



ISO
International Organization for Standardization

Authoritative Version: Adobe Acrobat Portable Document Format (PDF)



ANSI
American National Standards Institute

ANSI TC X3H2
Database
ISO/IEC JTC 1/SC 32
Data Management and Interchange
WG 3
Database Languages

Project: ANSI: 1234D — ISO: 1.32.3.4

Title: Subproject: "XML-Related Specs (SQL/XML)"

Status: For approval by NCITS TC H2 and subsequent submission to ISO/IEC JTC 1/SC 32

Author: Jim Melton

Contact: Jim Melton (jim.melton@acm.org)

Abstract: XML is rapidly becoming one of the most promising technologies for the exchange of data on the Internet and the World-Wide Web. SQL has for many years been the language of choice for managing traditional ("object-relational") data. It is inevitable that facilities will arise for using these two technologies together.

References:

- 1) [XML] <http://www.w3.org/TR/REC-xml>, *Extensible Markup Language 1.0 (REC-xml-19980210)*, 10 February, 1998.
- 2) [XML-QDM] <http://www.w3.org/TR/query-datamodel/>, *XML Query Data Model*, W3C Working Draft, 11 May, 2000.
- 3) [XML-QR] <http://www.w3.org/TR/xmlquery-req>, *XML Query Requirements*, W3C Working Draft, 31 January, 2000. (Distributed as H2-2000-024.)
- 4) [XML-des] H2-2000-184. *Database Event Stream (DES) in XML*, 18 March, 2000.

- 5) [XML-esd] H2-2000-185, *Event Stream Delivery (ESD) protocol*, 18 March, 2000.
- 6) [XML-rs] H2-2000-186, *A JDBC-oriented XML DTD for SQL result sets*, 18 March, 2000
- 7) [XML-schema/primer] <http://www.w3.org/TR/xmlschema-0/>, *XML Schema Part 0: Primer*, W3C Working Draft, 7 April, 2000. (Distributed as H2-2000-222.)
- 8) [XML-schema/struct] <http://www.w3.org/TR/xmlschema-1/>, *XML Schema Part 1: Structures*, W3C Working Draft, 7 April, 2000. (Distributed as H2-2000-223.)
- 9) [XML-schema/types] <http://www.w3.org/TR/xmlschema-2/>, *XML Schema Part 2: Datatypes*, W3C Working Draft, 7 April, 2000. (Distributed as H2-2000-224.)
- 10) [XML-namespaces] <http://www.w3.org/TR/REC-xml-names>, *Namespaces in XML*, W3C Recommendation, 14 January, 1999.
- 11) [XML-info] <http://www.w3.org/TR/xml-info>, *XML information Set*, W3C Working Draft, 26 July, 2000.
- 12) [XML-path] <http://www.w3.org/TR/xpath>, *XML Path Language (XPath) Version 1.0*, W3C Recommendation, 16 November, 1999.
- 13) [MED] FDIS 9075-10, *Information technology — Database languages — SQL — Part 10: Management of External Data (SQL/MED)*, September, 2000.
- 14) [Plan] ISO/IEC JTC1/SC21 N8091, *Elaboration of the Scope and Content of Project 1.21.3.4 (SQL3)*, dated 30 June, 1993, also distributed as document ISO/IEC JTC1/SC21/WG3 N1455, dated 20 November, 1992.
- 15) [SQL:1999] ISO/IEC 9075-1:1999, *Information technology — Database languages — SQL — Part 1: Framework (SQL/Framework)*, ISO/IEC 9075-2:1999, *Information technology — Database languages — SQL — Part 2: Foundation (SQL/Foundation)*, ISO/IEC 9075-3:1999, *Information technology — Database languages — SQL — Part 3: Call-Level Interface (SQL/CLI)*, ISO/IEC 9075-4:1999, *Information technology — Database languages — SQL — Part 4: Persistent Stored Modules (SQL/PSM)*, ISO/IEC 9075-5:1999, *Information technology — Database languages — SQL — Part 5: Host Language Bindings (SQL/Bindings)*, November, 1999.
- 16) [OLB:2000] ISO/IEC 9075-10:1999, *Information technology — Database languages — SQL — Part 10: Object Language Bindings (SQL/OLB)*, to be published (in 2000)

1. Discussion

1.1. SQL Standard Status

[SQL:1999] was published late in 1999, containing five parts of the multi-part standard being continuously developed under the project plan documented in [Plan]. Other parts of this multi-part standard continue to evolve. More recently, technical development and editing for [OLB:2000] was completed, with publication expected shortly. Other components, including [MED] are expected to progress before the end of 2000.

2. Needs

The SQL Project, ISO project 1.32.3.4, is documented in [Plan]; although the project plan has not been revised in some time, it remains valid and active. The project plan already authorizes further development of Database Language SQL, including creation of new subprojects and new parts as needed.

SQL continues to be arguably the most successful and important database language standard ever developed and forms the foundations of a multi-billion (US) dollar data management industry as well as vital infrastructure for many more billions of (US) dollars of commerce worldwide. SQL's success has been due in part to its relationship to the relational model of data, but also to the flexibility of the language in meeting varied business needs and its amenability to performance optimization.

