

FUNDAMENTOS DE BASES DE DATOS

Introducción a las bases de datos

Allan Murillo
Marlen Treviño
Yessenia Calvo

Agenda

- Contacto
- Lineamientos, programa y evaluación
- Bases de datos
- Sistemas gestores de bases de datos

CONTACTO

Contacto

- Marlen Treviño Villalobos



marlentrevi@gmail.com



2401-3137

LINEAMIENTOS, PROGRAMA Y EVALUACIÓN

Objetivos

- ***Objetivo General***

- Comprender los principios que rigen el modelado de las bases de datos, que permitan la integridad de la información que se almacena, mediante lecturas y talleres adecuados a las posibilidades del entorno.

- ***Objetivos Específicos:***

- Comprender los principios de diseño de las bases de datos, mediante la utilización oportuna de los mismos, que permitan la creación de modelos de bases de datos acordes a las necesidades del entorno.
- Diseñar modelos de bases de datos, mediante la aplicación de fundamentos teóricos y prácticos, logrando con ello aprovechar las posibilidades de los entornos informáticos que utilizan las bases de datos para su trabajo efectivo y eficiente.
- Conocer las instrucciones del lenguaje de SQL, mediante lecturas y aplicación de casos, que le permitan al estudiante el máximo aprovechamiento de las técnicas de búsqueda, creación y almacenamiento de datos.

Programa

Tema	Semana
Introducción a las bases de datos	1
Diseño de Modelos de bases de datos	2, 3 y 4
Normalización	5
Lenguaje SQL	6
Examen I	7
Lenguaje SQL	8 y 9
Algebra relacional	10
Lenguaje SQL	11, 12 y 13
Examen II	14

Evaluación

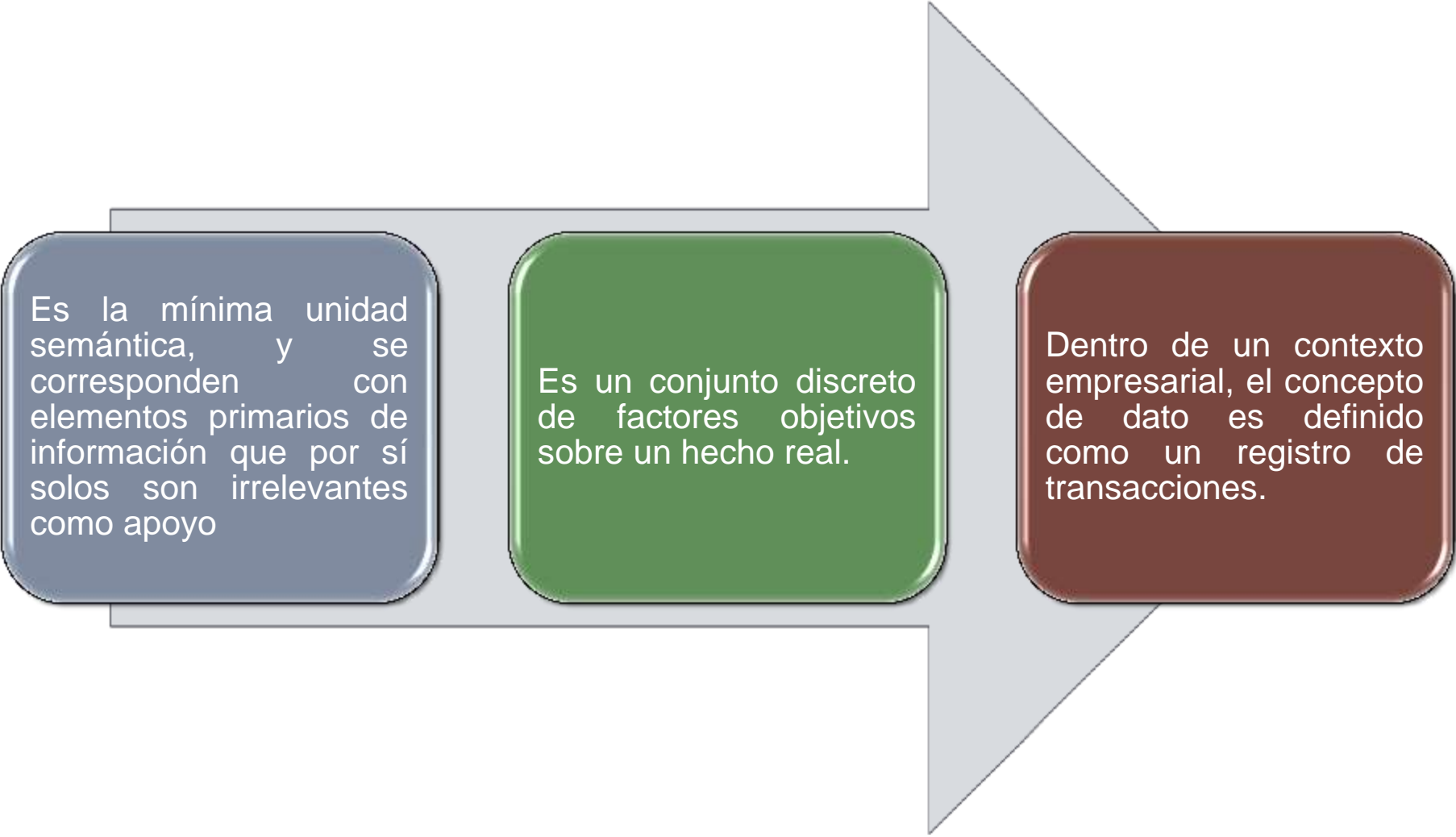
Tareas y Quices	20%
Proyecto	30%
Examen I	25%
Examen II	25%
TOTAL	100%

Materiales disponibles en:

<https://sites.google.com/site/fundamentosbasesdatos2013/>

BASE DE DATOS

Datos: Definición



Es la mínima unidad semántica, y se corresponden con elementos primarios de información que por sí solos son irrelevantes como apoyo

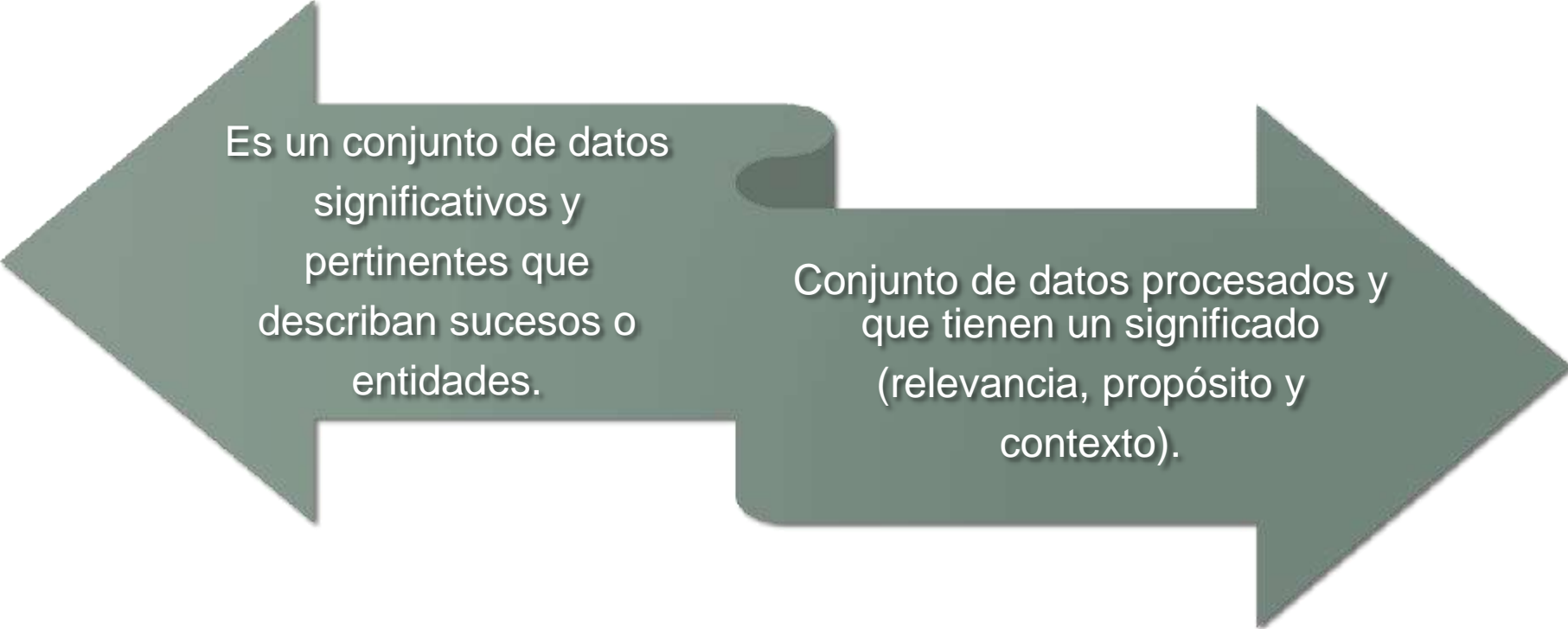
Es un conjunto discreto de factores objetivos sobre un hecho real.

