

Conceptos Básicos: Dato, Información y Base de datos



Contenido

1

Jerarquía DIKW

Concepto

2

Base de Datos

¿Qué es?
¿Es digital?

3

DBMS / Motor de Base de Datos

¿Qué es?
Utilidad
Características

4

Paradigma transaccional

ACID

5

Arquitectura técnica del DBMS

Diagramas

6

Modelo relacional

¿Qué es?
Componentes

Jerarquía
DIKW

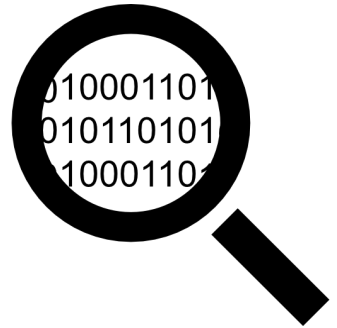
Conceptos básicos: Dato e Información



Fuente: Hey, J.: The Data, Information, Knowledge, Wisdom Chaim: The Metaphorical Link

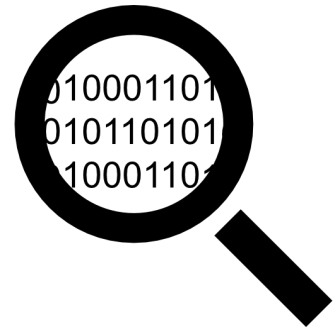
Jerarquía del Conocimiento (DIKW)

La "**Jerarquía del Conocimiento**" (también conocida como "Jerarquía DIKW" o "Pirámide del Conocimiento") podría ser definida como un conjunto de modelos para representar las relaciones aparentemente estructurales entre Datos, Información, Conocimiento, y en algunos casos Sabiduría.



Conceptos básicos: Dato e Información

¿Qué es un dato?



Conceptos básicos: Dato

El dato se refiere a la representación simbólica de una entidad, por ejemplo letras del alfabeto, números, puntos, dibujos, etc.

Estos datos por sí solos no tienen valor semántico,

Pero si se le procesa apropiadamente, este provee información importante ayudando en la toma de decisiones.



Conceptos básicos: Dato

- ✓ Representación Simbólica
- ✓ No tienen sentido semántico
- ✓ No transmiten Mensaje
- ✓ Describen situaciones, hechos



Conceptos básicos: Dato e Información

¿Que es información?



Conceptos básicos: Información

La información se refiere al conjunto de datos, que están organizados para transmitir un significado, con el propósito de reducir la incertidumbre e incrementar el **conocimiento**.

La Información favorece a la resolución de problemas puesto que permite una adecuada toma de decisiones.



Conceptos básicos: Información

- ✓ Conjunto de datos procesados y organizados
- ✓ Tiene significado, transmite un mensaje
- ✓ Incrementa el conocimiento
- ✓ Facilita la toma de decisiones



Bases de Datos

Base de datos

- Una Base de Datos es un **conjunto de datos relacionados y organizados** de tal manera que puedan responder a un propósito específico. Representan aspectos de la realidad (Dato != Información).
- Permiten **organizar los datos** con métodos eficientes para obtener la información requerida y **crear reglas para asegurarse de que los datos se almacenen de forma correcta y consistente**.
- El lenguaje que permite obtener datos de una Base de Datos se llama **SQL** (Structured Query Language).

Base de datos

**¿Toda base de datos es
digital?**

Base de datos



Agenda Tradicional

Una agenda de papel puede ser un ejemplo de base de datos no digital, contiene contactos y sus teléfonos de forma organizada para facilitar su consulta.

Base de datos

**¿Por qué usamos base de datos
en lugar de archivos de texto o
planillas de cálculo?**

Base de datos

Archivos Tradicionales (excel, csv, txt)

- ¿Concurrencia?
- ¿Consistencia?
- ¿Disponibilidad?
- ¿Integridad de los datos?

DBMS / Motor de Base de Datos

DBMS / Motor de Base de Datos

- Con el término “Base de Datos” también se hace referencia al software (motor de base de datos o **DBMS**) que permite administrarla, es decir, gestionar su tamaño, cargarle datos y consultarlos.
- Permiten **organizar** los datos con métodos eficientes para **obtener la información requerida** y crear reglas para asegurarse de que **los datos se almacenen de forma correcta y consistente**.
- El lenguaje que permite obtener datos de una Base de Datos administrada por un Motor de Base de datos relacional (RDBMS) se llama **SQL** (Structured Query Language).

