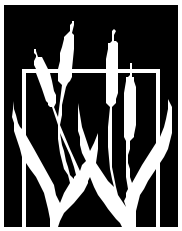

Whitemarsh
Information Systems Corporation

# Is SQL Really A Standard Anymore?

Whitemarsh Information Systems Corporation
2008 Althea Lane
Bowie, Maryland 20716
Tele: 301-249-1142
Email: mmgorman@wiscorp.com
Web: www.wiscorp.com

# Table of Contents



***Just what is a standard? It is something you can count on: for what it means, what it says, and what it does. It is from this simple and clear definition that I now question whether SQL is really a standard any more.***

## Background Note

Normally, I do not include my "official" standards title, that is, Secretary ANSI NCITS H2 Technical Committee on Database. However, in this case, I have created this article for TDAN with the special knowledge that comes to me because of my role as an officer of the ANSI (American National Standards Institute) NCITS (National Committee on Information Technology Standards) H2 Technical Committee on Database. I have been the committee's secretary since meeting "zero" in April 1978. I guess that means that I can view the entire past 23 years of data management standards efforts over its entire history.

There are a number of papers that supplement this article. They are all on the SQL page of the website, www.wiscorp.com. The SQL page can be reached from a link at the bottom of the home page. The link will take you to the Whitemarsh SQL page. There is a special section of papers that can be downloaded. These papers are all listed at the end of this paper.

## In the Beginning

Starting in 1978, the H2 committee was chartered to create a standard for the CODASYL (Committee on Data Systems and Languages) Network data model. The data definition language part of that effort was completed by the end of 1980. The data manipulation component of that standard was completed by the end of 1982. The standard then underwent "bureaucratic processing" and became an American National Standard in 1986.

The NDL standard was based on the CODASYL DDLC's (Data Definition Language Committee) work that was started in the early 1970s. The DDLC's effort was a offshoot of the database languages task group of the Codasyl Cobol Committee.

CODASYL as an organization, never made standards. Standards were done by the appropriate ANSI committee. Until the emergence of H2, ANSI committees only permitted themselves the right to subset CODASYL's work, never to invent or amend. The phrase, "CODASYL proposes and ANSI disposes" said it all. But H2, at its chartering meeting, gave itself the right to create specifications on its own. Thus, within three or so years, the CODASYL DDLC stopped all work.

In 1982, H2 was assigned the additional project of standardizing the relational data model. The effort went on for several meetings and finally, Phil Shaw of IBM suggested that the SQL/DS specification could be re-drafted and used as a base document. Phil Shaw redrafted the document and presented it to H2 for adoption as relational data language (RDL) base document. It was accepted.

During the next two years, RDL underwent substantial changes and it grew to quite a robust relational database language specification.

In 1984, however, it became quite obvious that the whole world was creating relational DBMSs that closely matched IBM's DB2. The H2 committee then decided to strip out a great deal of the RDL specification so that it could be readily implemented by the widest array of companies. By the end of 1984 that was accomplished. Then, the renamed document, SQL, underwent "bureaucratic processing" and it too became an American National Standard in 1986. It had two levels, Entry (level 1) and Full (level 2).

Simply put, the SQL86 standard was very simple. It was immediately implemented by a number of DBMS vendors. Because referential integrity was missing from SQL86, another version, SQL89, was quickly created and standardized. It too had two levels. Level 1 was largely SQL86 and Level 2 as the additional material for referential integrity.

In general, since it takes about two years for "bureaucratic processing," the technical work of SQL89 was really completed by the end of 1987. Work began immediately on a significant upgrade to the SQL89 standard. This work was completed about 1990 and was a standard by 1992. Hence, SQL92. Because SQL92 contained a significant quantity of new features, it too was divided into levels: Entry, Intermediate, and Full. The Entry level was essentially the full level of SQL89. Vendors immediately began implementation of SQL92 entry level**What Makes a Standard a Standard?**

Simple. Not implementation, but conformance. And, conformance is "known" only after conformance testing. Then and only then can users know with any degree of certainty that a vendor's product conforms to a standard.