Dentro de un contexto empresarial, el concepto de dato es definido como un registro de transacciones.

Datos: Definición

- El mundo está lleno de datos:
 - Pacientes/Doctores.
 - Alumnos/Profesores.
 - Cuentas bancarias.
 - Deportistas /Deportes.
 - Actores/Películas.
 - Recetas de cocina.
 - Superhéroes, etc.
 - Qué queremos hacer con ellos?
 - Consultarlos.
 - Actualizarlos.
- Con el propósito de obtener información

Información: Definición



Es un conjunto de datos
significativos y
pertinentes que
describan sucesos o
entidades.

Conjunto de datos procesados y
que tienen un significado
(relevancia, propósito y
contexto).

Información: Características

Exactitud

- En este sentido la información debe reflejar el evento epidemiológico al cual se refiere y su sistema de medición expresado con poca variabilidad.

Objetividad

- La información debe ser el producto de criterios establecidos que permitan la interpretación en forma estandarizada por diferentes personas en circunstancias diversas de tiempo y lugar.

Válida

- Se refiere a que la información ha de permitir medir en forma precisa el concepto que se estudia, con criterios uniformes.

Continuidad

- La información ha de ser generada en forma permanente de tal manera que exista la disponibilidad de los datos a través del proceso de vigilancia.

Completa

- Debe contener todos los datos y variables previamente establecidas para cumplir con su finalidad en cada evento epidemiológico .

Oportuna

- La información debe generarse y notificarse a la par con los acontecimientos de tal manera que permita la toma de decisiones y la actuación Inmediata.

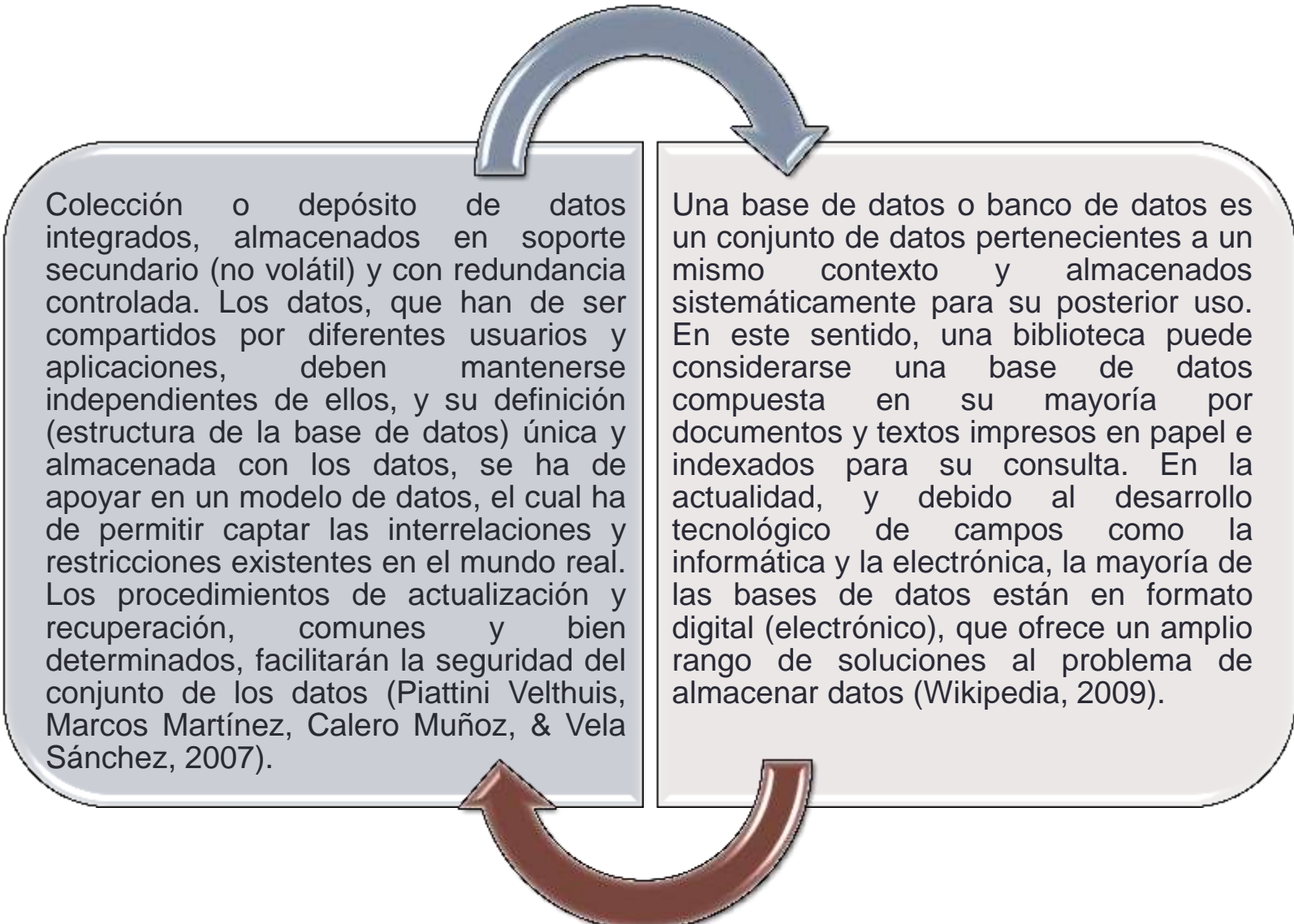
Comparable

- Que permita ser confrontada con datos similares.

Base de datos: Definición

- En grupos de tres personas, comente qué conoce sobre el concepto de Bases de datos y elabore un mapa conceptual.
- Luego, lea uno de los artículos que su profesor le facilitará y mejore su mapa conceptual.
- Posteriormente, cada grupo expone sus conocimientos y agrupamos los componentes en común para formar un solo concepto.

Base de datos: Definición



Colección o depósito de datos integrados, almacenados en soporte secundario (no volátil) y con redundancia controlada. Los datos, que han de ser compartidos por diferentes usuarios y aplicaciones, deben mantenerse independientes de ellos, y su definición (estructura de la base de datos) única y almacenada con los datos, se ha de apoyar en un modelo de datos, el cual ha de permitir captar las interrelaciones y restricciones existentes en el mundo real. Los procedimientos de actualización y recuperación, comunes y bien determinados, facilitarán la seguridad del conjunto de los datos (Piattini Velthuis, Marcos Martínez, Calero Muñoz, & Vela Sánchez, 2007).

Una base de datos o banco de datos es un conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso. En este sentido, una biblioteca puede considerarse una base de datos compuesta en su mayoría por documentos y textos impresos en papel e indexados para su consulta. En la actualidad, y debido al desarrollo tecnológico de campos como la informática y la electrónica, la mayoría de las bases de datos están en formato digital (electrónico), que ofrece un amplio rango de soluciones al problema de almacenar datos (Wikipedia, 2009).

