**Informe sobre Base de Datos (BD): normativa, teoría, y evolución**

**Introducción** Las bases de datos (BD) son sistemas organizados que permiten almacenar, gestionar y recuperar información de manera eficiente. Se utilizan en casi todos los ámbitos, desde bancos hasta redes sociales, asegurando que los datos sean accesibles y seguros. En este informe, exploraremos las reglas de Codd, la normalización de bases de datos y el modelo relacional, junto con una línea de tiempo sobre la evolución de estas tecnologías.

**Las 12 Reglas de Codd** Edgar F. Codd estableció 12 reglas que definen los principios fundamentales de una base de datos relacional:

1. **Regla de la información**: Todos los datos deben almacenarse en tablas.
   * *Ejemplo:* Una tabla "Clientes" con columnas ID, Nombre y Teléfono.
2. **Regla del acceso garantizado**: Cada dato debe ser accesible vía combinación de tabla, columna y clave.
   * *Ejemplo:* Buscar un cliente por su ID.
3. **Tratamiento sistemático de valores nulos**: Los valores nulos deben manejarse de forma uniforme.
   * *Ejemplo:* Un campo "Teléfono" puede estar vacío sin afectar otros datos.
4. **Catálogo en línea basado en el modelo relacional**: La estructura de la base de datos debe almacenarse en tablas.
   * *Ejemplo:* La tabla "Esquema" almacena información sobre otras tablas.
5. **Regla del sublenguaje de datos**: Debe existir un lenguaje completo para manipular la base de datos (ej: SQL).
   * *Ejemplo:* Uso de SQL para consultar datos.
6. **Regla de actualización**: Cualquier vista derivada de otras tablas debe ser actualizable.
   * *Ejemplo:* Vista "Clientes\_Activos" basada en "Clientes".
7. **Alta, baja y modificación de vistas**: Las vistas deben poder modificarse igual que las tablas base.
   * *Ejemplo:* Modificar un cliente en "Clientes\_Activos" afecta "Clientes".
8. **Independencia física de los datos**: Cambios en el almacenamiento físico no afectan la estructura lógica.
   * *Ejemplo:* Migrar datos de disco a nube sin afectar consultas.
9. **Independencia lógica de los datos**: Cambios en la estructura no afectan la forma en que se acceden los datos.
   * *Ejemplo:* Agregar una columna "Email" sin afectar consultas existentes.
10. **Independencia de integridad**: Las reglas de integridad deben definirse en la BD y no en la aplicación.
    * *Ejemplo:* Restricción de "Edad >= 18" en la tabla "Clientes".
11. **Independencia de distribución**: Las consultas deben funcionar sin importar cómo esté distribuida la BD.
    * *Ejemplo:* Datos repartidos en varios servidores sin afectar la consulta.
12. **Regla de la no subversión**: No deben existir formas de acceder a los datos sin seguir las reglas de la BD.
    * *Ejemplo:* No se deben permitir accesos no controlados fuera de SQL.

**Normalización de Base de Datos** La normalización es un proceso que mejora la organización de los datos eliminando redundancias y asegurando integridad. Se realiza en varias formas normales (1FN, 2FN, 3FN, etc.).

Entre las formas más comúnmente aplicadas tenemos:

* **Primera Forma Normal (1FN):** Se logra eliminando grupos de datos repetitivos y asegurando que cada columna almacene un único valor atómico.
  + *Ejemplo:* Una tabla "Pedidos" con múltiples productos en una sola celda debe dividirse en otra tabla "DetallePedidos".
* **Segunda Forma Normal (2FN):** Se consigue eliminando dependencias parciales. Todas las columnas deben depender completamente de la clave primaria.
  + *Ejemplo:* En una tabla "DetallePedidos" (IDPedido, IDProducto, NombreProducto, Cantidad), "NombreProducto" depende solo de "IDProducto", por lo que debe moverse a una tabla "Productos".
* **Tercera Forma Normal (3FN):** Se logra eliminando dependencias transitivas. Ninguna columna debe depender de una columna que no sea clave primaria.
  + *Ejemplo:* Si una tabla "Empleados" almacena "IDDepartamento" y "NombreDepartamento", "NombreDepartamento" debe moverse a una tabla "Departamentos".

**Modelo Relacional** El modelo relacional organiza los datos en tablas con filas y columnas, permitiendo establecer relaciones entre ellas mediante claves primarias y foráneas. Cada tabla representa una entidad y las relaciones permiten estructurar y gestionar los datos de manera eficiente. Este modelo es el más utilizado en bases de datos modernas debido a su flexibilidad y facilidad para realizar consultas mediante SQL.

**Gestión de Bases de Datos** La gestión de bases de datos (DBMS) abarca todas las tareas necesarias para administrar y mantener una base de datos eficiente y segura. Esto incluye:

* **Seguridad:** Control de accesos y permisos para proteger la información.
* **Respaldo y recuperación:** Copias de seguridad periódicas para prevenir pérdidas de datos.
* **Optimización:** Indexación y estrategias de almacenamiento para mejorar el rendimiento.
* **Integridad de datos:** Garantizar que los datos sean precisos y consistentes mediante reglas de validación.

**Línea de Tiempo: Evolución de las Bases de Datos**

| **Fecha** | **Suceso** | **Personaje/Entidad** |
| --- | --- | --- |
| 1960 | Creación del modelo jerárquico | IBM |
| 1970 | Propuesta del modelo relacional | Edgar F. Codd |
| 1974 | Desarrollo de SQL | Donald Chamberlin y Raymond Boyce |
| 1979 | Primer sistema relacional comercial (Oracle) | Oracle Corp. |
| 1983 | Creación de PostgreSQL | Michael Stonebraker |
| 1995 | Popularización de MySQL | MySQL AB |
| 2000 | Introducción de NoSQL | Eric Brewer |
| 2009 | Aparición de MongoDB | 10gen (MongoDB Inc.) |
| 2012 | Big Data revoluciona las BD | Varios |
| 2023 | Avances en bases de datos en la nube | AWS, Google, Microsoft |

**Anécdotas Curiosas:**

1. IBM rechazó inicialmente el modelo relacional, pero luego se convirtió en líder en DB relacionales.
2. SQL inicialmente se llamaba SEQUEL, pero fue renombrado por problemas de marca.
3. PostgreSQL surgió de un proyecto para mejorar Ingres.
4. Google usa una BD interna llamada Spanner, que mantiene transacciones a nivel global.
5. Facebook almacena billones de datos en MySQL, optimizado con RocksDB.

**Conclusión** La evolución de las bases de datos, así como la estandarización de sus normativas, ha permitido el desarrollo de diversas tecnologías incluyendo a las inteligencias artificiales, ya que el almacenamiento, organización y acceso eficiente a grandes volúmenes de datos han sido fundamentales para entrenar y mejorar dichos modelos.