

Análisis de Sistemas

Materia:
Base de Datos I

Docente contenidista: MANSILLA, Hugo Orlando

Revisión: Coordinación



Contenido

Te Damos la Bienvenida	4
¿Qué es la información?	5
Tipos de información.....	6
Información vs. Datos (procesados).....	7
Clasificación de los datos	9
Clasificación de los datos según su posibilidad de acceso	10
Formas de almacenamiento de datos	12
Definición de Base de Datos.....	14
Tipos de Bases de Datos.....	15
Bibliografía	21
Para ampliar la información	21

Clase 1



iTe damos la bienvenida a la materia
Base de Datos I!

En esta clase vamos a ver los siguientes temas:

- Información.
- Información vs. Datos.
- Clasificación de datos.
- Formas de almacenamiento de datos.
- Definición de base de datos.
- Tipos de base de datos.

Te Damos la Bienvenida

Hoy comenzamos a recorrer el mundo de los datos: vivimos inmersos en la Era de la Información, en la que es crucial comprender cómo se componen y organizan los datos que disponemos de cada rompecabezas de información que es la realidad.

*La información es un compendio de datos organizados
que nos permite hacer distinciones entre distintos
escenarios de nuestra realidad.*

Durante nuestra materia comprenderemos la importancia de los datos y de su correcta organización para tener disponible la información que necesitemos de manera ordenada y concreta.

¿Qué es la información?

Empecemos por una definición formal de “información”:

*La información es un conjunto organizado **de datos procesados**, que constituye un mensaje que cambia del estado de conocimiento del **sujeto o sistema** que recibe dicho mensaje*

Algunas definiciones más de información:

Adalberto Chiavenato afirmaba que la información consiste en un conjunto de datos que poseen un significado, de modo tal que reducen la incertidumbre y aumentan el conocimiento de quien se acerca a contemplarlos.

Estos datos se encuentran disponibles para su uso inmediato y sirven para clarificar incertidumbres sobre determinados temas.

Ferrell y Hirt, por su parte, dicen que esos datos y conocimientos están estrictamente ligados con mejorar nuestra toma de decisiones. Si un individuo se encuentra bien informado sobre un aspecto, seguramente su decisión podrá ser más acertada que la de uno que no lo esté.

Otros autores que han definido la información son **Czinkota** y **Kotabe**, que quienes la definen como un conjunto de datos que han sido clasificados y ordenados con un propósito determinado.

Uniendo todas las teorías sobre el concepto, llegamos a la conclusión que la información son datos sobre un suceso o fenómeno particular que, al ser ordenados en un contexto, sirven para disminuir la incertidumbre y aumentar el conocimiento sobre un tema específico.

Tipos de información

Existen muchos tipos de información. Algunos de ellos son:

- **Información meteorológica**

Reúne datos sobre temperatura y precipitaciones, y otros que hacen referencia al clima de una zona. Sirve para predecir el tiempo en un lugar. Quienes se encargan de ordenar esta información se llaman meteorólogos.

- **Información financiera**

Conjunto de datos sobre una economía de mercado que permiten establecer parámetros de tipo económicos, así como analizar la solidez económica de un país y predecir las operaciones que convendría realizar.

- También puede conseguirse información a través de **material de estudio sobre temas específicos**, como documentales, libros de historia, sociología, matemáticas, etc., que permiten resolver dudas y ofrecen definiciones claras sobre los conceptos que se desea estudiar.

Información vs. Datos (procesados)



Como se desprende de la definición de información, podemos notar que se habla de “datos procesados”.
Ahora bien, ¿qué diferencia hay entre dato e información?

Empecemos por definir qué se entiende por dato.

Un **dato** es una representación simbólica de un atributo o característica de una entidad o de un hecho de la realidad.

En la realidad podemos distinguir entre entidades que pueden ser tangibles o intangibles. Una **entidad** es cualquier elemento de la realidad que podamos describir a través de sus características. Por ejemplo, una persona que puedo describir por su nombre, estado civil, edad, etc.