There are two conformance testing models: The "Underwriters Laboratory (UL)" model and the "Consumers Union (CU)" model. Under the "UL" model, vendors pay an independent organization to do conformance testing. Vendors undergo testing for three main reasons: the public will only buy tested products, liability insurance companies demand it, wholesalers will only purchase tested products for resale to the end user. There is no such environment today for SQL.

Under the CU model, an independent organization, funded by consumers performs testing and reports the results of the tests to the buying public. They, in turn, use that information to make informed buying decisions. Again, there is no vendor who has decided to conduct conformance testing in the hopes of selling conformance test results to a buying public.

But, from the late 1980s through 1996 there was conformance testing. This was accomplished by the United States Government Department of Commerce's National Institute of Standards and Technology (NIST). NIST conducted the tests in support of public law that was originally known as the "Brooks Act," and later under other laws that were passed in the 1990s. The force behind the testing was that no Federal agency was able to buy a DBMS unless it passed conformance tests. Conformance meant the possibility of sales.

The benefits derived from the NIST conformance tests were well documented. A NIST commissioned study showed that there were about $35 million in savings from a program that only cost

about $600 thousand. But, in 1996, NIST started to dismantle its data management standards program. The publically stated reason was "costs." Obviously, that wasn't true.

Vendors, academics, industry, other Governments, and government agencies all expressed shock and dismay at NITS's data management standards dismantlement actions. Meetings were held with NIST, the Department of Commerce, and members of Congress. All were to no avail. The dismantlement was complete by the end of 1996.

In May of 1996, I wrote an article for the Outlook section of the Washington Post. It was unpublished as it was considered too technical. The key parts of the article were:

> "Because of NIST's FY-97 and beyond plans, SQL's conformance tests and certifications, that is, those beyond the SQL shell will be left to the ANSI/SQL vendors. They however have no motivation whatsoever to perform full and complete testing nor self policing. Only the largest buyer has that motivation, and in the case of ANSI/SQL the largest buyer is the United States Government.
>
> "Simply put, without robust test development and conformance testing by NIST, DBMS will return to the days of vendor specific, conflicting features and facilities that will lock Federal agencies into one vendor, or make DBMS frightfully expensive acquire, use, and dislodge.
>
> "If vendors did believe in level playing fields, highly interchangeable databases and application programs, they would all be protesting NIST's abandonment of ANSI/SQL certification and conformance testing. NIST's conformance tests and certifications leveled the playing field. Is it mere conincidence that the two biggest DBMS vendors (75% market share) are silently sitting on the side lines? Could they just be biding their time till NIST dismantles its ANSI/SQL certification and conformance testing program before bringing their vendor specific and unique features to market?
>
> "Once these new vendor-locking features are deployed through aggressive marketing campaigns, Agencies will be unable to know they are unique because NIST won't have the conformance tests and certifications for these newly developed features. Agencies will then either be condemned to the unsophisticated data management of the 1980s or be sucked into using the new sophisticated, but vendor locking data management features of the future. No longer will the playing field be level. Rather, only the few very big vendors will remain. Competition will drop, prices will rise, and portability of programs, data and trained staff will end."

Five full years have passed since that article was sent to The Washington Post. The key predictions of the article were:

- Fewer vendors
- Higher costs
- More features implemented in a proprietary way

## Fewer vendors

As to the first prediction, the list of vendor members of H2 in December 1994 is presented in Table 1. The second column of that table are indications as to whether the vendor is still a member of H2. It can be inferred that these vendors are still interested in pursuing SQL standards work. While it certainly cannot be stated that dropping membership was due to the elimination of SQL testing, two observations can be made. First, there is no longer a well defined practically implementable subset of SQL features that can be proven via conformance tests. Second, because there was no outside "buyer" to shape the content of the Core level of SQL99, it was enlarged to such an extent that to implement it all is close to impossible for all vendors except for two or three. In short, the size of core is a natural barrier to practical product development.

