**BASE DE DATOS RELACIONALES**

Las bases de datos relacionales han transformado la manera en que se almacenan, organizan y gestionan los datos en sistemas informáticos modernos. Este informe aborda tres aspectos fundamentales: las 12 Reglas de Codd, la normalización y el modelo relacional, pilares que garantizan eficiencia, integridad y accesibilidad de la información.

En 1984 Edgar F. Codd, creador de del Modelo Relacional publicó las 12 Reglas que un verdadero Sistema Relacional de Bases de Datos debería cumplir. En la práctica algunas de estas reglas son difíciles de implementar, así que un sistema podrá considerarse más relacional cuanto más siga estas reglas. Son un estándar de calidad para diseñar y evaluar bases de datos.

1. Regla de la información: Todos los datos se representan como valores en tablas.
2. Regla del acceso garantizado: Cada valor en una tabla es accesible mediante la combinación de nombre de tabla, clave primaria y nombre de columna.
3. Tratamiento sistemático de valores nulos: Se permite la presencia de valores nulos para denotar datos desconocidos o inaplicables.
4. Catálogo dinámico en el nivel relacional: Los metadatos se almacenan en tablas accesibles mediante lenguaje relacional.
5. Regla del sublenguaje completo: El sistema debe soportar un lenguaje como SQL para definir, manipular y consultar datos.
6. Actualización de vistas: Las vistas derivadas deben ser actualizables lógicamente.
7. Inserción, actualización y eliminación de alto nivel: Estas operaciones deben poder ejecutarse en conjuntos de datos.
8. Independencia física de los datos: Los cambios en la estructura física de almacenamiento no afectan la forma en que los datos se acceden.
9. Independencia lógica de los datos: Cambios en la estructura lógica no requieren modificaciones en las aplicaciones.
10. Independencia de la integridad: Las restricciones de integridad se definen en el sistema relacional, no en las aplicaciones.
11. Independencia de la distribución: La distribución física de los datos no afecta su manejo lógico.
12. No subversión: Ninguna herramienta de bajo nivel puede evitar las restricciones de integridad.

La normalización es un método para estructurar los datos dentro de una base de datos. Este proceso implica crear tablas y definir relaciones entre ellas siguiendo principios diseñados para salvaguardar la integridad de la información y mejorar la flexibilidad del sistema, eliminando redundancias y dependencias inconsistentes. Esto se logra dividiendo tablas grandes en tablas más pequeñas y vinculándolas mediante relaciones. Las fases de normalización son:

1. Primera Forma Normal (1FN):

* Los datos deben ser atómicos, sin grupos repetitivos.
* Cada fila debe ser identificada por una clave primaria.

1. Segunda Forma Normal (2FN):

Cumple la 1FN.

Todas las columnas deben depender completamente de la clave primaria (eliminando dependencias parciales).

1. Tercera Forma Normal (3FN):

Cumple la 2FN.

No debe haber dependencias transitivas (una columna depende de otra no clave).

1. Forma Normal de Boyce-Codd (BCNF):

Refina la 3FN asegurándose de que todas las dependencias funcionales sean en clave candidata.

1. Cuarta y Quinta Forma Normal (4FN y 5FN):

Eliminan redundancias complejas, como dependencias multivaluadas y de unión.

Beneficios de la normalización:

Eliminación de redundancia: Evita duplicación de datos.

Consistencia e integridad: Previene anomalías en actualizaciones, eliminaciones o inserciones.

Eficiencia: Optimiza el uso del espacio de almacenamiento y facilita las consultas.

Respecto al Modelo Relacional para el Modelado y la Gestión de Bases de Datos. Es un modelo relacional, desarrollado por Codd, organiza los datos en tablas (relaciones) que representan entidades y sus atributos. Es el estándar más utilizado en bases de datos debido a su simplicidad y robustez.