DBMS / Motor de Base de Datos

DBMS (Database Management system) es una colección de **software muy específico**, cuya función es servir de **interfaz** entre la base de datos, el usuario y las distintas aplicaciones utilizadas.

El objetivo es manejar un conjunto de datos para convertirlos en **información relevante** para la organización, ya sea a nivel operativo o estratégico.

DBMS / Motor de Base de Datos

Características:

- ✓ Es un componente de Software
- ✓ Garantiza disponibilidad y accesibilidad
- ✓ Permite el acceso concurrente a los datos
- ✓ Asegura la integridad transaccional
- ✓ Protege los datos y los accesos
- ✓ Optimiza la performance y el rendimiento.

DBMS / Motor de Base de Datos

DEBE CUMPLIR CON EL PARADIGMA TRANSACCIONAL (ACID)

- Concurrency
- Consistency
- Availability
- Integrity of the data

Paradigma Transaccional

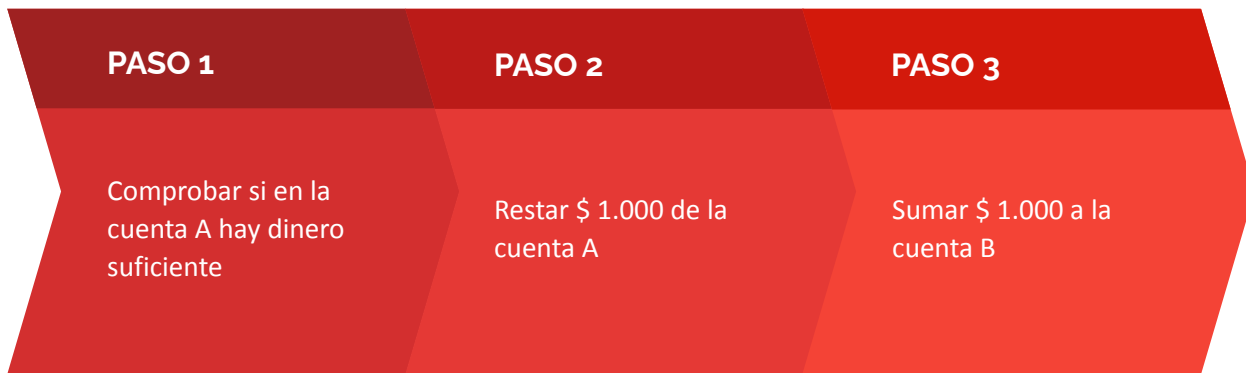
INTEGRIDAD O PARADIGMA TRANSACCIONAL (ACID)

- Atómicas (**A**tomicity): Todas las modificaciones de una transacción deben ser ejecutadas o bien, ninguna de ella. Se ejecuta “todo o nada”.
- Consistentes (**C**onsistency): La transacción comienza a ejecutarse en un estado válido (consistente) y finaliza en otro estado válido. Para eso, deben cumplirse todas las reglas de integridad que hayan sido definidas.
- Aisladas (**I**solation): Dos o más transacciones *concurrentes* se ejecutan sin afectarse entre sí. Si ambas transacciones deben trabajar sobre el mismo dato una de ellas deberá esperar a que la otra termine.
- Durables (**D**urability): Asegura que una vez realizada la operación va a persistir (quedarán almacenados los cambios) por más que falle el sistema. Uso del Transaction Log del DBMS.

Paradigma Transaccional

- Una **transacción** es un conjunto de operaciones, pertenecientes a una misma tarea, que se realizan sobre una base de datos y representan un cambio en los datos.