## Higher costs

As to the second prediction, higher costs, it was hoped that there would be sufficient cost data available to identify a general DBMS product configuration, and then to show the rate of increase in prices over the key vendors during the 1990s and into 2001. The expectation was that with the elimination of conformance tests and with the emergency of proprietary feature sets that vendors would be able to charge greater prices for their products during the second half of the 1990s than for the first half. However, it seems that nobody keeps paper any more. I called a number of organizations and Government agencies. None seem to keep records earlier than "yesterday." So, I cannot report at all on the second prediction.

In retrospect, comparing the rate in increase in costs is probably not very helpful in any event given the rise in personnel costs. If a DBMS feature cuts personnel time by 25% and if there are 50 professional staff, and if the cost of a staff over the 12 year period ranged from $100K to $200K fully burdened, then the cost increase in DBMSs would have to be very significant to overshadow the rise in prices. The staff cost is $90 million. If a DBMS caused a 25% savings, that is, $22.5 million but the cost of the supporting DBMS ranged from $50K per year to $100K per year over that time frame, then staying with the DBMS and forgoing the productivity increase would be a bad decision indeed as

the DBMS increased cost was only $600K over the time while the potential staff savings would have been $22.5 million.

## More Features Implemented in a Proprietary Way

As to the third prediction, more features implemented in a proprietary way. This is best understood through an exposition of the features in SQL86, SQL89, and then SQL92. Table 2 lists those features. SQL92, the third column, was itself divided into three levels: Entry, Intermediate, and Full. Since virtually all the vendors applied for and were tested only at SQL92 Entry level, conformance at any level above that is unknown. In column 3, the features of SQL92 entry level are listed at the top of the column. All the untested features are listed in the rest of the column.

SQL99 is the newest SQL standard. It contains a great deal of features well beyond SQL92. Table 3 lists the names of the key features. SQL99 is actually a family of SQL related standards. There is the foundation standard and then there are the different parts that define wholly contained additional collections of features. These features are fully defined in the SQL documents listed in the references at the end of this paper.

Since the acceptance of SQL99 as a world wide data management standard, vendors have been hard at work implementing its various aspects and features outside the standard that are of interest to their customer base. Since the key to data interoperability is the existence of standards conforming DBMSs, a survey was conducted in the first part of 2001 as to the intentions of the SQL vendors regarding conformance to the SQL99 standard. Table 5 presents the results of that survey.

From Table 4, it can be clearly seen that across the first 11 questions, the majority of vendors are implementing in each area. Regrettable, however, this table also clearly shows that the vast majority of the vendors are not intending to implement these new features in a standards compliant manner

## Summary and Conclusions

From the predictions made in May 1996, clearly the count of vendors has dramatically reduced and the most modern of feature set from the remaining SQL vendor's products are untested. Neither the predictions nor the fulfillment of those predictions should come as a surprise. It is normal behavior on the part of vendors attempting to gain a wider and wider market share in a proprietary manner.

The only organization that was capable of preventing this proprietary consequence was NIST. NIST chose for totally unsupportable to stop conformance test development and conformance testing. Not only is this unfortunate for consumers at large and for "Corporate America," it is especially unfortunate for Federal Government Agencies who are mandated by Congress to have interoperable systems. The only way they can achieve this requirement is to choose one vendor for the entire agency.

This would lead to ridiculous situations like the Sybase Army, the IBM Navy, the Oracle EPA, and the Microsoft Justice Department (wouldn't that be interesting). Or maybe, the Oracle Federal Government. Then there of course is all the State and Local Governments that are similarly affected.

A key question must therefore be asked: *Is SQL a Real Standard Any More*? And, if by standard it is meant that there exists the three essential components of any successful standard,

- A significant market share vendor community
- Publically developed, available, and evolving standards
- Enforced conformance tests

Then while the first two exist, the third does not. Hence SQL no longer represents a successful information technology standard. SQL could, however again achieve the status of a successful standard if one or more large organizations with buying-clout defined conformance tests, conducted conformance testing, and then only purchased conforming products.

