

A Brief History of SQL

by
Frank Kalis, guest columnist, SQLServerCentral.com
edited and annotated by
Alan G. Labouseur

The original concept behind relational databases was first published by Edgar Frank Codd (an IBM researcher, commonly referred to as E.F.Codd) in a paper, "Derivability, Redundancy, and Consistency of Relations Stored in Large Data Banks" (RJ599), dated 08/19/1969. However, what is commonly viewed as the first milestone in the development of relational databases is a publication by Codd entitled "A Relational Model of Data for Large Shared Data Banks" in Communications of the ACM (Vol. 13, No. 6, June 1970, pp. 377-87). This was only a revised version of the 1969 paper. This article awoke massive public opinion in both the academic community and industry in the feasibility and usability of relational databases for commercial products.

Several other articles by Codd throughout the 1970s and 1980s are still viewed almost as gospel for relational database implementation. One of these articles is the famous so-called "12 rules" for relational databases, which was published in two parts in Computerworld. Part 1 was named "Is Your DBMS Really Relational?" (Published 10/14/1985); part 2 was called "Does Your DBMS Run by the Rules?" (10/21/1985). Codd continuously added new rules to these 12 originals and published them in his book "The Relational Model for Database Management, Version 2" (Addison-Wesley, 1990).

But to continue with the evolution of SQL and relational databases we must take a step back in time to the year 1974.

In 1974 Donald Chamberlin and others developed System R for IBM as a first prototype of a relational database. The Query Language was named SEQUEL (Structured English QUery Language).

System R also became part of IBM's prototype SEQUEL-XRM during 1974 and 1975. It was completely rewritten in 1976 – 1977. In addition there were new features like multi-table and multi-user capabilities implemented. The result of this revision was quickly named SEQUEL/2, but had to be renamed to SQL due to legal reasons because Hawker Siddeley Aircraft Company claimed the trademark SEQUEL for themselves.

In 1978 systematic tests were performed to prove real-world usability on customers' systems. It became a big success for IBM because this new system proved both useful and practical. With this result, IBM began to develop commercial products based on System R. SQL/DS came out in 1981; DB2 hit the streets in 1983.

Although IBM had done most of the research work, it was in fact a small and unknown software company named "Relational Software" to first release a RDBMS in 1979, two years before IBM. This unknown software company was later renamed Oracle. As an interesting side-note,

Relational Software released its product as Version 2. Obviously a brilliant marketing move: no one had to worry about a buggy and/or unstable Version 1 of this new kind of software product. [*Note from Alan*: Aston-Tate did this as well, releasing their first database product as “dBase II”.] The victory of SQL was well on its way.

Being the de facto standard, SQL became also an official standard through the American National Standards Institute (ANSI) certification in 1986 (X3.135-1986). But this standard could only be viewed as a cleaned-up version of DB2’s SQL dialect. Just one year later, in 1987, this was followed the standardization by the International Standards Organization (ISO 9075-1987). Only two years later, ANSI released a revised standard (X3.135-1989); so did ISO with ISO/IEC 9075:1989.

Partially due to the commercial interests of the software firms, many parts of the standard were left vague and unclear. This standard was viewed as a least common denominator and missed its intention. It was some 150 pages long.

To strengthen and establish the standard, ANSI revised SQL89 thoroughly, and released the SQL2 standard in 1992 (X3.135-1992). This time they did it right. Several weaknesses of SQL89 were eliminated. Further conceptual features were standardized, although at that time they were far beyond the possibilities of all relational databases. The new standard was some 500 pages long. But even today there is no single product available that fully complies with SQL92. Due to this disparity there were three level of conformity introduced:

1. Entry level conformance. Only small changes compared to SQL89.
2. Intermediate conformance. At least 50% of the standard has to be implemented.
3. Full conformance.

SQL99, also known as SQL3, was next. This standard addresses some of the modern, previously ignored features of certain modern SQL systems. There are object-relational database models, call-level interfaces, and integrity management. SQL99 replaces SQL92 levels of conformance with its own: Core SQL99 and Enhanced SQL99.

SQL99 is split into 5 parts:

1. Framework (SQL / Framework)
2. Foundation (SQL / Foundation)
3. Call Level Interface (SQL / CLI)
4. Persistent Stored Modules (SQL / PSM)
5. Host Languages Binding (SQL / Bindings)

Another impressive feature of SQL99 is the number of pages of this standard, some 2,000 pages long. [*Note from Alan*: Is this really a good thing, 2000 pages? I don’t think so; not at all.]

SQL has established itself as standard of database query language. So what’s next? Well, at a minimum SQL4 is due to be released in the next century.

There were 2003 and 2008 SQL Standard Releases. The most recent is SQL 2011.