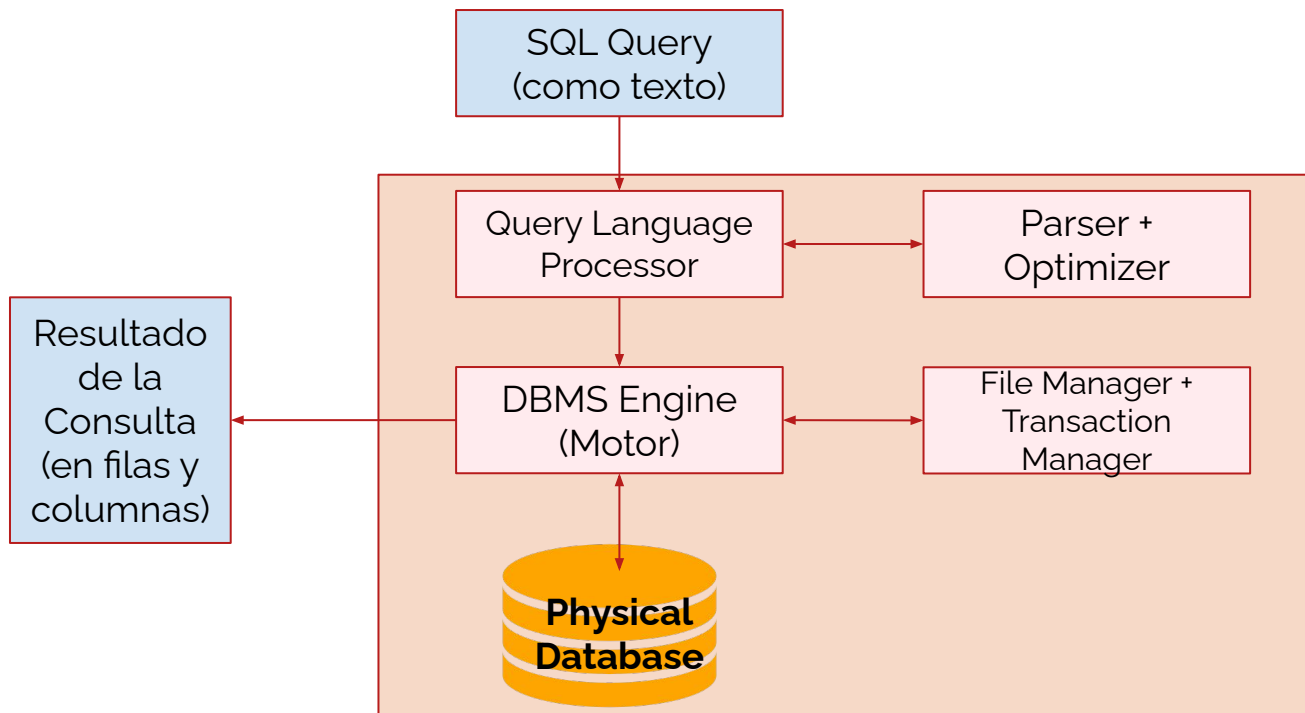
Ejemplo de una Transacción bancaria



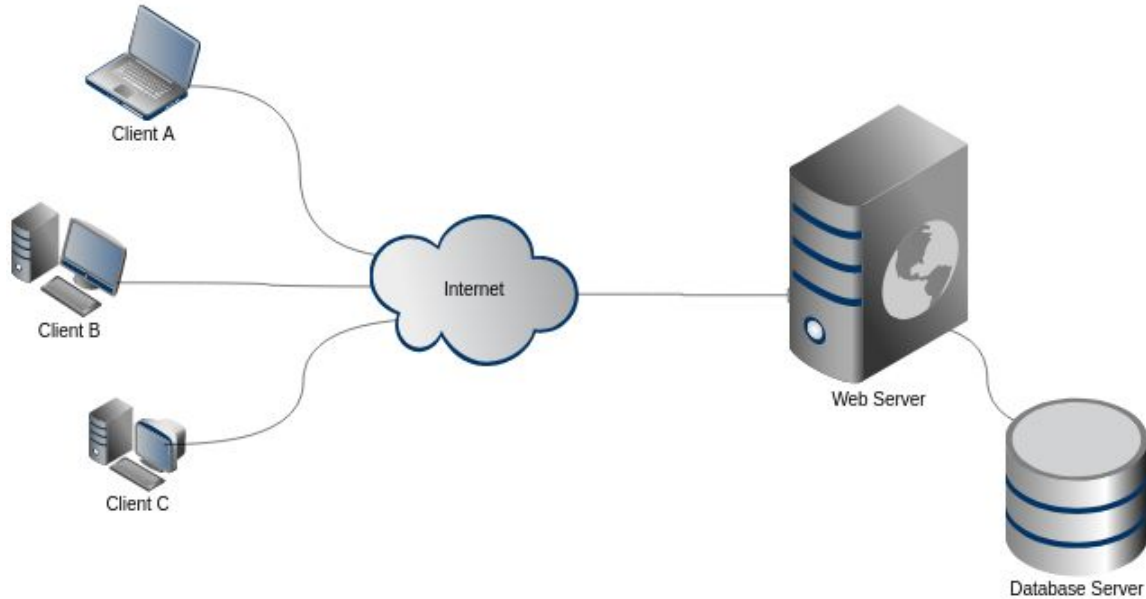
- ¿Qué pasa si el proceso falla entre el paso 2 y el 3 (o sólo se ejecutan los pasos 1 y 2)?
- ¿Qué obtiene alguien que quiere consultar su Saldo entre los pasos mencionados?