## Related Readings

Note: These materials may be obtained directly through the links, or from the Whitemarsh website (http://www.wiscorp.com/documentsdownloadsfree.html). There are also a large number of SQL related documents from the Whitemarsh SQL page, http://www.wiscorp.com/sql99.html.

| Related Readings | | |
|---|---|---|
| **Nbr** | **Title** | **Description** |
| 1. | The Role of NIST in Data Standardization (http://www.tdan.com/i002fe04.htm) | This paper was published on TDAN in 1997. The paper is a position paper written by key members of the database standards community and was approved by the key standards Vps of IBM, Sybase, Informix, and Oracle. This paper was delivered to NIST, and key members of Congress. |
| 2. | Great News, The Relational Model is Dead (http://www.tdan.com/i008hy02.htm) | This paper, published on TDAN outlined the changes to the SQL99 standard that, in my mind, clearly make it not relational. |

| Related Readings | | |
|---|---|---|
| **Nbr** | **Title** | **Description** |
| 3. | SQL 1999 Impact On Data Administration | This presentation reviews the key features of SQL99 and the effect these key features have on database design, data administration, database administration, and database application development. |
| 4. | SQL 1999 | This presentation is a brief tutorial on the features of SQL99. Many of the examples contained in the presentation were taken from presentations created by IBM and Oracle that are also on the Whitemarsh website. |
| 5. | Essential Paradigm for IT Standards Success | This paper reviews the key components that are required for successful IT standards. When these requirements exist, there is standards success. When one or more of the requirements are missing, the IT standard is unsuccessful. |
| 6. | Supporting Public Law for Data Management Standards | This very brief document summarizes the public law that supports IT standards within the Federal Government. Adhering to these public laws clearly requires data-based interoperability, which in turn requires standards conforming DBMSs that can only be proven through conformance tests and conformance testing. |
| 7. | Chronological Listing of Key Documents Related to NIST and Data Management Standards | This long document contains a chronological listing of and abstracts of over 140 documents, reports, letters, E-mails, related to data managements standards, conformance tests, and conformance testing. |
| 8. | September 1997 Presentation to NCR DAMA, entitled, "NIST Guts Data Management Standards Infrastructure" | This presentation was delivered to the National Capital Region chapter of DAMA. The presentation reviews the requirements for successful IT standards, the role for a key agency like NIST, the role that NIST played for the 20 years prior to 1996, and the effects of NIST's actions that dismantled data managements standards conformance test development and conformance testing. |

**Table 1. List of SQL Vendors Belong to ANSI NCITS H2 in 1994 and in 2001**

| Vendors | ANSI NCITS H2 Membership | |
|---|---|---|
| | December 1994 | April 2001 |
| Bull/HN | Yes | No |
| CINCOM Systems | Yes | No |
| Computer Associates | Yes | Yes |
| Digital Equipment Corporation | Yes | No |
| Hewlett Packard | Yes | No |
| Information Dimensions | Yes | No |
| Informix, Inc. | Yes | Yes |
| International Business Machines | Yes | Yes |
| InterSystems | Yes | Yes |
| Microsoft Corporation | Yes | Yes |
| Oracle Corporation | Yes | Yes |
| Progress Software Corporation | Yes | No |
| Sybase | Yes | Yes |
| Tandem Computers | Yes | No |
| The ASK Group | Yes | No |
| UniSQL | Yes | No |
| Unisys Corporation | Yes | No |
| Watcom | Yes | No |

**Table 2. Features from SQL:1986, -89, and -92**

| **SQL/1986** | **SQL/1989** | **SQL/1992** |
|---|---|---|
| Basic features, that is | SQL/1986 plus | SQL/1989 plus |
| Tables<br>Columns<br>Views<br>Basic relational operations<br>Some integrity constraints<br>Language bindings to<br>COBOL, FORTRAN, C, etc. | Partial Referential Integrity | **Entry Level SQL92**<br>Enhanced constraints<br>Full Referential Integrity<br>Table constraints<br>Schema manipulation<br>Row & Table constraints<br><br>**Beyond Entry Level SQL92**<br>Assertions<br>Bit data type<br>CASE<br>Character Sets<br>Connection Management<br>DATETIME<br>Domains<br>Dynamic SQL<br>Get Diagnostics<br>Grouped operations<br>Information Schema<br>Multiple module support<br>National character sets<br>Natural joins (inner & outer)<br>Subqueries in check clauses<br>Temporary tables<br>Transaction Management<br>Union and intersect |