Base de datos: Historia

- Ver video
 - Historia de la base de datos

Base de datos: Características

Integrada

- Se entiende que una base de datos puede considerarse como una unificación de varios archivos de datos independientes

Compartida

- Se entiende que partes individuales de la Base de Datos pueden compartirse entre varios usuarios distintos, en el sentido que cada uno de ellos puede tener acceso a la misma parte de la Base de Datos y utilizarla con propósitos diferentes

Vistas

- Diferentes usuarios pueden tener una perspectiva o vista diferente de los datos. Una vista es un subconjunto de los datos o datos virtuales derivados de los almacenados en la base de datos

Base de datos: Características

Redundancia controlada

- No existan duplicidades perjudiciales ni innecesarias. Las redundancias físicas, convenientes muchas veces a fin de responder a objetivos de eficiencia, son tratadas por el mismo sistema, de modo que no puedan producirse incoherencias. Esto significa que en las bases de datos NO se espera tener redundancia lógica, pero si se admite cierta redundancia física por motivos de eficiencia

Metadata

- En las bases de datos, la descripción, y en algunos casos, también una definición y documentación completa sobre los datos, se almacenan junto con estos, de modo que estos están documentados, y cualquier cambio que se produzca debe quedar recogido en el sistema

Independencia

- Los mecanismos de abstracción permiten que exista una independencia entre los datos y los programas. Esta abstracción de los datos de una base de datos está asociada con el modelo de datos que permite definir un SGBD.

Sistema de base de datos vs. Sistema de archivos

Sistema de base de datos

Se mantiene un **único almacén de datos** que se define una sola vez y al cual tienen acceso muchos usuarios

Naturaleza autodescriptiva de los sistemas de base de datos: el sistema no solo contiene la base de datos misma, sino también una **definición o descripción completa de esta**.

Separación entre los programas y los datos: los programas de acceso del SGBD deben ser independientes de cualquier archivo específico. **La estructura de los archivos de datos está en el catálogo aparte de los programas de acceso**

Abstracción de los datos: el SGBD provee al usuario una **representación conceptual de la data** que no incluye (necesariamente) detalles de su implementación

Sistema de archivos

Cada usuario define e implementa los archivos requeridos para una aplicación específica

La definición de los datos es parte de los programas de aplicación, por tanto dichos **programas sólo pueden trabajar con una base de datos específica**

La estructura de los archivos de datos viene integrada en los programas de acceso, así que **cualquier modificación de la estructura de un archivo puede requerir la modificación de todos los programas que tienen acceso a dicho archivo**.

Cada archivo puede estar definido por su **longitud de registros** (# de bytes por registro) y cada campo puede ser especificado por su byte de comienzo dentro de un registro y su longitud en byte.

Sistema de base de datos vs. Sistema de archivos

Sistema de base de datos

Comportamiento de datos y procesamiento de transacciones multiusuarios: los SGBD deben permitir el **acceso simultáneo a varios usuarios**. El SGBD debe incluir software de control de concurrencia para asegurarse de que las actualizaciones de un dato sean las **correctas**, además de asegurarse de que estas actualizaciones estén disponibles para todos lo Usuarios

Manejo de múltiples vistas de los datos: **cada uno de los usuarios de la base de datos puede requerir una perspectiva o vista diferente de la misma**. Una vista puede ser un subconjunto de la base de datos o contener datos virtuales que se deriven de los archivos de la base de datos, pero que no estén almacenados explícitamente.

Sistema de archivos

En los sistemas de archivos existen muchas **incoherencias en los datos**, debido a la falla en las actualizaciones.

En los sistemas de archivos **no existen diferentes vistas de datos**

SISTEMA GESTOR DE BASE DE DATOS

SGBD: Definición

- Colección de datos relacionados entre sí, estructurados y organizados, y un conjunto de programas que acceden y gestionan esos datos.
- Es un software específico que permite a los usuarios crear, mantener y manipular la BD (incorporando transparente e internamente los mecanismos necesarios para su funcionamiento correcto) .
- Objetivo principal:
 - Proporcionar una forma de almacenar y recuperar la información de una base de datos de manera que sea tanto práctica como eficiente.

SGBD: Características



SGBD: Características



SGBD: Características

- Integridad

- Calidad de la información: *“los datos deben estar estructurados reflejando adecuadamente los objetos, relaciones y las restricciones existentes en la parcela del mundo real que modela la base de datos”*
- SGBD debe asegurar que los datos se almacenen correctamente
- SGBD debe asegurar que las actualizaciones de los usuarios sobre la base de datos se ejecutan correctamente y que se hacen permanentes
- Ejemplo
 - El saldo de una cuenta bancaria no puede nunca ser más bajo de una cantidad predeterminada (por ejemplo 25 US).

SGBD: Características



SGBD: Características

- Seguridad
 - Objetivo: *Sólo pueden acceder a la información las personas y procesos autorizados y en la forma autorizada.*



- Ejemplo
 - En un sistema Bancario, el personal de nóminas necesita ver sólo esa parte de la DB que tiene información acerca de varios empleados del banco.
 - No necesita acceder a la información acerca de las cuentas de los clientes.

SGBD: Características



SGBD: Características

- Redundancia

- En las bases de datos se almacenan grandes cantidades de datos, que son definidos una sola vez y que pueden ser accedidas por varios usuarios a la vez, teniendo todos los datos integrados y creando una dependencia de datos a la organización y no a los departamentos o núcleos de la misma, eliminando la redundancia de datos y estableciendo una mínima duplicidad de los datos.

Ejemplo

La dirección y número de teléfono de un cliente particular puede aparecer en un archivo que contenga **cuentas de ahorros** y en un archivo que contenga registros de una **cuenta corriente**.

Esta redundancia conduce a un **almacenamiento y coste de acceso más alto**. Además, puede conducir a la **inconsistencia de los datos**. "Es decir, las diversas copias de los mismos datos pueden NO coincidir"



SGBD: Características

- En parejas, utilice una hoja de excel para responder las siguientes preguntas:
 - Escriba los videojuegos más conocidos por ustedes
 - Identifique en qué género se encuentran: Acción, Carreras, Deportes, Simulación, etc
 - Qué tipo de consola se utiliza para jugar
 - Clasificación:
 - Mayores de 3 años
 - Mayores de 10 años
 - Mayores de 12 años
 - Mayores de 18 años
 - Cuál fue la empresa encargada de desarrollar el videojuego
- Luego, unifica tus respuestas con otro grupo
- Identifica en ese grupo de datos la característica de redundancia

SGBD: Servicios

Creación y definición de la BD

- Especificación de la estructura, el tipo de los datos, las restricciones y relaciones entre ellos mediante lenguajes de definición de datos. Toda esta información se almacena en el diccionario de datos, el SGBD proporcionará mecanismos para la gestión del diccionario de datos.

Manipulación de los datos

- Realizando consultas, inserciones y actualizaciones de los mismos utilizando lenguajes de manipulación de datos.

Acceso controlado a los datos de la BD

- Mediante mecanismos de seguridad de acceso a los usuarios.

Mantener la integridad y consistencia de los datos

- Utilizando mecanismos para evitar que los datos sean perjudicados por cambios no autorizados.

Acceso compartido a la BD

- Controlando la interacción entre usuarios concurrentes.

Mecanismos de respaldo y recuperación

- Para restablecer la información en caso de fallos en el sistema.

