

Concepto de Bases de datos:

Una base de datos es un contenedor que permite almacenar grandes volúmenes de información, el contenedor la mantiene bien organizada para poder realizar las operaciones de gestión sobre los datos.

- Almacenar.
- Eliminar
- Buscar
- Consultar
- Operaciones propias

Datos:

Un dato es una representación simbólica (entero, alfabética, algorítmica, espacial, visual, etc.) de un atributo o variable cuantitativa o cualitativa. Los datos describen hechos empíricos, sucesos y entidades.

Es un valor o referente, los datos representan la información que el desarrollador manipula en la construcción de una solución o en el desarrollo de un algoritmo. Como por ejemplo edad.

Información:

La información está constituida por un grupo de datos ya supervisados y ordenados, que sirven para construir un mensaje basado en un cierto fenómeno o ente. La información permite resolver problemas y tomar decisiones, ya que su aprovechamiento racional es la base del conocimiento. Como por ejemplo empleado.

Sistema de gestión de bases de datos:

Es un sistema de software que permite el almacenar, modificar, eliminar, consultar y extracción información en una base de datos, además proporciona herramientas para gestión de datos, junto con el análisis de los datos.

Motores de bases de datos.

Motor de base de datos es el servicio principal para almacenar, procesar y proteger los datos. Motor de base de datos es usado para crear bases de datos relacionales para el procesamiento de transacciones en línea o datos de procesamiento analíticos en línea.

- DB2: Este SGBD es propiedad de IBM, bajo la cual se comercializa el sistema de gestión de base de datos. Utiliza XML como motor, además el modelo que utiliza es el jerárquico en lugar del modelo relacional que utilizan otros gestores de bases de datos. Es el único de los gestores que hemos comentado que no es relacional.
- MongoDB: Es una base de datos Open Source de gran rendimiento, escalable, schema-free y orientada a documentos (esquemas de datos tipo JSON). Hay drivers preparados para usar esta base de datos desde lenguajes como PHP, Python, Perl, Ruby, JavaScript, C++ y muchos más.
- PostgreSQL: Es un sistema de gestión de base de datos relacional orientada a objetos y libre, publicado bajo la licencia BSD, como muchos otros proyectos de código abierto, el desarrollo de PostgreSQL no es manejado por una empresa y/o persona, sino que es dirigido por una comunidad de desarrolladores que trabajan de forma desinteresada, altruista, libre y/o apoyada por organizaciones comerciales. La comunidad PostgreSQL se denominada el PGDG (PostgreSQL Global Development Group).
- Hypertable : Es un sistema de almacenamiento distribuido de datos de alto rendimiento diseñado para soportar aplicaciones que requieran máximo rendimiento, escalabilidad y eficiencia. Se ha diseñado y modelado a partir del proyecto BigTable de Google y se enfoca sobre todo a conjuntos de datos de gran escala.
- Apache CouchDB: Como en el caso de MongoDB, este proyecto está destinado a ofrecer una base de datos orientada a documentos que se pueden consultar o indexar en modo MapReduce usando JavaScript. CouchDB ofrece una API JSON RESTful a la que se puede acceder desde cualquier entorno que soporte peticiones HTTP.
- Neo4j: Es un motor de persistencia completamente transaccional en Java que almacena los datos mediante grafos, y no mediante tablas. Neo4j ofrece una escalabilidad masiva. Puede manejar grafos de varios miles de millones de nodos/relaciones/propiedades en una única máquina, y se puede escalar a lo largo de múltiples máquinas.

