

(2)

bases de datos relacionales

(2.1) el modelo relacional

(2.1.1) introducción

Edgar Frank Codd definió las bases del modelo relacional a finales de los 60. En 1970 publica el documento *"A Relational Model of data for Large Shared Data Banks"* (*"Un modelo relacional de datos para grandes bancos de datos compartidos"*). Actualmente se considera que ese es uno de los documentos más influyentes de toda la historia de la informática. Lo es porque en él se definieron las bases del llamado **Modelo Relacional de Bases de Datos**. Anteriormente el único modelo teórico estandarizado era el **CodasyI** que se utilizó masivamente en los años 70 como paradigma del modelo en red de bases de datos.

Codd se apoya en los trabajos de los matemáticos **Cantor** y **Childs** (cuya teoría de conjuntos es la verdadera base del modelo relacional). Según Codd los datos se agrupan en **relaciones** (actualmente llamadas **tablas**) que es un concepto que se refiere a la estructura que aglutina datos referidos a una misma entidad de forma independiente respecto a su almacenamiento físico.

Lo que Codd intentaba fundamentalmente es evitar que las usuarias y usuarios de la base de datos tuvieran que verse obligadas a aprender los entresijos internos del sistema. Pretendía que los usuarios/as trabajaran de forma sencilla e independiente del funcionamiento físico de la base de datos en sí. Fue un enfoque revolucionario.

Aunque trabajaba para **IBM**, esta empresa no recibió de buen grado sus teorías (de hecho continuó trabajando en su modelo en red **IMS**). De hecho fueron otras empresas (en especial **Oracle**) las que implementaron sus teorías. Pocos años después el modelo se empezó a utilizar cada vez más, hasta finalmente ser el modelo de bases de datos más popular. Hoy en día casi todas las bases de datos siguen este modelo.

(2.1.2) objetivos

Codd perseguía estos objetivos con su modelo:

Independencia física. La forma de almacenar los datos, no debe influir en su manipulación lógica. Si la forma de almacenar los datos cambia, los usuarios no tienen siquiera porque percibirlo y seguirán trabajando de la misma forma con la base de datos. Esto permite que los usuarios y usuarias se concentren en qué quieren consultar en la base de datos y no en cómo está realizada la misma.

Independencia lógica. Las aplicaciones que utilizan la base de datos no deben ser modificadas porque se modifiquen elementos de la base de datos. Es decir, añadir, borrar y suprimir datos, no influye en las vistas de los usuarios. De una manera más precisa, gracias a esta independencia el esquema externo de la base de datos es realmente independiente del modelo lógico.

Flexibilidad. La base de datos ofrece fácilmente distintas vistas en función de los usuarios y aplicaciones.

Uniformidad. Las estructuras lógicas siempre tienen una única forma conceptual (las tablas).

Sencillez. Facilidad de manejo (algo cuestionable, pero ciertamente verdadero si comparamos con los sistemas gestores de bases de datos anteriores a este modelo).

(2.1.3) historia del modelo relacional

Año	Hecho
1970	Codd publica las bases del modelo relacional
1971-72	Primeros desarrollos teóricos
1973-78	Primeros prototipos de base de datos relacional. Son el System R de IBM. En ese sistema se desarrolla Sequel que con el tiempo cambiará su nombre a SQL.
1974	La Universidad de Berkeley desarrolla Ingres , SGBD relacional basado en cálculo relacional. Utilizaba el lenguaje Quel desarrollado en las universidades y muy popular en la época en ámbitos académicos.
1978	Aparece el lenguaje OBE (<i>Query By Example</i>) lenguaje de acceso relacional a los archivos VSAM de IBM
1979	Aparece Oracle , el primer SGBD comercial relacional (ganando en unas semanas al System/38 de IBM). Implementa SQL y se convertirá en el sistema gestor de bases de datos relacionales líder del mercado. Codd revisa su modelo relacional y lanza el modelo RM/T como un intento de subsanar sus deficiencias.
1981	Aparece Informix como SGBD relacional para Unix
1983	Aparece DB2 , el sistema gestor de bases de datos relacionales de IBM
1984	Aparece la base de datos Sybase que llegó a ser la segunda más popular (tras Oracle)

Año	Hecho
1986	ANSI normaliza el SQL (SQL/ANSI). SQL es ya de hecho el lenguaje principal de gestión de bases de datos relacionales.
1987	ISO también normaliza SQL. Es el SQL ISO(9075)
1988	La versión 6 de Oracle incorpora el lenguaje procedimental PL/SQL
1989	ISO revisa el estándar y publica el estándar SQL Addendum . Microsoft y Sybase desarrollan SQL Server para el sistema operativo OS/2 de Microsoft e IBM . Durante años Sybase y SQL Server fueron el mismo producto.
1990	Versión dos del modelo relacional (RM/V2) realizada por Codd. Propuesta de Michael Stonebraker para añadir al modelo relacional capacidades de orientación a objetos.
1992	ISO publica el estándar SQL 92 (todavía el más utilizado)
1995	Manifiesto de Darwen y Date en el que animan a reinterpretar el modelo relacional desde una perspectiva de objetos. Aparece el modelo objeto/relacional. Aparece MySQL una base de datos relacional de código abierto con licencia GNU que se hace muy popular entre los desarrolladores de páginas web.
1996	ANSI normaliza el lenguaje procedimental basado en SQL y lo llaman SQL/PSM . Permite técnicas propias de los lenguajes de programación estructurada. Aparece el SGBD abierto PostgreSQL como remodelación de la antigua Ingres, utilizando de forma nativa el lenguaje SQL (en lugar de Quel).
1999	ISO publica un nuevo estándar que incluye características más avanzadas. Se llama SQL 99 (también se le conoce como SQL 200)
2003	ISO publica el estándar SQL 2003 . En él se añade SQL/PSM al estándar.
2006	Estándar ISO. SQL 2006
2008	Estándar ISO. SQL 2008