SGBD: Actores

Administrador de la base de datos

Diseñador de la base de datos

Usuarios finales

Desarrolladores de herramientas

SGBD: Actores

Administrador de la base de datos

Diseñador de la base de datos

Usuarios finales

Desarrolladores de herramientas

SGBD: Actores

- **Administrador** de la base de datos (ABD)
 - Responsable de administrar los recursos del SGBD (nivel técnico):
 - BD + SGBD + Otro Software (aplicaciones/programas de acceso)
 - Las funciones del ABD incluyen:
 - Definir/Modificar la **estructura** de la BD y **restricciones** de los datos
 - Crear/Modificar **estructuras de almacenamiento** y **métodos de acceso**
 - Conceder/Denegar **permisos de acceso** y controlar dicho acceso a datos (seguridad de la BD)
 - Definir **planes de copias de seguridad** de los datos de la BD
 - Garantizar el funcionamiento correcto del sistema y proporcionar servicio técnico al usuario (respuesta lenta del sistema)
 - Adquirir los recursos software y hardware necesarios

SGBD: Actores

Administrador de la base de datos

Diseñador de la base de datos

Usuarios finales

Desarrolladores de herramientas

SGBD: Actores

- **Diseñadores** de la base de datos
 - Antes de implementar la BD, interactúan con sus futuros usuarios
 - Recogen y comprenden sus necesidades y objetivos --- Requisitos
 - **Identifican datos** que almacenar en la base de datos
 - Eligen estructuras para representar y almacenar los datos
 - Construyen...
 - **Vistas** que satisfacen requisitos de cada grupo de usuarios
 - **Diseño final de BD** que satisface necesidades de todos los usuarios (resultado de la integración de las diferentes vistas)

SGBD: Actores

Administrador de la base de datos

Diseñador de la base de datos

Usuarios finales

Desarrolladores de herramientas

SGBD: Actores

- **Usuarios finales**
 - Ocasionales o Sofisticados
 - Acceso esporádico y distinto cada vez; usan lenguaje de consulta
 - Paramétricos o Normales
 - Accesos constantes, repetitivos
 - Usan “transacciones programadas” para ellos
 - Avanzados o Especializados
 - Implementan sus propias aplicaciones especializadas para cumplir sus complejos requisitos
 - Ingenieros, científicos, analistas de empresa, ...
 - Autónomos
 - Usan BD personales, a través de aplicación/paquete comercial específico

SGBD: Actores

Administrador de la base de datos

Diseñador de la base de datos

Usuarios finales

Desarrolladores de herramientas

SGBD: Actores

- **Desarrolladores de herramientas**
 - Diseñan e Implementan paquetes software que:
 - Facilitan diseño y uso de los SGBD, y
 - Permiten aumentar el rendimiento de los SGBD
 - Herramientas para
 - Diseño de BD, de aplicaciones de BD, de interfaces de usuario...
 - Creación de prototipos de aplicaciones de BD
 - Realizar simulaciones y generar datos de prueba
 - Supervisión de rendimiento del sistema
 - ...

SGBD: Aplicaciones



BANCA

- Para información de los clientes, cuentas y préstamos, y transacciones.



LÍNEAS AÉREAS

- Para reservas e información de planificación.



UNIVERSIDADES

- Para información de los estudiantes, matriculas de las asignaturas y cursos.



RECURSOS HUMANOS

- Para información sobre los empleados, salarios, impuestos y beneficios, y para la generación de nominas.

SGBD: Aplicaciones



TELECOMUNICACIONES

- Para guardar un registro de las llamadas realizadas, generación mensual de facturas, manteniendo el saldo de las tarjetas telefónicas prepago y para almacenar información sobre redes de comunicaciones.



FINANZAS

- Para almacenar información sobre grandes empresas, ventas y compras de documentos formales financieros, como bolsa y bonos.



VENTAS

- Para información de clientes productos y compras.



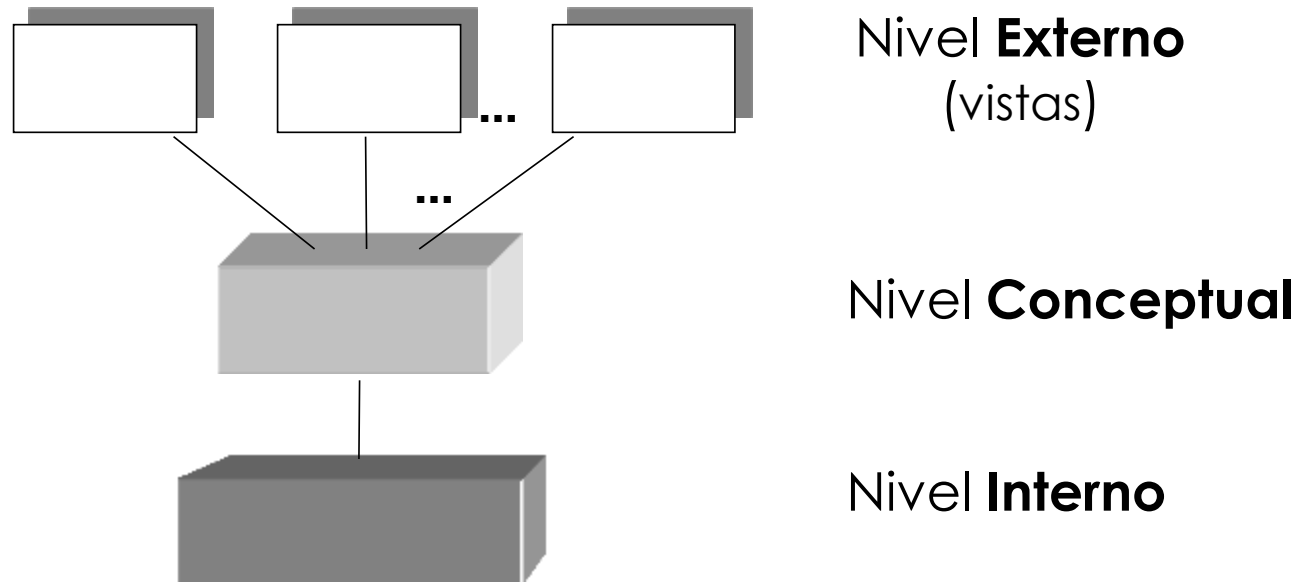
PRODUCCIÓN

- Para la gestión de la cadena de producción y para el seguimiento de la producción de elementos en las factorías, inventarios de elementos en almacenes y pedidos de elementos.

SGBD: Aplicaciones

- Ver el video
 - The usefulness of DBMS

SGBD: Niveles de arquitectura



SGBD: Niveles de arquitectura

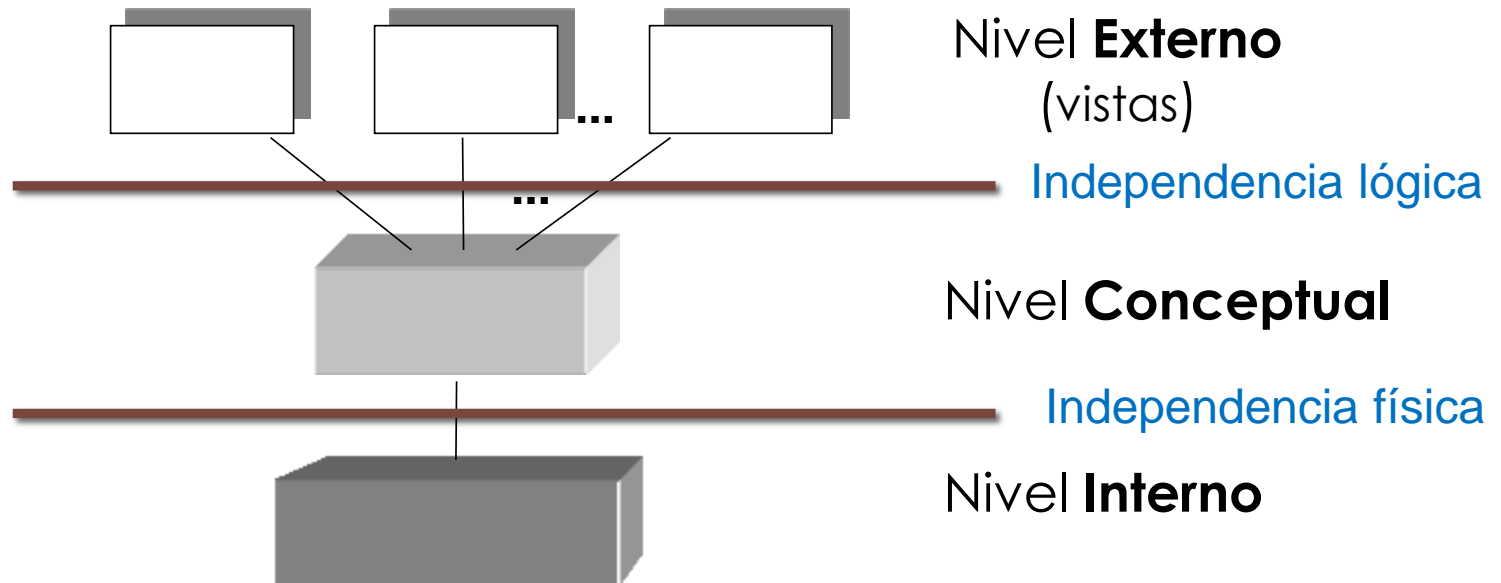
- Nivel interno o físico
 - El más cercano al **almacenamiento físico**, es decir, tal y como están almacenados en el ordenador. Describe la estructura física de la BD mediante un esquema interno. Este esquema se especifica con un modelo físico y describe los detalles de cómo se almacenan físicamente los datos: los archivos que contienen la información, su organización, los métodos de acceso a los registros, los tipos de registros, la longitud, los campos que los componen, etcétera.
- Nivel externo o de visión
 - Es el más cercano a los usuarios, es decir, es donde se describen varios **esquemas externos o vistas de usuarios**. Cada esquema describe la parte de la BD que interesa a un grupo de usuarios en este nivel se representa la visión individual de un usuario o de un grupo de usuarios.
- Nivel conceptual
 - Describe la **estructura de toda la BD** para un grupo de usuarios mediante un esquema conceptual. Este esquema **describe las entidades, tributos, relaciones, operaciones de los usuarios y restricciones, ocultando los detalles de las estructuras físicas de almacenamiento**. Representa la información contenida en la BD.