Arquitectura Técnica del DBMS

DBMS - ESQUEMA DE FUNCIONAMIENTO DEL MOTOR



DBMS - ESQUEMA ACCESO AL SERVICIO DEL MOTOR



DBMS Mercado Actual

DBMS: MERCADO ACTUAL

Ranking Motores -LINK: <https://db-engines.com/en/ranking>

359 systems in ranking, August 2020

Rank			DBMS	Database Model	Score		
Aug 2020	Jul 2020	Aug 2019			Aug 2020	Jul 2020	Aug 2019
1.	1.	1.	Oracle +	Relational, Multi-model i	1355.16	+14.90	+15.68
2.	2.	2.	MySQL +	Relational, Multi-model i	1261.57	-6.93	+7.89
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model i	1075.87	+16.15	-17.30
4.	4.	4.	PostgreSQL +	Relational, Multi-model i	536.77	+9.76	+55.43
5.	5.	5.	MongoDB +	Document, Multi-model i	443.56	+0.08	+38.99
6.	6.	6.	IBM Db2 +	Relational, Multi-model i	162.45	-0.72	-10.50
7.	↑ 8.	↑ 8.	Redis +	Key-value, Multi-model i	152.87	+2.83	+8.79
8.	↓ 7.	↓ 7.	Elasticsearch +	Search engine, Multi-model i	152.32	+0.73	+3.23
9.	9.	↑ 11.	SQLite +	Relational	126.82	-0.64	+4.10
10.	↑ 11.	↓ 9.	Microsoft Access	Relational	119.86	+3.32	-15.47

DBMS: MERCADO ACTUAL



- Desarrollado de forma independiente y luego adquirido por Oracle
- Multiplataforma
- Open Source y licencia de tipo "Copy Left"
- Usado por grandes empresas como Facebook, Twitter, Wikipedia, etc.

DBMS: MERCADO ACTUAL



Microsoft®
SQL Server®

- Desarrollado por la empresa Microsoft
- No es multiplataforma, solo corre en sistemas operativos Windows
- No es open source y para usarlo se debe pagar una licencia.
- Muy usado en entornos corporativos donde el sistema operativo de base es Windows.

DBMS: MERCADO ACTUAL



- Desarrollado por la universidad de Berkeley
- Multiplataforma
- Open source y con licencia libre incluso para usos comerciales
- Se caracteriza por su alta performance y su versatilidad en los tipos de datos que permite almacenar.

GARTNER MAGIC QUADRANT FOR OPERATIONAL DATABASE MANAGEMENT SYSTEMS

Figure 1. Magic Quadrant for Operational Database Management Systems



Source: Gartner (November 2019)

GARTNER MAGIC QUADRANT FOR DATA MANAGEMENT SOLUTIONS FOR ANALYTICS

Figure 1. Magic Quadrant for Data Management Solutions for Analytics



Source: Gartner (January 2019)

Modelo Relacional

Modelo Relacional

El modelo relacional se utiliza para el modelado y la gestión de bases de datos. Su idea fundamental es el uso de relaciones. Estas relaciones podrían considerarse en forma lógica como conjuntos de datos llamados tuplas.