* Componentes clave:

1. Tablas (Relaciones):

Las tablas están formadas por filas (tuplas) y columnas (atributos).

1. Claves:

* Clave primaria: Identifica de manera única cada fila en una tabla.
* Clave foránea: Enlaza tablas para representar relaciones.

1. Reglas de integridad:

* Integridad de entidad: La clave primaria no debe contener valores nulos.
* Integridad referencial: Las claves foráneas deben coincidir con valores existentes en la tabla referenciada.
* Operaciones relacionales:

El modelo relacional utiliza un conjunto de operaciones algebraicas para manipular los datos:

* Selección: Filtra filas específicas.
* Proyección: Muestra columnas seleccionadas.
* Unión: Combina filas de dos tablas.
* Intersección: Encuentra filas comunes entre tablas.
* Diferencia: Identifica filas presentes en una tabla pero no en otra.
* Ventajas del modelo relacional:
* Simplicidad: La estructura basada en tablas es intuitiva.
* Escalabilidad: Maneja grandes volúmenes de datos de manera eficiente.
* Flexibilidad: Permite la modificación y expansión de datos sin afectar las aplicaciones.
* Compatibilidad: Amplio soporte por sistemas como MySQL, PostgreSQL, y Oracle.

Línea de tiempo de los detalles:

**1960s**

Nacimiento de la informática y almacenamiento de datos:

Los sistemas de bases de datos jerárquicos y en red eran predominantes. Estos modelos eran complejos y carecían de flexibilidad.

**1970**

Introducción del modelo relacional (Edgar F. Codd):

Publicación del artículo seminal *"A Relational Model of Data for Large Shared Data Banks"*. Codd propone una forma estructurada de almacenar datos en tablas y manipularlos usando álgebra relacional.

**1971-1974**

Concepto de normalización de bases de datos:

Codd introduce la idea de formas normales para organizar datos y reducir redundancias.

Define las tres primeras formas normales: 1FN, 2FN y 3FN.

Más adelante, la Forma Normal de Boyce-Codd (BCNF) se establece como un refinamiento de la 3FN.

**1979**

Las 12 Reglas de Codd:

Codd formula estas reglas para definir qué requisitos debe cumplir un sistema de gestión de bases de datos relacional. Estas reglas sirven como guía para evaluar y diseñar sistemas relacionales.

**1980s**

Evolución del modelo relacional:

Comercialización de sistemas relacionales como Oracle (1979), IBM DB2 (1983), y Microsoft SQL Server (1989).

Los sistemas relacionales comienzan a dominar frente a los modelos jerárquicos y en red.

**1986-1987**

SQL como estándar:

El lenguaje SQL (Structured Query Language), basado en el modelo relacional, se estandariza por ANSI e ISO.

SQL se convierte en el lenguaje universal para interactuar con bases de datos relacionales.

**1990s**

Expansión de la normalización:

Se adoptan las formas normales más avanzadas (4FN y 5FN) para manejar relaciones más complejas y dependencias multivaluadas.

**2000s**

Avances en bases de datos relacionales:

Se integran características modernas, como la gestión de datos distribuidos y la optimización para grandes volúmenes de datos (Big Data).

Herramientas como PostgreSQL y MySQL ganan popularidad.

**Presente (2020s)**

Persistencia del modelo relacional:

A pesar del auge de bases de datos NoSQL para ciertos casos de uso, las bases relacionales siguen siendo el estándar para sistemas que requieren integridad y transacciones confiables.

Nuevas implementaciones como bases de datos en la nube (AWS RDS, Google Cloud SQL).

Modelos híbridos que combinan características relacionales y no relacionales.

En conclusión, el modelo relacional es fundamental en la gestión moderna de bases de datos. Las 12 Reglas de Codd proporcionan un marco conceptual, mientras que la normalización garantiza integridad y eficiencia. Comprender estos conceptos es esencial para diseñar sistemas de datos sólidos y escalables.