SGBD: Niveles de arquitectura

- El SGBD debe de transformar cualquier petición de usuario (esquema externo) a una petición expresada en términos de esquema conceptual, para finalmente ser una petición expresada en el esquema interno que se procesará sobre la BD almacenada. El proceso de transformar peticiones y resultados de un nivel a otro se denomina **correspondencia o transformación**, el SGBD es capaz de interpretar una solicitud de datos y realiza los siguientes pasos:



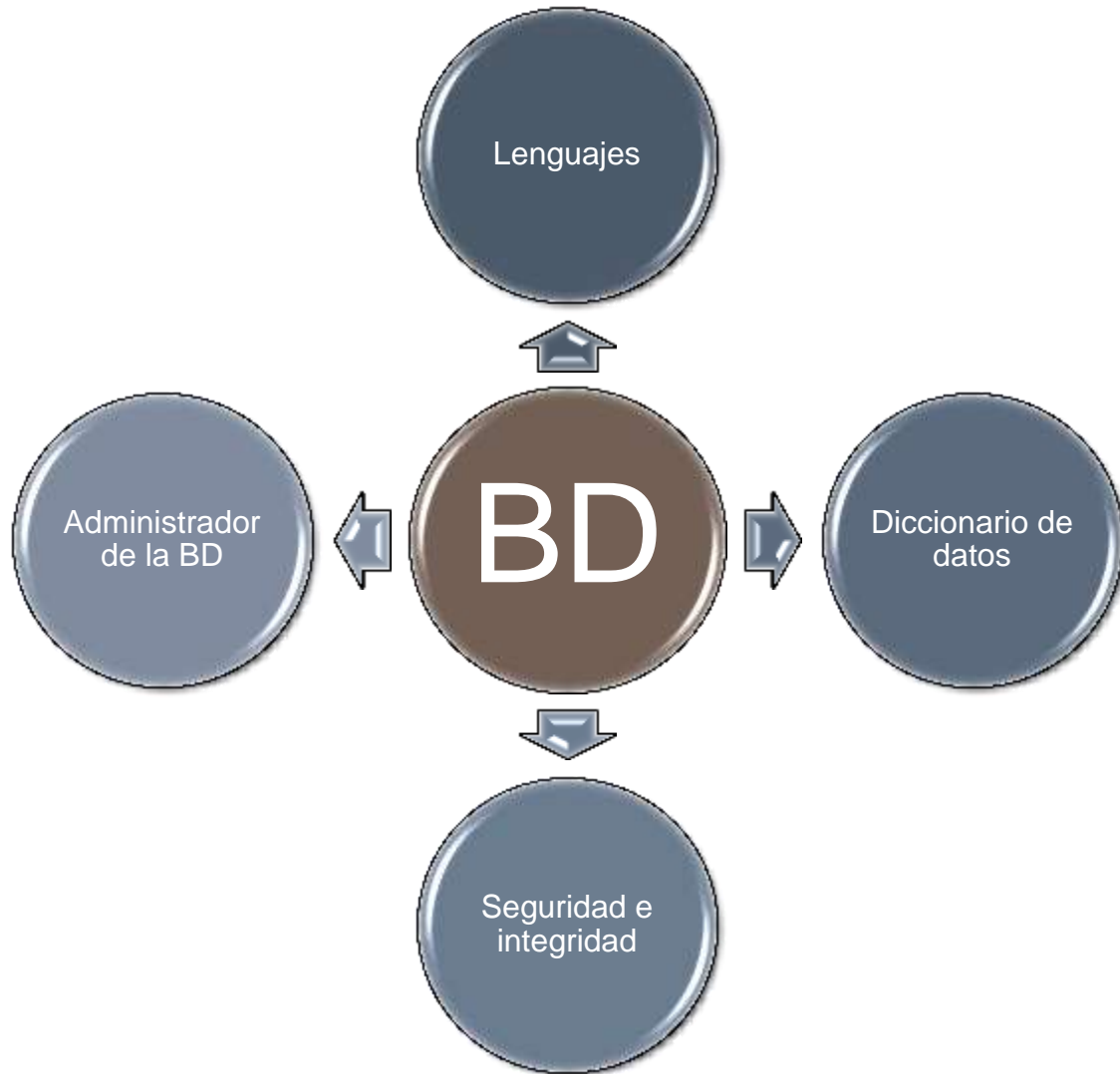
SGBD: Independencia de datos



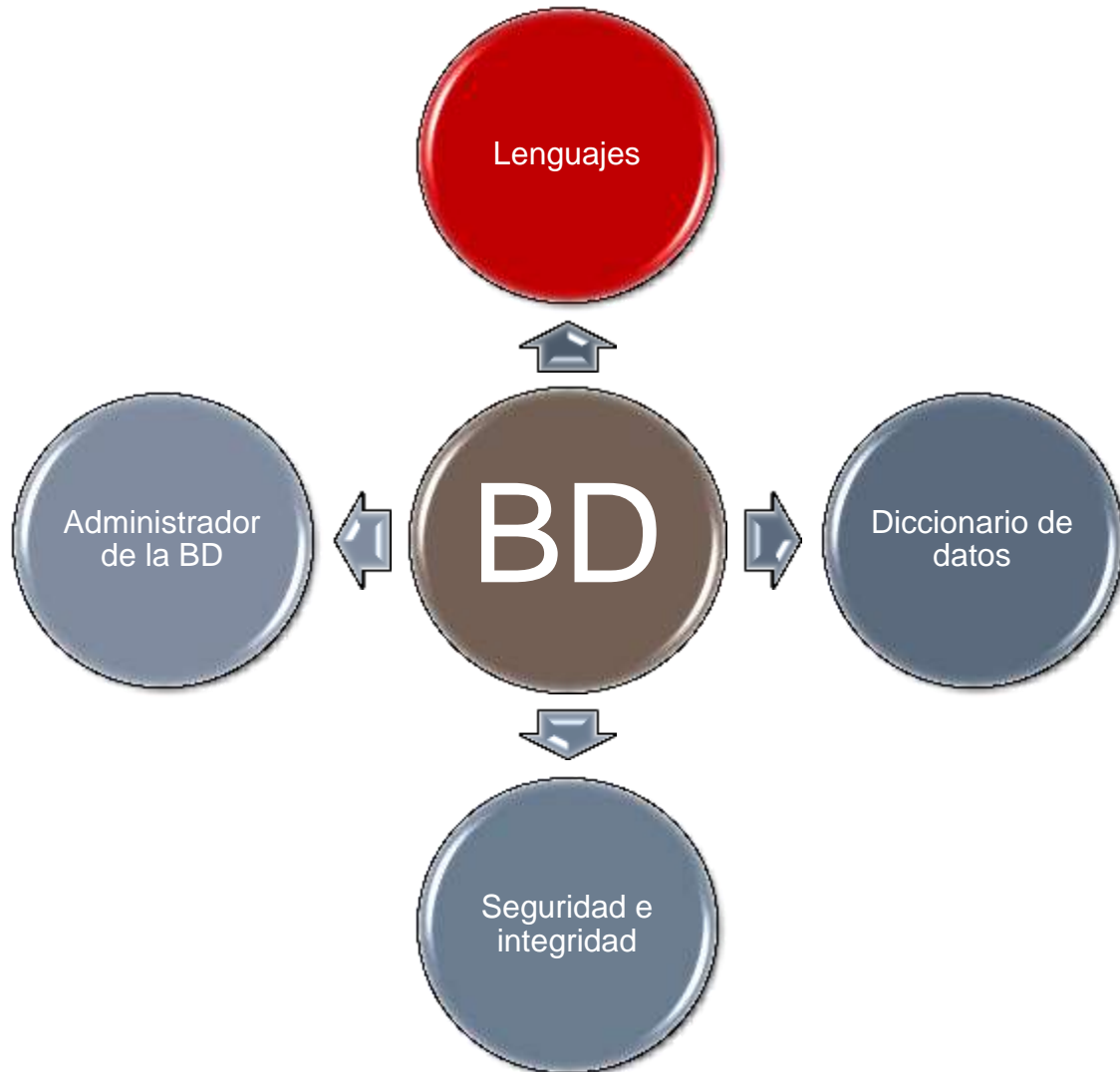
SGBD: Independencia de datos

- La independencia de datos es la capacidad para modificar el esquema en un nivel del sistema sin tener que modificar el esquema del nivel inmediato superior.
- Así, las aplicaciones están “aisladas” (no deben preocuparse) de cómo están los datos estructurados y almacenados.
- **Independencia lógica de los datos:**
 - Protección ante cambios en la estructura lógica de los datos.
 - Se podrá modificar el esquema conceptual para ampliar la BD o para reducirla, por ejemplo, si se elimina una entidad, los esquemas externos que no se refieran a ella no se verán afectados.
- **Independencia física de los datos:**
 - Protección ante cambios en las estructuras físicas de los datos.
 - Por ejemplo, se pueden reorganizar los archivos físicos con el fin de mejorar el rendimiento de las operaciones de consulta o de actualización, o se pueden añadir nuevos archivos de datos porque los que había se han llenado.
 - La independencia física es más fácil de conseguir que la lógica, pues se refiere a la separación entre las aplicaciones y las estructuras físicas de almacenamiento.

SGBD: Componentes



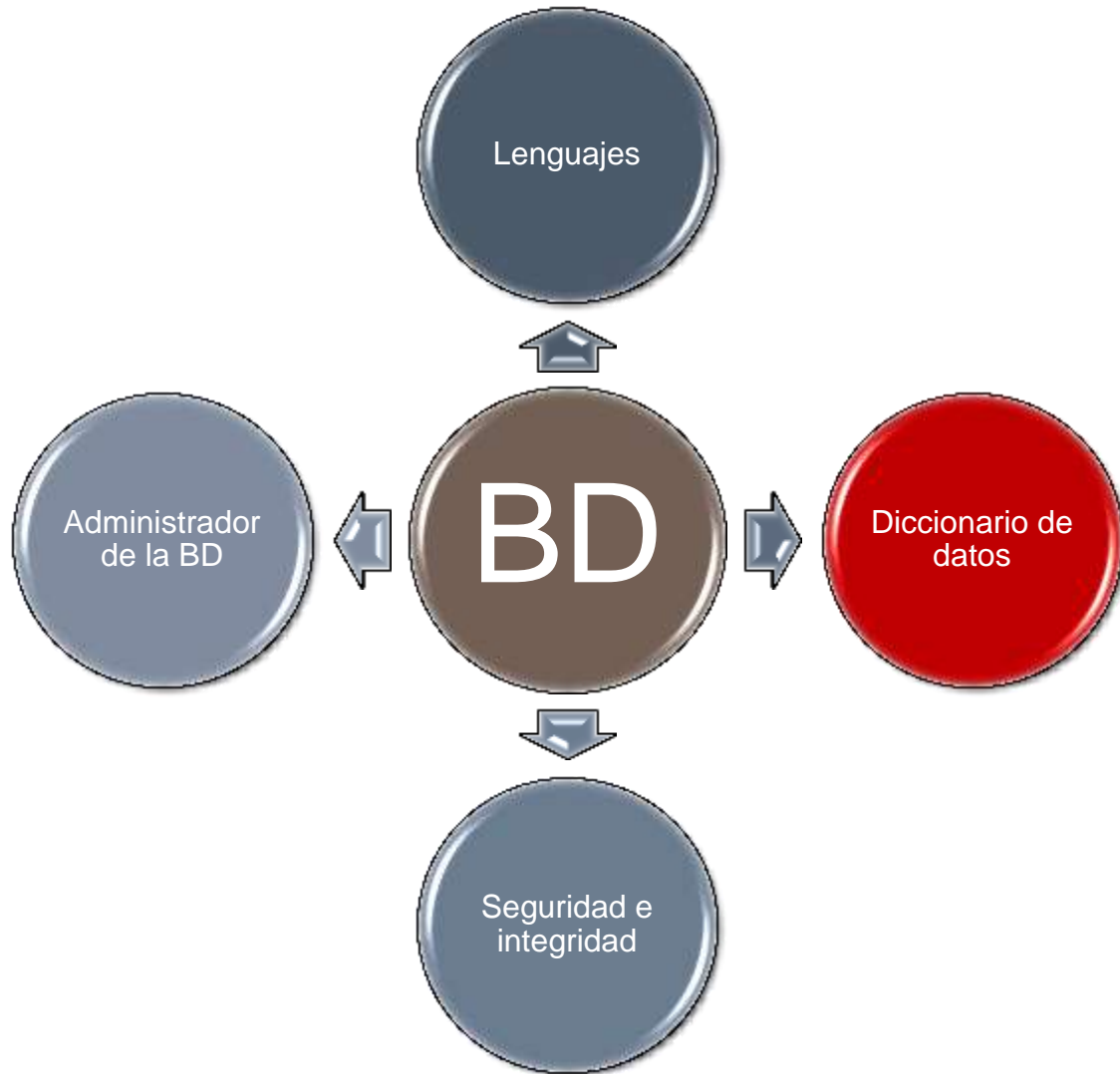
SGBD: Componentes