Un **hecho** es un suceso que puedo describir a través de lo ocurrido. Por ejemplo, en el suceso donde Argentina salió campeón del Mundial de Qatar: podemos inferir en qué día fue la final, contra qué rival se jugó, cuál fue el resultado, las alineaciones de ambos equipos, etc.



Como verán, podemos representar cualquier entidad o hecho a través de sus datos.

Pero ¿podemos decir que cada entidad o hecho es información en sí?

*Un dato en sí mismo no constituye información.
Es el procesamiento de los datos lo que
nos proporciona información*

De esta distinción se desprende que, si no los procesamos, los datos pueden quedar aislados y no constituir información en sí.

Por ejemplo, si describo un auto sin indicar todos los componentes, solo serán datos aislados de un vehículo que no serán información del mismo, ya que tendré solo datos parciales.

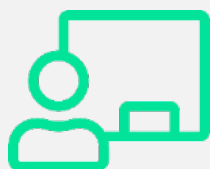
Lo mismo si digo algún dato aislado sobre un hecho de la realidad. Por ejemplo, si digo que hay 90% de humedad, no implica que esté lloviendo: tendría que agregar información acerca de la cantidad de precipitaciones y poder medirlas. Es decir, la humedad es solo un dato aislado y no constituye por sí sola la información de que esté lloviendo.

Ahora que ya sabemos lo que son los datos, podemos volver a la definición que nos dice que la información es el conjunto de **datos procesados**.

Esta afirmación implica que debemos clasificar y organizar los datos para que tengan sentido y, en conjunto, se transformen en información.

Dicha clasificación y organización va a surgir de nuestro análisis, donde evaluaremos:

- qué tipos de datos están involucrados (números, textos, fechas, etc.),
- en qué orden tenemos que organizarlos y
- cómo hacer para que la información sea concreta y fácil de acceder.



Trabajaremos todo esto a través del modelado de datos que veremos durante toda la materia, y que es la columna vertebral de los sistemas de gestión de información.



Clasificación de los datos

Podemos clasificar los datos según su representación de la siguiente manera:

- **Datos numéricos**
Pueden ser números enteros o reales. Por ejemplo, la edad de una persona, el precio de un producto, la cantidad de goles de un partido de fútbol, etc.
- **Datos alfanuméricos**
Constituidos por letras y números, pueden ser una palabra o un conjunto de palabras. Por ejemplo, el nombre y apellido de una persona, el nombre del árbitro de la final del mundo, la descripción de un producto, etc.
- **Datos de fechas**
Cualquier fecha del calendario. Por ejemplo, fecha de nacimiento, fecha de vencimiento de un pago, fecha de un partido de fútbol, etc.
- **Datos lógicos**
Cualquier dato que solo sea verdadero o falso. Por ejemplo, si quiero saber si un alumno aprobó o no su examen, podría tener un dato que sea "Aprobó", y que allí solo informe si es verdadero o falso.
- **Datos complejos**
Existen otras clasificaciones de datos como pueden ser los archivos adjuntos de una publicación, fotografías, videos, etc.

¡No dejemos de practicar!



Te invitamos a que clasifiques los datos de lo que más te gusta. Puede ser un objeto o una actividad. ¿Te animás?

Compartilo con tus compañeros en el foro para intercambiar experiencias.

Clasificación de los datos según su posibilidad de acceso

La información suele clasificarse acorde a cuatro criterios de acceso: **confidencial, restringido, uso interno y público.**

Estos criterios obedecen a una categorización generalizada y previamente establecida por las **Normas ISO**. De aquí que la misma cobre una importancia vital en las empresas como también en los espacios institucionales.

Como es bien sabido, la información no es más que un conjunto de datos que revisten una importancia vital para las personas, considerando que los mismos representen un contenido verídico y, sobre todo, útil.

Adecuarse a un sistema de clasificación de la información es uno de los pasos principales que toda empresa lleva a cabo con el fin de mantener sus datos en perfecto orden y sincronización.

La clasificación de la información se presenta como una regla indescriptible a seguir en toda organización con el fin de que pueda ser adecuada de forma correcta y, así, se resguarde todo tipo de dato que represente utilidad para la misma.