El objetivo es pensar en cada relación como si fuese una tabla que está compuesta por registros (cada fila de la tabla sería un registro o "tupla") y columnas (también llamadas "campos").

¿QUÉ ES UNA BASE DE DATOS RELACIONAL?

Son bases de datos basadas en el “modelo relacional” (Codd, 1970).

El modelo relacional se basa en los conceptos de álgebra relacional, de ahí su nombre.

- La información se organiza en **tablas** compuestas de filas y columnas que están relacionadas de acuerdo a las entidades o conceptos que representan.
- A las tablas se le asocian reglas que son aplicadas cuando se intenta cargarles algún dato:
 - Cuantos campos tiene y cómo se llaman
 - De qué tipo son (textos, números enteros, fechas, etc.)
- Al menos una columna debe funcionar como **Clave Primaria**.
- Puede contener una **Clave Foránea** que haga referencia a la Clave Primaria de otra tabla.

Modelo Relacional

Modelo Relacional - Componentes

1. Tablas
2. Filas (o Tuplas o Registros)
3. Columnas (o Campos)

Modelo Relacional

Tabla

- Tiene un nombre único que permite identificarla
- Estructura principal que almacena datos
- Abstracción de una **entidad**.
- Contiene registros/tuplas

Modelo Relacional

Fila / Registro / Tupla

- Se almacenan “dentro” de las tablas
- Todos los registros contienen datos de los mismos tipos y en el mismo orden.
- Está compuesto por una o más columnas
- Cada registro representa UN elemento puntual.

Modelo Relacional

Columna / Campo

- Cada registro de una tabla contendrá múltiples columnas
- Cada columna tendrá un tipo de dato definido
- Las columnas almacenan datos puntuales
- Cada columna tendrá un nombre que la identificará unívocamente dentro de la tabla

Modelo Relacional

Ejemplo de Aplicación

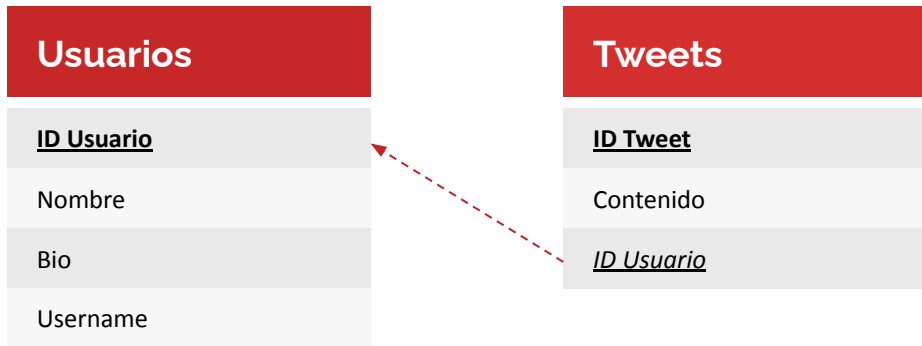
Twitter



Modelo Relacional

EJEMPLO DE UNA BASE DE DATOS RELACIONAL: TWITTER

- Las principales **entidades** o conceptos: Usuarios y Tweets. Cada una representada por una **tabla**.
- La tabla Usuarios tendrá el ID del usuario (Clave Primaria), su nombre y otros datos personales.
- La tabla Tweets tendrá el ID del Tweet (Clave Primaria), el texto o contenido del Tweet y el ID del usuario que lo escribió (Clave Foránea de la tabla de Usuarios).



Conclusión

CONCLUSIONES

- Las bases de datos relacionales nos permiten representar objetos de negocio, y cómo se relacionan dichos objetos.
- Esta modelización simplificada de la realidad, nos permite almacenar datos, relacionarlos, y organizarlos, de manera que permita convertir esos datos en información.
- Las bases de datos, administradas por Motores de Bases de datos o DBMS, son una herramienta tecnológica clave en el mundo de los datos, y también en el mundo de las aplicaciones. Sin esta herramienta no podríamos almacenar, ordenar y acceder a los datos de manera ágil, rápida y segura.
- Hoy en día, contamos con muchas opciones de DBMS, tanto para soluciones complejas o para comenzar a trabajar con ellas.