SGBD: Componentes

- Todos los SGBD ofrecen lenguajes e interfaces apropiadas para cada tipo de usuario: administradores, diseñadores, programadores de aplicaciones y usuarios finales.
- Los lenguajes van a permitir al administrador de la BD **especificar los datos que componen la BD, su estructura, las relaciones que existen entre ellos, las reglas de integridad, los controles de acceso, las características de tipo físico y las vistas externas de los usuarios.**
- Los lenguajes del SGBD se clasifican en:
 - **Lenguaje de definición de datos (LDD o DDL):** se utiliza para especificar el esquema de la BD, las vistas de los usuarios y las estructuras de almacenamiento. Es el que define el esquema conceptual y el esquema interno. Lo utilizan los diseñadores y los administradores de la BD.
 - **Lenguaje de manipulación de datos (LMD o DML):** se utilizan para leer y actualizar los datos de la BD. Es el utilizado por los usuarios para realizar consultas, inserciones, eliminaciones y modificaciones.
 - Los hay procedurales, en los que el usuario será normalmente un programador y especifica las operaciones de acceso a los datos llamando a los procedimientos necesarios. Estos lenguajes acceden a un registro y lo procesan. Las sentencias de un LMD procedural están embebidas en un lenguaje de alto nivel llamado anfitrión.
 - No procedurales son los lenguajes declarativos. En muchos SGBD se pueden introducir interactivamente instrucciones del LMD desde un terminal, también pueden ir embebidas en un lenguaje de programación de alto nivel. Estos lenguajes permiten especificar los datos a obtener en una consulta, o los datos a modificar, mediante sentencias sencillas. Las BD relacionales utilizan lenguajes no procedurales como SQL (Structured Query Language) o QBE (Query By Example).