- Riak: Riak es una base de datos ideal para aplicaciones web y combina: Una tienda con un valor clave descentralizado, un motor map/reduce flexible, una interfaz de consultas HTTP/JSPN amigable.
- Oracle Berkeley DB: Se trata de un motor de bases de datos embebidas que proporciona a los desarrolladores persistencia local, rápida y eficiente con una administración nula. Oracle Berkeley DB es una librería que se enlaza directamente en nuestras aplicaciones y permite realizar llamadas simples a funciones en lugar de enviar mensajes a un servidor remoto para mejorar el rendimiento.
- Apache Cassandra: Es probablemente uno de los proyectos NoSQL más concoidos del mercado. Se trata de una base de datos distribuida de segunda generación con alta escalabilidad que está siendo usada por gigantes como Facebook (que es quien la ha desarrollado), Digg, Twitter, Cisco y más empresas. El objetivo es ofrecer un entorno consistente, tolerante a fallos y de alta disponibilidad a la hora de almacenar datos.
- Memcached: Es un almacén del tipo in-memory key-value para pequeñas cadenas de datos arbitrarios (textos, objetos) de resultados de llamadas a base de datos, llamadas a API, o renderizado de páginas. Está orientado a acelerar aplicaciones web dinámicas al aliviar la carga de la base de datos.
- Firebird: Es una base de datos relacional que puede ser utilizada en Linux, Windows y varias plataformas UNIX, y ofrece un alto rendimiento y potente soporte de lenguajes para procedimientos almacenados y triggers.
- Redis: Es una base de datos avanzada del tipo fast key-value que está escrita en C y se puede usar como memcached, por delante de una base de datos tradicional, o bien por sí sola de forma independiente. Tiene soporte para varios lenguajes de programación y se utiliza en proyectos muy populares como GitHub o Engine Yard. También hay un cliente PHP llamado **Rediska** que permite gestionar bases de datos redis.
- HBase: Es u almacén distribuido del tipo column-oriented que puede ser también denominado como la base de datos Hadoop. El proyecto está dirigido a ofrecer tablas enormes de “miles de millones de filas, y millones de columnas”. Dispone de un gateway RESTful que soporta XML, Protobug y opciones de codificación binaria de datos.
- Keyspace: Se trata de un almacén del tipo key-value con replicación consistente y que funciona sobre sistemas operativos Windows. Keyspace ofrece una alta disponibilidad al enmascarar los fallos de servidor y red y al aparecer como un único servicio de alta disponibilidad.

- 4store: 4store es un motor de almacenamiento de bases de datos y de consultas que mantiene datos en formato RDF. Está escrito en ANSI C99, está diseñado para funcionar en sistemas UNIX y ofrece una plataforma de alto rendimiento, escalable y estable.
- MariaDB: Es una rama compatible hacia atrás de MySQL® Database Server. Incluye soporte para la mayoría de los motores de almacenamiento Open Source, y además para el propio motor de almacenamiento Maria.
- Drizzle: Es un fork de MySQL que se centra en ser una base de datos eficiente y estable, sobre todo optimizada para aplicaciones de Internet y que siguen la filosofía Cloud Computing.
- HyperSQL: Es un motor de bases de datos relacionales SQL escrito en Java, ofrece un pequeño pero veloz motor de bases de datos que dispone de tablas in-memory y basadas en disco, y que soporta los modos embebido y servidor. Además, dispone de herramientas como una consola de comandos SQL y una interfaz gráfica para las consultas.
- MonetDB: Es un sistema de bases de datos para aplicaciones de alto rendimiento dirigidas a la minería de datos, OAP, GIS, búsquedas XML, y recolección de información a partir de ficheros de texto y multimedia.
- Persevere: Es un motor de almacenamiento de objetos y un servidor de aplicaciones (ejecutándose en Java/Rhino) que ofrece almacenamiento de datos dinámicos JSON para el desarrollo rápido de aplicaciones de Internet orientadas a datos y basadas en JavaScript.
- eXist-db: eXist-db está desarrollada a través de la tecnología XML. Almacena datos CML según el modelo de datos de este estándar, y se caracteriza por un procesamiento eficiente y basado en índices de XQuery.
- Otras DB: Gladius (base de datos PHP flat-file), CloudStore, OpenQM (base de datos multivalor), ScarletDME (base de datos multivalor), SmallSQL (Motor de bases de datos Java Desktop SQL), LucidDB, HyperGraphDB (base de datos de grafos), InfoGrid (base de datos de grafos), Apache Derby, hamsterdb, H2 Database, EyeDB, txtSQL, db4o, Tokyo Cabinet, Project Voldemort.