Información Confidencial

Es aquel cúmulo de datos que presentan un contenido generalizado pero que solo puede estar a la vista de los altos ejecutivos; por lo general, son catalogados de confidencial todos y cada uno de los datos que se disponen en el departamento de Recursos Humanos, ya que contempla todas las nociones sobre los empleados.

Esta información se marca como confidencial con el fin de salvaguardar los intereses y la privacidad de cada miembro del capital humano.

Información de Uso interno

Es la información que se dispone en cada uno de los departamentos y que es del conocimiento del personal que labora en estos.

En efecto, esta información está compuesta por datos que solo pueden ser entendidos y operados por el personal que la maneja y crea con sus actividades. De allí que suela ser exclusiva de cada área, no pudiendo ser compartida con los demás departamentos, a menos que los ejecutivos determinen lo contrario.

Información para Todo público

Una vez que se realiza la depuración de los datos y se los clasifica por su importancia y nivel necesario de protección para resguardarlos de daños de terceros, quedan datos que pueden ser expuestos a todo el personal.

Es más: como parte de toda organización, siempre existe un conjunto de datos que debe ser explicado al público para que pueda darse a conocer como empresa.

Formas de almacenamiento de datos

Un poco de historia

Más allá de los datos, lo importante es que podamos almacenarlos para consultarlos en cualquier momento.

Históricamente, antes del advenimiento de la computadora, para acceder a los datos teníamos que recurrir a bibliotecas y libros que contuvieran la información ya organizada.

También podíamos acceder a periódicos en papel a partir de grandes archivos donde se encontraban clasificados y ordenados por fecha. También tuvimos la posibilidad, ya más adelante, de poder ver videocasetes, donde quizá podíamos ver a través de documentales la información que necesitábamos, y esos datos estaban allí almacenados.



La computadora revolucionó todo. Los datos podían almacenarse ya en dispositivos físicos como disquetes, inicialmente, y luego CD y DVD, hasta llegar a los pendrives y los discos externos.

También las computadoras comenzaron a traer discos rígidos cada vez con más capacidad y velocidad.

Ya con la llegada de internet y el surgimiento de servicios en la nube, hoy día podemos almacenar nuestros datos en servidores que quizá estén del otro lado del planeta.

Ahora bien, lo que acabamos de describir es dónde podemos almacenar los datos (cosa no menor), pero, para poder procesarlos, **¿cómo los almacenamos?**

Formas de almacenamiento:

- Archivos de texto plano.
- Documentos de Word.
- Planillas de Excel.
- **Base de Datos.**

Y acá llegamos a lo que estábamos esperando.

Todo comenzó con el procesamiento de archivos que soluciona gran parte de los problemas, pero no todos: cuando tenemos datos complejos y mucho volumen de información, el sistema de archivos queda obsoleto y se necesita contener la información en una fuente única de datos que permita organizarlos de manera coherente, concisa y segura.

Allí entran las **Bases de Datos.**

Definición de Base de Datos

Hagamos un compendio de distintas definiciones de lo que es una base de datos:

- Una base de datos es un conjunto de información que pertenece a un mismo contexto, ordenada de modo sistemático, para su posterior recuperación, análisis y/o transmisión.
- Una base de datos es una herramienta para recopilar y organizar información. Las bases de datos pueden almacenar información sobre personas, productos, pedidos u otras cosas. Muchas bases de datos comienzan como una lista en una hoja de cálculo o en un programa de procesamiento de texto. A medida que la lista aumenta su tamaño, empiezan a aparecer redundancias e inconsistencias en los datos. Cada vez es más difícil comprender los datos en forma de lista, y los métodos de búsqueda y extracción son limitados. Una vez que aparecen estos problemas, se recomienda transferir los datos a una base de datos.
- Una base de datos es una recopilación organizada de información o datos estructurados que normalmente se almacena en forma electrónica en un sistema informático.
- Una base de datos no solo se encarga de almacenar datos, sino también de conectarlos entre sí en una unidad lógica. Es un conjunto de datos estructurados que pertenecen a un mismo contexto y que se utiliza para administrar de forma electrónica grandes cantidades de información.
- De este compendio de definiciones podemos ver un común denominador en que todas hablan de recopilación organizada de datos y de un mismo contexto. Por ejemplo, puede tratarse de la base de datos de una biblioteca o de una escuela. También se destaca que se puede mejorar la calidad de la información al estar ordenada, y que permite grandes volúmenes de información a los que se puede acceder y extraer de manera rápida.