## Table 3, Key SQL:1999 New Features and Brief Descriptions

| SQL 1999 Feature | Brief Description |
|---|---|
| SQL Data Type: String (BLOB, or Character (CLOB)) | The ability to store either bit images or large character documents |
| SQL Data Type: Boolean | The ability to specify boolean data types, logic, and supporting rules |
| SQL Data Type: Ref Types | The ability to have a DBMS generated or column value based pointer as a reference between rows of different tables. |
| SQL Data Type: Arrays | The ability to have an ordered list of values within a column. Each value may be a RefType. Each may also be a ROW data structure |
| SQL Data Type: ROW Data Structures | The ability to have groups of "subcolumns" within a column. Each may be an array or a RefType |
| SQL Data Type: User Defined Types | The ability to completely define an non-traditional data type such as nautical distance. |
| Triggers | The ability to specify the instigation of an action as a consequence of a state change in the database |
| Savepoints | The ability to have cascading sets of "soft" commits that can be rolled back until there is a traditional hard-commit |
| Roles Security Enhancement | The ability to define additional layers and kinds of security and the assignment of persons fulfilling the defined roles |
| Recursion | The ability to fully model nested relationships such as hierarchies for organizations. |
| Information Schema | A virtual database defined as virtual tables and real SQL views on the virtual tables that contain the complete set of metadata in support of defined databases. |

| SQL 1999 Feature | Brief Description |
|---|---|
| Call Level Interface | The complete specification of a DBMS vendor independent set of database access routines similar to that contained in the Microsoft ODBC specification. |
| SQL Multi-Media: Full Text | The complete set of data structures, special full-text operations and SQL routines that support the loading, accessing, and maintenance of full-text type of data such as books, manuscripts. |
| SQL Multi-Media: Spatial | The complete library or set of data structures and routines that support spatial data types and operations on those data types |
| SQL Multi-Media: Still Image | The complete library or set of data structures and routines that support still image data types and operations on those data types |
| SQL Programming Language | A complete SQL DBMS encapsulated programming language that includes traditional assignment, looping, branching, If..Then...Else, and CASE type constructs. |
| Transaction, Connection, Session, and Diagnostics Management | The ability to specify sessions and the management of those sessions in support of centralized or distributed type batch-type processing. |
| SQL/MED | The routines and facilities in support of the management of types and classes of data that exists outside the domain of SQL. |

## Table 4. March 2000 Implementation Directions of SQL Vendors for SQL:1999

| Feature that through standard or non-standard ways accomplishes: | Percent Distribution Across SQL Vendors | | | |
|---|---|---|---|---|
| | 1. Have Already Imple-mented | 2. Currently Imple-menting | 3. Investi-gating | 4. Definitely Not Imple-menting |
| Nested data structures within columns (e.g., array, Ref Types, ROW, UDTs, et al) | 100 | | | |
| BLOB, CLOB | 100 | | | |
| Methods within Columns | 90 | 5 | 5 | |
| Triggers | 100 | | | |
| Savepoints | 90 | 5 | 5 | |
| Recursion | 35 | | 55 | 10 |
| SQL/CLI | 90 | | 10 | |
| SQL/PSM | 90 | 10 | | |
| SQL/Full Text | 80 | | | 20 |
| SQL/Spatial | 55 | | 15 | 30 |
| SQL/MED | 55 | 15 | 15 | 15 |
| Export data through XML | 40 | 15 | 30 | 15 |
| Export schema et al through XML | 40 | 10 | 40 | 10 |

SQL Vendors on the ANSI NCITS H2 Technical Committee on Database: Compaq, Computer Associates, Informix, Corel, International Business Machines, Intersystems, Microsoft, NCR, Oracle, Pervasive, Progress, and Sybase