SGBD: Componentes



SGBD: Componentes

- El diccionario de datos es el **lugar** donde se **deposita información acerca de todos los datos que forman la BD**. Es una guía en la que se describe la BD y los objetos que la forman.
- El diccionario contiene las **características lógicas** de los sitios donde se almacenan los datos del sistema, incluyendo nombre, descripción, alias, contenido y organización. Identifica los procesos donde se emplean los datos y los sitios donde se necesita el acceso inmediato a la información.

Un diccionario de datos debe cumplir las siguientes características:

Debe soportar las descripciones de los modelos conceptual, lógico, interno y externo de la BD.

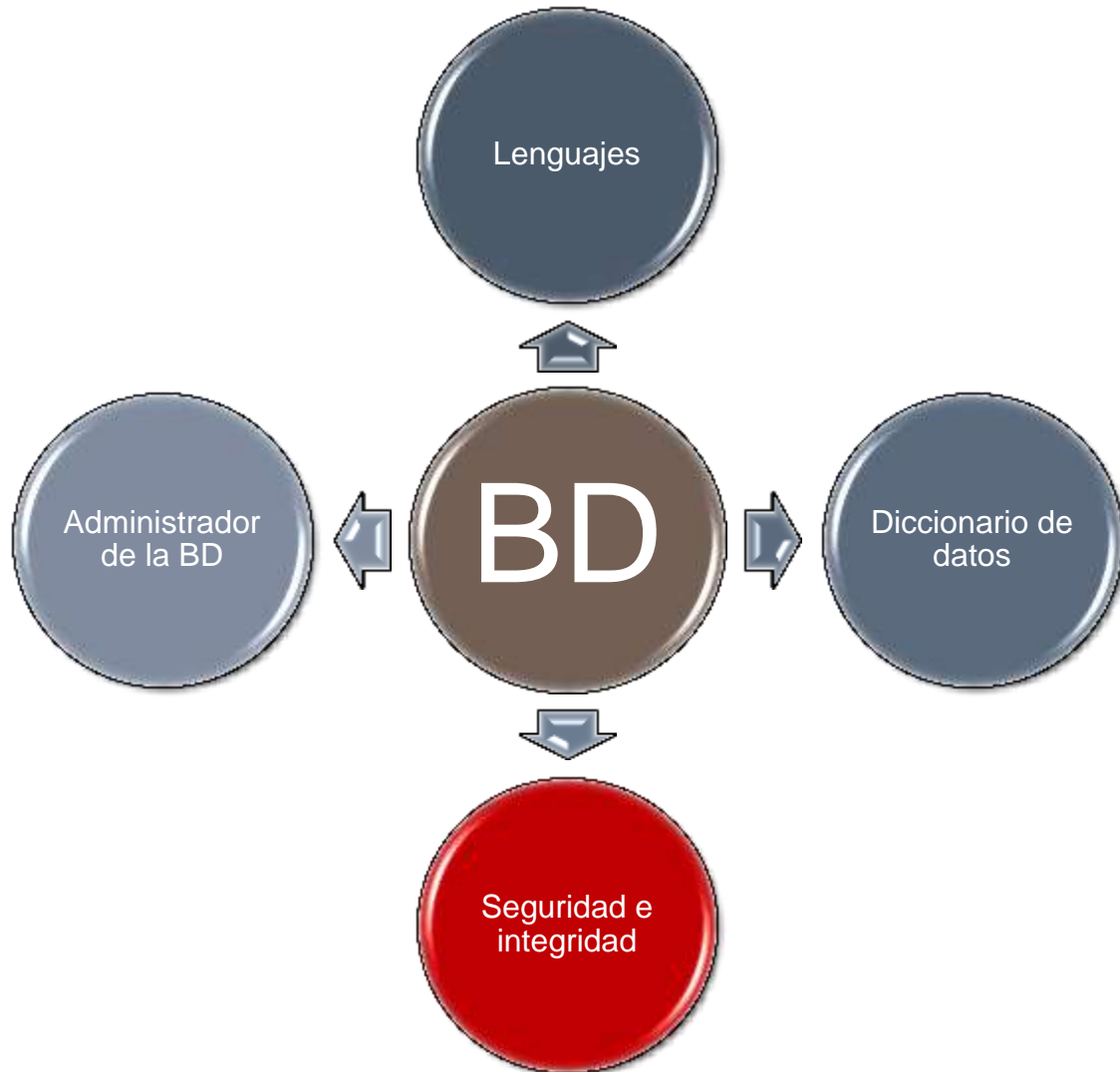
Debe estar integrado dentro del SGBD.

Debe apoyar la transferencia eficiente de información al SGBD. La conexión entre los modelos interno y externo debe ser realizada en tiempo de ejecución.

Debe comenzar con la reorganización de versiones de producción de la BD. Además debe reflejar los cambios en la descripción de la BD.

Debe estar almacenado en un medio de almacenamiento con acceso directo para la fácil recuperación de información.

SGBD: Componentes



SGBD: Componentes

Debe garantizar la protección de los datos contra accesos no autorizados, tanto intencionados como accidentales. Debe controlar que sólo los usuarios autorizados accedan a la BD.

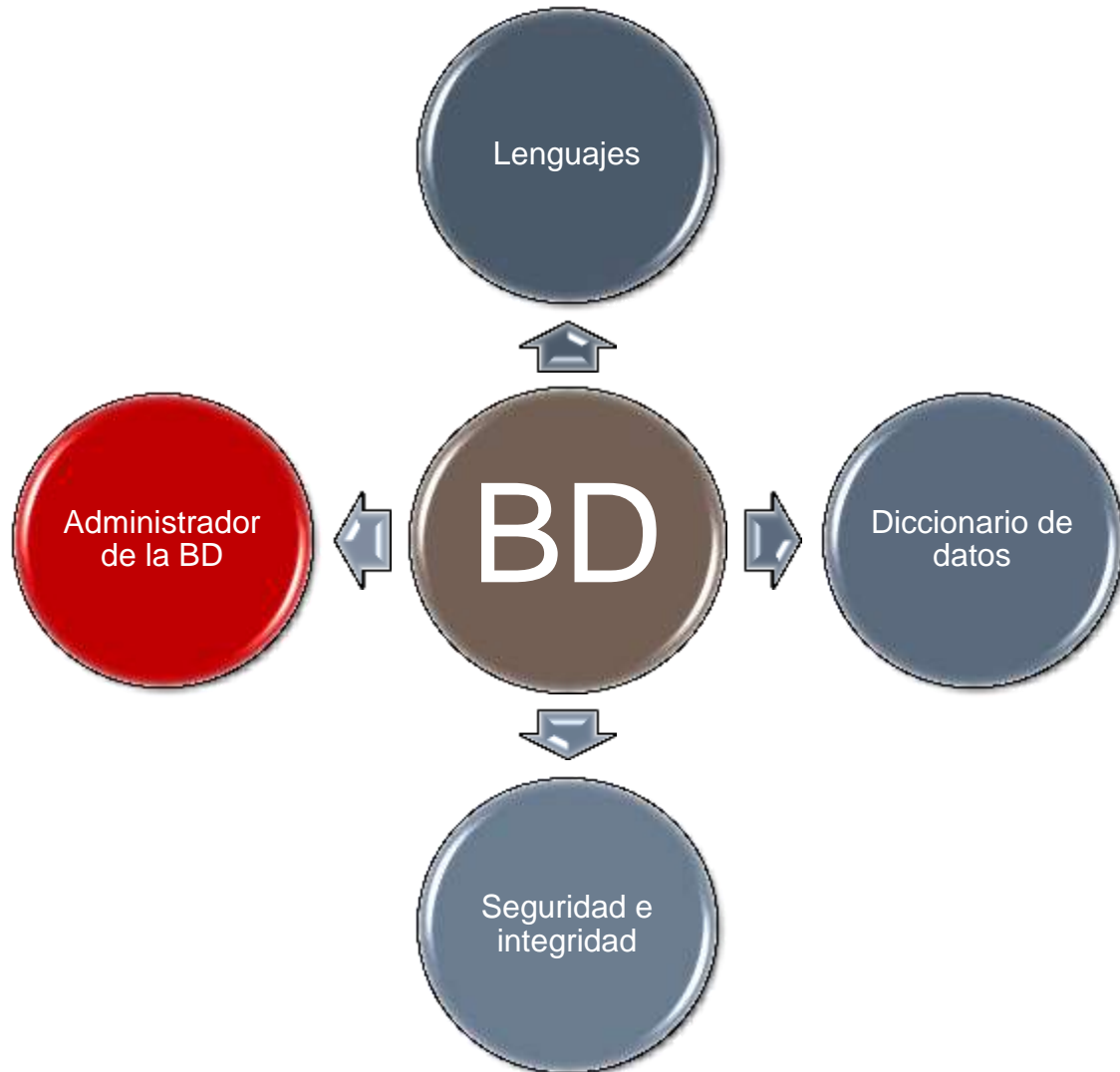
Los SGBD ofrecen mecanismos para implantar restricciones de integridad en la BD. Estas restricciones van a proteger la BD contra daños accidentales