Tipos de Bases de Datos

Existen diferentes tipos de base de datos según capacidad de almacenamiento, robustez y costos. Conozcamos los principales hoy día:

- **Relacional:**

Almacena y proporciona acceso a datos relacionados entre sí. Se basa en el modelo relacional, una forma intuitiva y directa de representar los datos en tablas. En una base de datos relacional, cada fila en una tabla es un registro único. Las columnas de la tabla contienen los atributos o características de los datos y cada registro suele tener un valor para cada atributo, lo que simplifica la creación de relaciones entre los datos.

El lenguaje de consulta estructurado (SQL) es la interfaz standard de usuario de programa de aplicación para una base de datos relacional.

Las más conocidas:

Oracle: con esta base de datos se puede navegar en cualquier sistema operativo. Tiene una gran variedad de perfiles con buena experiencia.



Teradata: sobresale de todas las demás gracias a su capacidad para almacenar y analizar datos. Siempre es utilizada en grandes instalaciones.



IBM DB2: suele ser una de las más utilizadas por Unix/Linux por su gran rapidez y capacidad para almacenar cualquier tipo de información.



SQL Server: solo tiene compatibilidad con sistemas Windows.



MySQL: será la base de datos que utilizaremos en la materia. Es un sistema de gestión de bases de datos relacionales de código abierto con un modelo cliente-servidor.



NoSQL (no relacional): no utiliza el esquema tabular de filas y columnas que se encuentra en la mayoría de los sistemas de bases de datos tradicionales. Son eficaces para los problemas de rendimiento Big Data (grandes conjuntos de datos distribuidos). Además, son más efectivas cuando una organización debe analizar grandes porciones de datos no estructurados, o datos que se almacenan en múltiples servidores virtuales en la nube.

Las más conocidas:

Mongo DB: es un gestor de base de datos de código abierto y orientado a documentos. Mongo DB Inc. ofrece una suite integrada de servicios cloud de base de datos, así como soporte comercial. Se suele utilizar para almacenar grandes volúmenes de datos.



Redis: abreviatura de Remote Dictionary Server. Es un sistema de gestión de bases de datos de código abierto NoSQL de clave-valor.



- **Orientada a objetos:**

Es un tipo de base de datos adecuado para elementos que se crean con lenguajes de programación orientados a objetos, como C++ o Java.

Se organiza en torno a objetos en lugar de acciones, y a datos en lugar de lógica.

Por ejemplo, un registro multimedia en una base de datos relacional puede ser un objeto de datos definible, a diferencia de un valor alfanumérico.

Una de las más utilizadas:

ObjectDB: es un sistema de gestión de base de datos orientado a objetos, multiplataforma. Requiere usar una API standard de Java, JPA (Jakarta Persistence) o JDO (Java Data Object). Está pensado para proporcionar rendimiento y aplicaciones más rápidos. Además, se presenta como la opción más productiva para desarrollar aplicaciones de Base de Datos Java usando la API Java Persistence.



- **Gráfica:**

Base de datos orientada a gráficos (o base datos gráfica).

Es un tipo de base de datos NoSQL que utiliza la teoría de los grafos para almacenar, mapear y consultar relaciones.

Son colecciones de nodos y bordes, donde cada nodo representa una entidad y cada borde representa una conexión entre nodos.

Actualmente están siendo muy utilizadas por las empresas para analizar interconexiones y, por ejemplo, para obtener datos sobre clientes en las redes sociales.

Una de las más conocidas:

InfiniteGraph: gestor de base de datos orientado a grafos, multiplataforma, escalable y habilitado para la nube. Está pensado para soportar un rendimiento alto y optimizado para grandes conjuntos de datos, altamente conectados y complejos.