## Table 5. March 2001 Conformance Intentions of SQL Vendors for SQL:1999

| Question | Percent Distribution of Vendors to Questions and Standards Compliance Sub-questions | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | | | 2 | | | | 3 | | | | 4 | | | |
| | A | B | C | D | A | B | C | D | A | B | C | D | A | B | C | D |
| 1. Nested data structures within columns | 22 | 11 | 22 | | 11 | | 22 | | 11 | | | | | | | |
| 2. BLOB, CLOB | 11 | 22 | 33 | | 11 | 11 | 11 | | | | | | | | | |
| 3. Methods within Columns | 22 | 11 | 11 | | 11 | | 33 | | 11 | | | | | | | |
| 4. Triggers | | 44 | 44 | | | | | | 11 | | | | | | | |
| 5. Savepoints | 22 | 22 | 22 | | | | | | | 11 | 11 | | 11 | | | |
| 6. Recursion | 11 | 11 | | | 11 | | | | 11 | 22 | 22 | | | | 11 | |
| 7. SQL/CLI | 22 | 33 | 11 | | | | | | 11 | 11 | | | | | 11 | |
| 8. SQL/PSM | 11 | 11 | 44 | | | | | | 11 | 11 | | | | | 11 | |
| 9. SQL/Full Text | 11 | 11 | 22 | | | | | | 11 | 11 | 11 | | | | 22 | |
| 10. SQL/Spatial | 11 | 11 | 11 | | | | | | 11 | 11 | 22 | | | | 22 | |

| Question | Percent Distribution of Vendors to Questions and Standards Compliance Sub-questions | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | | | 2 | | | | 3 | | | | 4 | | | |
| | A | B | C | D | A | B | C | D | A | B | C | D | A | B | C | D |
| 11. SQL/MED | | 11 | 11 | | | | 11 | | 11 | 11 | 22 | | | | 22 | |

| Question | Percent Distribution of Vendors to Questions and Standards Compliance Sub-questions | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | | | 2 | | | | 3 | | | | 4 | | | |
| | A | B | C | D | A | B | C | D | A | B | C | D | A | B | C | D |
| 12. Export data through XML | | | 11 | 33 | 11 | | | 33 | 11 | | | | | | | |
| 13. Import data through formatted XML | | | 11 | 44 | | | | 22 | 11 | | | 11 | | | | |
| 14. Import schema DDL through XML | | | 11 | 33 | | | | 33 | 11 | | | 11 | | | | |
| 15. Export schema DDL et al through XML | | | 11 | 33 | | | | 44 | 11 | | | | | | | |
| 16. SQL/J support. E.g., JDBC | 33 | 33 | | | | | | | 11 | | 11 | | | | 11 | |
| Summary Percentages | 11 | 14 | 17 | 9 | 3 | 1 | 5 | 8 | 10 | 6 | 6 | 1 | 1 | | 7 | |
| | 25 | | 26 | | 4 | | 13 | | 16 | | 7 | | 1 | | 7 | |
| | 51 | | | | 17 | | | | 23 | | | | 8 | | | |

**Implementation Answer Key**

| | |
|---|---|
| 1 | Have already implemented one or more aspects of "area," |
| 2 | Implementing one more aspects of "area", |
| 3 | Investigating but have not decided, |
| 4 | Definitely NOT implementing. |

**Standards Compliance Answer Key**

| | |
|---|---|
| A | Intend to comply with both the syntax and semantics of SQL1999 |
| B | Will largely comply with both the syntax and semantics of SQL1999 |
| C | Do not know, at this time, if our implementation will comply with the syntax and semantics of SQL1999 |
| D | No standard currently exists in this area or feature. |

| | |
|---|---|
| **SQL Vendors on ANSI NCITS H2 Technical Committee on Database** | IBM, Oracle, Sybase, Informix, NCR, Computer Associates, Microsoft, FileTek, InterSystems, Compaq, Pervasive |