Proporciona herramientas y mecanismos para la planificación y realización de copias de seguridad y restauración.

Debe ser capaz de recuperar la BD llevándola a un estado consistente en caso de ocurrir algún suceso que la dañe.

Debe asegurar el acceso concurrente y ofrecer mecanismos para conservar la consistencia de los datos en el caso de que varios usuarios actualicen la BD de forma concurrente.

SGBD: Componentes



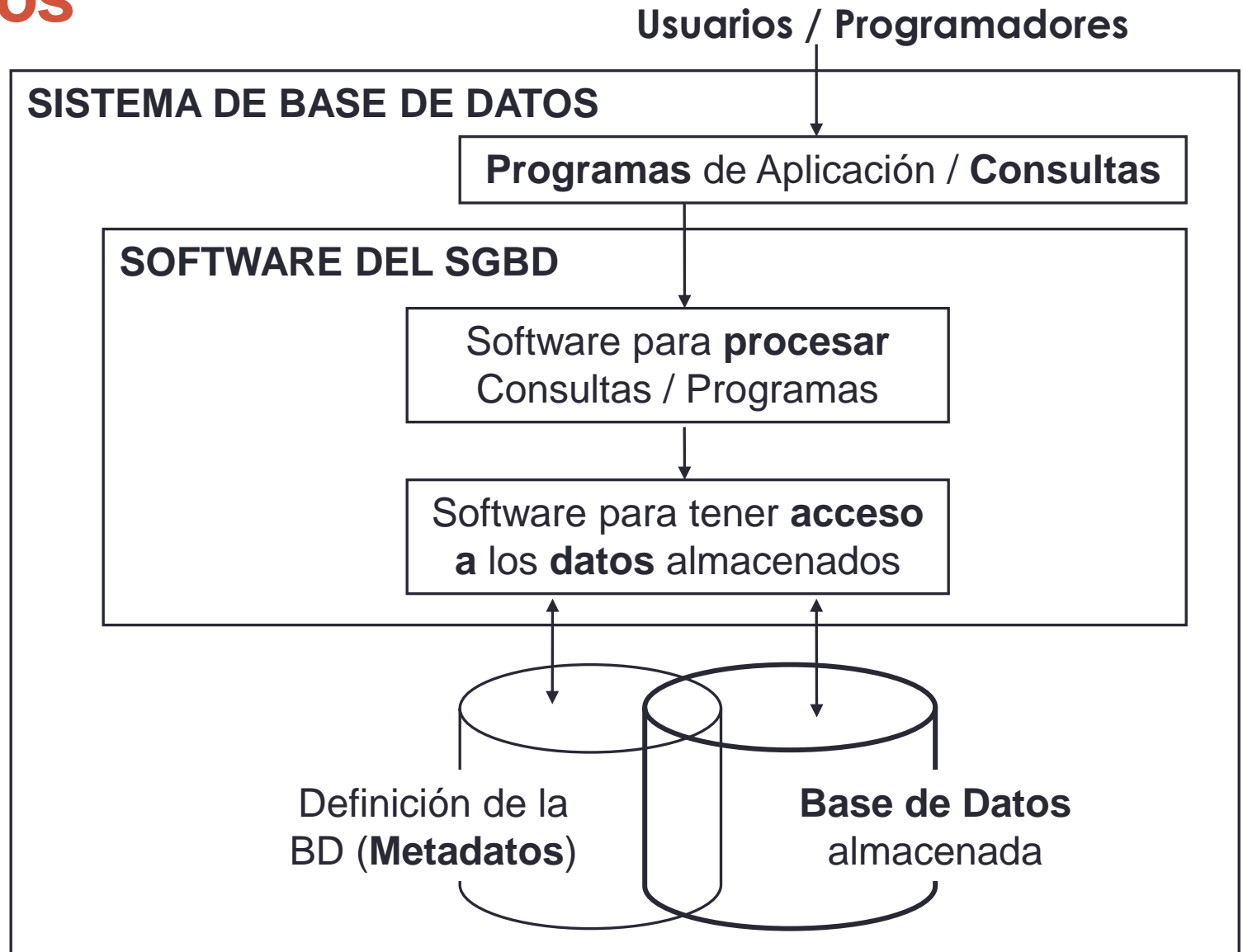
SGBD: Componentes

- El DBA tiene una gran responsabilidad ya que posee el máximo nivel de privilegios. Será el encargado de crear los usuarios que se conectarán a la BD. En la administración de una BD siempre hay que procurar que haya el menor número de administradores, a ser posible una sola persona.

SGBD: Clasificación

- **Según el Modelo de Datos en el que está basado**
 - Relacional, Red, Jerárquico,
 - Orientado a Objetos,
 - Objeto/Relacional, ...
- **Según el nº de usuarios simultáneos**
 - Monousuario
 - Multiusuario
- **Según el nº de lugares en que se almacenan datos**
 - Centralizado
 - Distribuido (SGBDD)
 - SGBDD homogéneo: mismo software de SGBD en todos los sitios
 - SGBDD heterogéneo (ej. Multi-Base de Datos o BD Federadas)
- **Según su propósito**
 - de propósito General
 - de propósito Específico: construido para un tipo concreto de aplicaciones

Entorno simplificado de un **Sistema de Base de Datos**



SGBD: Ejemplos

Algunos de los Sistemas de Gestión de Base de Datos más conocidos:

- SQL
- DB2
- PostgreSQL
- ORACLE
- INGRES
- INFORMIX
- MySQL
- SYBASE
- PARADOX
- DBASE
- ACCESS
- FOXPRO

Repaso

- ¿Por qué elegirías un sistema de BD en lugar de almacenar la información en ficheros del sistema operativo?
- ¿Cuál es la diferencia entre una base de datos y un sistema gestor de base de datos?
- Explica la diferencia entre independencia lógica e independencia física.
- Explica la diferencias entre esquema externo, conceptual y físico.
- ¿Qué es el diccionario de datos? Explique sus características
- ¿Cuáles son las formas en que se clasifican las bases de datos?

Tarea

- En parejas investigue sobre un Sistema Gestor de Base de datos y prepare un exposición de máximo 10 minutos.
 - SQL
 - PostgreSQL
 - ORACLE
 - MySQL