- **En la nube:**

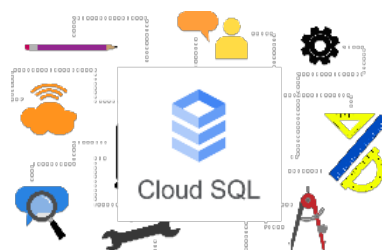
Se trata de una base de datos que se ha optimizado o creado para un entorno virtualizado, ya sea en una nube híbrida, pública o privada.

Tiene la ventaja de pagar por capacidad de almacenamiento y el ancho de banda para cada usuario. Además, ofrece escalabilidad y alta disponibilidad.

Además, brinda a las empresas la oportunidad de admitir aplicaciones de **software como servicio**.

Las más conocidas:

Google Cloud SQL: el servicio de base de datos en la nube de Google se centra en dos productos principales: GoogleCloud SQL, que Google describe como una infraestructura de base de datos MySQL completamente relacional, y Google BigQuery, herramienta de análisis para ejecutar consultas en grandes conjuntos de datos almacenados en la nube.



Microsoft Azure: Microsoft utiliza su tecnología de servidor SQL para proporcionar una base de datos relacional, permitiendo que los clientes accedan a una base de datos SQL, ya sea en su nube o en instancias de servidor SQL en máquinas virtuales. Microsoft también hace hincapié de las bases de datos híbridas que combinan datos tanto en las instalaciones del cliente como en la nube de Azure a través de SQL Data Sync. Microsoft tiene un servicio de nube alojada en la base de datos NoSQL llamada Tables, mientras que BLOB (almacenamiento de objeto binario grande) se ha optimizado para archivos multimedia, como audio y video.



Otra **clasificación importante de las bases de datos** está relacionada con dónde se encuentran sus componentes principales y los datos, e incluye las bases de datos centralizadas y las distribuidas, cuyas características son las siguientes:

- **Centralizada**

La información (datos) se almacena en una ubicación centralizada y los usuarios de diferentes ubicaciones pueden acceder a estos datos.

En este tipo de base de datos se aplican diversos procedimientos de autenticación para la verificación y validación de los usuarios finales.

Asimismo, los procedimientos de la aplicación proporcionan un número de registro que mantiene un registro de uso de los datos.

- **Distribuida**

Es el tipo de bases de datos más utilizado actualmente. Las diferentes partes de la base de datos se almacenan en varias ubicaciones físicas, y el procesamiento se dispersa o replica en diferentes puntos de la red.

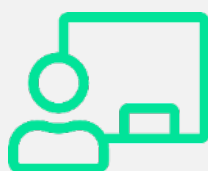
Hay dos tipos de bases de datos distribuidas: homogéneas y heterogéneas. En las homogéneas, todas las ubicaciones físicas tienen el mismo hardware subyacente y ejecutan los mismos sistemas operativos y aplicaciones de base de datos.

En las heterogéneas, el hardware, los sistemas operativos o las aplicaciones pueden ser diferentes en cada una de las situaciones.



Hemos llegado así al final de esta clase en la que vimos:

- Concepto de información.
- Diferencias entre información y dato.
- Clasificación de datos.
- Formas de almacenamiento de datos.
- Definición de base de datos.
- Tipos de bases de datos.



Te esperamos en la **clase en vivo** de esta semana.
No olvides realizar el **desafío semanal**.

¡Hasta la próxima clase!

Bibliografía

Korth, H., Silberschatz, A., y Sudarshan, S., (2002). Fundamentos de bases de datos. Editorial Mc Graw Hill.

Para ampliar la información

Concepto de Base de Datos:

<https://concepto.de/base-de-datos/>

Conceptos básicos sobre base de datos:

<https://support.microsoft.com/es-es/office/conceptos-b%C3%A1sicos-sobre-bases-de-datos-a849ac16-07c7-4a31-9948-3c8c94a7c204>

¿Qué es una base de datos?

<https://www.oracle.com/mx/database/what-is-database/>

¿Qué es una base de datos relacional?

<https://www.oracle.com/mx/database/what-is-a-relational-database/>