In recent years, the success of the World-Wide Web has led to requirements for better management of data that does not lend itself as well to relational structures and operations than SQL has traditionally offered. [XML] has become one of the most important new technologies for representing much data on the Web. The World-Wide Web Consortium (W3C) has established formal processes for developing Recommendations — *de facto* standards — based on XML and related technologies. [XML-schema/primer], [XML-schema/struct], and [XML-schema/types] are the most recent public W3C documents directed at defining XML language with strong data typing. [XML-QDM] and [XML-QR] are the most recent public W3C documents associated with the work in progress to develop an XML query capability. In short, XML represents nothing short of a paradigm shift in data communications and management; like every paradigm shift, this one has the potential to greatly enhance the businesses that depend on data management and communications, as well as the ability to marginalize existing business segments and open new opportunities for new players.

One of the most intriguing and urgent requirements to arise from the appearance of XML is a well-defined relationship between XML and SQL. Vast quantities of business data are currently stored in SQL database systems and great demand exists for the ability to present that data in XML form to various client applications. By contrast, increasing amounts of less-traditional data ("documents") are being produced in XML formats and there is tremendous pressure to allow that data to be queried concurrently with traditional ("object-relational") data. In addition, the growing need to allow disparate systems to exchange data has caused significant attention to be paid to the use of XML as a "canonical data format" between such systems (e.g., on the Web).

In short, business, industry, and society will find ways to use XML and SQL together. Standards in this area will eventually emerge, but a focussed effort to develop such standards will have significant benefits for all parties involved. The W3C is appropriately focussing on XML-centered standards, including a Schema facility and a Query facility. It is correspondingly appropriate for the SQL standards world to address the relationship between XML and SQL, preferably in cooperation with the W3C and not in competition with it.

3. Existing Practice

In mid-2000, enterprises using SQL and XML together do so with a highly eclectic and *ad hoc* mixture of technologies, virtually all proprietary in nature. The W3C's work on XML Schema and XML Query is proceeding at a very rapid rate, but the requirements of business have encouraged early adoption of stop-gap measures (such as on-going use of Document Type Descriptions, or DTDs, and several XML-oriented query languages, such as XQL, XML-QL, and Quilt).

Particularly in the area of representing SQL data in XML format, there are no known *de facto* standards; instead, virtually every user of SQL data in an XML context has been creating new DTDs (or other metadata representations) for each specific purpose.

4. Expected Stability

While it can be argued that the very newness of XML and its necessary relationship with SQL dictates against attempts to standardize technologies, the extremely rapid rate of adoption of XML speaks for such standardization. The commercial community needs standards in the area of XML and SQL together in order to maximize benefits of both technologies and such standards will either emerge to the benefit of all concerned or they will emerge far more chaotically with accompanying difficulties.

The SQL standards community has a unique opportunity to work on standards directly associated with the relationship between XML and SQL, and must do so alongside the other recognized force in the XML standards area, the W3C.

5. Program of Work

Create a new subproject and a new part of the SQL standard, provisionally to be named *Information technology — Database languages — SQL — Part (number to be assigned): XML-Related Specifications (SQL/XML)*. This new part of SQL may, but will not necessarily, contain:

- Specifications for the representation of SQL data (specifically rows and tables of rows, as well as views and query results) in XML form, and *vice versa*.
- Specifications associated with mapping SQL schemata to and from XML schemata.
- Specifications for the representation of SQL Schemas in XML.
- Specifications for the representation of SQL actions (insert, update, delete).
- Specifications for protocols related to the transport of XML when used with SQL.
- Specifications of the (perhaps "a") manner in which SQL language can be used with XML.

All efforts will be made to ensure that the specifications in this part will neither duplicate nor conflict with W3C specifications.

6. Justification for Subproject Request

Various criteria have been specified for the approval of program extensions (that is, subdivisions and minor enhancements) of existing projects. Those criteria relative to the subdivision of the existing SQL project [Plan] are satisfied as followed:

1. The rationale for the SQL/XML part of Database Language SQL is given in the "Needs" paragraph above. This proposed work is within the scope of already-authorized work in the SQL project description. [Plan] support for work of this nature is embedded in statements such as the following:

- “[Database Language SQL] will consider support for...other high-level tools for modern information management” (Description, 2nd paragraph)
 - “Specific candidate features include, but are not limited to, the following:”, followed by “distributed data support;” and “interfaces to additional programming languages;” (Description, 3rd paragraph and following list)
 - “[SQL] will specify a...language for the...query of persistent complex objects” (Description, 2nd paragraph)
2. Consensus on the need to do this work is evident from discussions held at the ISO/IEC JTC1/SC32/WG3 meeting held in July, 2000, at Warwick, GBR, and from previous discussions within WG3 and SC32 generally.
 3. A candidate base document does not exist. It is inappropriate and presumptuous at this stage to provide such a candidate base document, since the contents of that document must be agreed as part of the development process. It is anticipated that the material to be accepted for a candidate base document will be generated as work is initiated and progresses.
 4. The SQL editor, Mr. Jim Melton, is willing to be the subproject editor for *Information technology — Database languages — SQL — Part (number to be assigned): XML-Related Specifications (SQL/XML)*.

End of paper