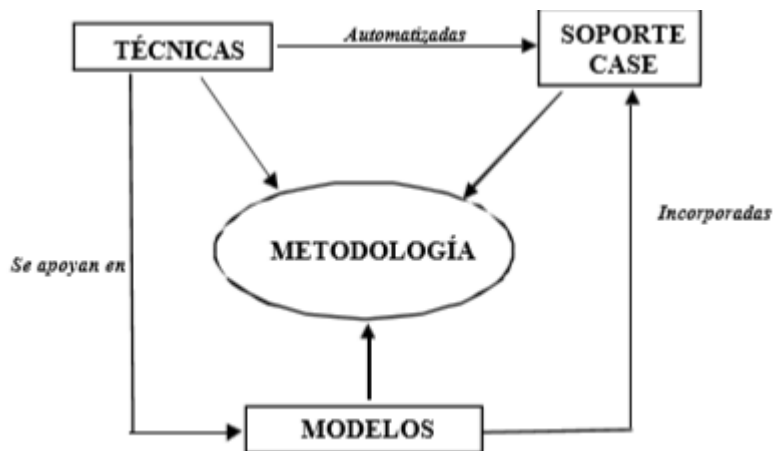
Administrador de la Base de Datos DBA:

Es el actor que administra las tecnologías de la información y la comunicación, siendo responsable de los aspectos técnicos, tecnológicos, científicos, inteligencia de negocios y legales de las bases de datos.

Bases de Datos Relacionales: Es una colección de elementos de datos organizados en un conjunto de tablas formalmente descritas desde la que se puede acceder a los datos o volver a gestionarlos de muchas maneras diferentes sin tener que reorganizar las tablas de la base de datos.

¿Qué es una metodología ¿Qué es una metodología y para qué sirve? y para qué sirve?: "Una metodología es un conjunto de procedimientos, técnicas y ayudas a la documentación para el desarrollo de un producto software".

- a) actividades a seguir en el desarrollo de principio a fin de la BD
- b) qué es lo que hay que realizar en cada actividad indicando
- c) qué se necesita como entrada,
- d) qué se produce como salida e
- e) quién está involucrado.



Modelos de datos como Modelos de datos como instrumentos de diseño:

MODELADO: Describir la estructura de la BD de forma que se represente fielmente la parcela del mundo real que se va a almacenar. Proceso de abstracción que se apoya en un modelo de datos (MD).

Modelos de datos como Modelos de datos como instrumentos de diseño:

TIPOS DE ABSTRACCIÓN:

La abstracción busca las propiedades comunes de un conjunto de objetos reduciendo la complejidad y ayudando a entender el mundo real.

- Clasificación/Particularización
- Agregación/Desagregación
- Generalización/Especialización
- Asociación/Disociación
- Jerarquía de Abstracciones

¿Qué es un Modelo de Datos (MD)? : Un modelo de datos proporciona un conjunto de conceptos, reglas y convenciones que nos permiten especificar y manipular los datos que queremos almacenar en la BD. Todo modelo de datos se compone de una parte estática y una parte dinámica.

COMPONENTES DE UN MD

- ESTÁTICA

- Conjunto de objetos
- Conjunto de asociaciones entre ellos
- Conjunto de restricciones
 - inherentes (propias del modelo)
 - semánticas (propias del usuario)

- DINÁMICA

- recuperación
- actualización

COMPONENTES DE UN MD

RESTRICCIONES: Limitaciones impuestas a la estructura del modelo o a los datos que invalidan ciertas ocurrencias de la BD

- Restricciones**
 - . Inherentes**
limitaciones impuestas a la estructura del modelo
 - . Semánticas**
limitaciones impuestas a los valores de los atributos o a las características de las interrelaciones

A. de Miguel y M. Piattini. Fundamentos y Modelos de Bases de Datos. Capítulo 3. Ed. 2. RA-MA. 1999.

A. de Miguel, M. Piattini y E. Marcos. Diseño de Bases de Datos Relacionales. Capítulos 1 y 8. RA-MA. 1999.

M. Piattini, E. Marcos, C. Calero y B. Vela. Tecnología y Diseño de Bases de Datos. Capítulos 4 y 13. RA-MA 2006.