tecnologías.

Dentro de estas técnicas tenemos:

Archivos, Sistemas de Archivos, Acceso y manipulación de archivos, Índices

Pero...

Por qué no es suficiente utilizar las herramientas anteriores y es necesario

emplear un DBMS?

No es lo mismo?

Cuál es la diferencia?

No es lo mismo, un sistemas de archivos aún cuando pensemos que contiene lógicamente archivos y que se cuenta con índices para acceder los registros en ellos, carece de mucha funcionalidad que se emplea en la mayoría de las aplicaciones, aunque como se mencionó anteriormente, un DBMS emplea sistemas de archivos e índices para la manipulación de datos.

La funcionalidad adicional que provee un DBMS surge en base de algunos inconvenientes al emplear sistemas de archivos únicamente:

Redundancia de datos e inconsistencias (Redundancy and Inconsistency): formatos, duplicidad de información (alto costo de almacenamiento y acceso) e incongruencia entre datos o copias de datos a lo largo del sistema.

Dificultad de acceso (Access): en un sistema de archivos no se pueden obtener aquellos datos que no estén implantados en un programa, se carece de niveles de abstracción.

Aislamiento de datos (Isolation): debido al factor tiempo y los requerimientos que van surgiendo se puede llegar a tener un problema al intentar separar un conjunto de datos porque ya se tiene un enredo en los archivos y se podría dar el caso en que dos usuarios estén manipulando la misma información pero de distinta manera.

Integridad (Integrity): si queremos asociar dos datos, por ejemplo un alumno con una materia que esté cursando, debemos asegurarnos que ambas entidades existan, de lo contrario el alumno parecerá cursando un curso fantasma y viceversa. Para ello se emplean "restricciones de consistencia" (consistency constraints)

Atomicidad (Atomicity): el problema clásico de transacciones bancarias, u ocurre toda la transacción o no ocurre nada pero no puede quedarse a medias.

Acceso concurrente (Concurrent-access): garantizar un buen tiempo de respuesta, que todos los usuarios puedan acceder y/o modificar la información;

esto no es fácil porque también hay que considerar que aunque los datos son los mismos, las aplicaciones no necesariamente lo son.

Seguridad (Security): no toda la información debe estar disponible a todos los usuarios, algunos usuarios solo tendrán permisos de lectura, esto es relativamente sencillo de resolver aplicando "roles" pero el problema aumenta cuando en lugar de pensar en términos de usuarios pensamos en términos de aplicaciones ya que el número de roles y sus combinaciones aumenta y mantener las restricciones de seguridad se torna complicado.

Podemos entonces extender la definición de DBMS como un sistema robusto que es capaz de emplear algoritmos de almacenamiento y recuperación de información para poder implementar un modelo de datos de manera física garantizando que todas las transacciones que se realizan con respecto a dichos datos sean "ácidas" (Atomicity, Consistency, Isolation, Durability).