

# SQL Standards Status and Directions

September 2009  
Sydney Rdb Technical Forum

**Keith W. Hare**  
JCC Consulting, Inc.  
600 Newark Road, P.O. Box 381  
Granville, OH 43023 USA  
+1.740.587.0157  
[Keith@jcc.com](mailto:Keith@jcc.com)

Copyright 2009, JCC Consulting, Inc.  
All rights reserved

# Abstract

While the SQL Standard has not been at the forefront of the computing press, it is still progressing and expanding.

- What is SQL?
- What distinguishes SQL?
- ISO and ANSI
- Current Status
- Directions

# Who Am I?

- Muskingum College, 1980, BS in Biology and Computer Science
- Senior Consultant with JCC Consulting, Inc. since 1985 – high performance database systems
- Ohio State – Masters in Computer & Information Science, 1985
- SQL Standards committees since 1988
- Vice Chair, INCITS H2 since 2003
- Convenor, ISO/IEC JTC1 SC32 WG3 since 2005

# What is SQL?

SQL is a language for defining data bases and manipulating the data in those data bases

- SQL Standard uses SQL as a name, not an acronym
  - Might stand for **S**tandard **Q**uery **L**anguage
  - Might stand for **S**QL **Q**uery **L**anguage
- SQL queries are independent of how the data is actually stored – specify what data you want, not how to get it



# What is a Data Base?

- A data base is a way of storing and retrieving data in a coherent manner
- Hierarchical – early 1960s – IBM's IMS
- Network – late 1960s
- Relational – 1970s – System R, Ingres, Oracle
- SQL – 1980s – DB2, Ingres, Oracle, Sybase, Microsoft, MySQL...

# SQL Data Bases

- Data stored in columns and tables
- Relationships represented by data
- Data Manipulation Language
- Data Definition Language
- Transactions

# Data Manipulation Language (DML)

- Data manipulated with Select, Insert, Update, & Delete statements
- Select T1.Column1, T2.Column2 ...  
From Table1, Table2 ...  
Where T1.Column1 = T2.Column1 ...
- Data Aggregation
- Compound statements
- Functions and Procedures

# Data Definition Language

- Create Table (Column1 Datatype1, Column2 Datatype 2, ...)
- Constraints to define and enforce relationships
  - Primary Key
  - Foreign Key
  - Etc.
- Triggers to respond to Insert, Update , & Delete
- Stored Modules
- Alter ...
- Drop ...
- Security and Access Control

# Transactions – ACID Properties

- **Atomic** – All of the work in a transaction completes (commit) or none of it completes
- **Consistent** – A transaction transforms the database from one consistent state to another consistent state. Consistency is defined in terms of constraints.
- **Isolated** – The results of any changes made during a transaction are not visible until the transaction has committed.
- **Durable** – The results of a committed transaction survive failures

# Other Data Base Features

The SQL Standard specifies the logical data definition and manipulation, not the physical data structures.

- Does not specify how to define:
  - Storage Areas/Files/Table Spaces
  - Indexes
  - Query Optimization strategies
  - Transaction Journals/Logs
  - Backup & Restore
  - Etc.

# What Distinguishes SQL?

Several factors distinguish SQL from other query and data storage technologies:

- Persistent data accumulated over long periods
- Complex update and read-only transactions
- Concurrent update and read-only access
- Null values and three-valued logic
- Integration with other technologies



# ISO and ANSI

In the international arena, the SQL Standard is developed by ISO/IEC JTC1 SC32 WG3.

■ Officers:

- Convenor – Keith W. Hare – USA
- Editor – Jim Melton – USA

■ Active participants are:

- Australia – Australian Bureau of Statistics
- Canada – Standards Council of Canada
- China – Chinese Electronics Standardization Institute
- Germany – DIN Deutsches Institut für Normung e. V.
- Great Britain – British Standards Institution
- Japan – SQL working group of JIS (Japan Industrial Standards)
- Netherlands
- USA – ANSI INCITS H2



# Translation

- ISO/IEC
  - ISO – International Organization for Standardization
  - IEC – International Electrotechnical Commission
- JTC1 – Joint Technical Committee 1
  - Information Technology
  - There is no JTC2
- SC32 – SubCommittee 32
- WG3 – Working Group 3 – the committee responsible for data base language standards

# USA – ANSI INCITS H2

## ■ Officers

- Chair: Don Deutsch – Oracle
- Vice Chair: Keith Hare – JCC Consulting, Inc.
- Secretary – Mike Gorman – Whitemarsh
- International Representative – Krishna Kulkarni – IBM
- Editor – Jim Melton – Oracle

- H2 was organized in 1978 to standardize database languages.

## ■ Current Participants

- Bentley Systems
- Computer Associates
- George Washington University
- HP
- IBM
- Ingres
- Intersystems
- JCC Consulting, Inc.
- Johns Hopkins University Applied Physics Laboratory
- Microsoft
- Terradata (was NCR/Terradata)
- Oracle Corporation
- Software Workshop, Inc
- Sybase
- Whitemarsh Information Systems

# Translation

- ANSI – American National Standards Institute
- INCITS – InterNational Committee for Information Technology Standards
- H2 – the committee responsible for database standards

# Current Status

- Brief History
- SQL 2003
- SQL 2008
- Validation Test
- SQL/MM

# Brief History

The following is a brief history of the SQL Standard's major revisions:

- SQL-86 (ANSI) and SQL-87 (ISO)
- SQL-89
- SQL-92
- SQL:1999
- SQL:2003
- SQL:2003 Technical Corrigendum
- SQL:2008

# SQL:1999

- INCITS/ISO/IEC 9075-1 1999 Information Technology - Database Languages - SQL - Part 1: Framework (SQL/Framework)
- INCITS/ISO/IEC 9075-2 1999 Information Technology - Database Languages - SQL - Part 2: Foundation (SQL/Foundation)
- INCITS/ISO/IEC 9075-3 1999 Information Technology - Database Languages - SQL - Part 3: Call Level Interface (SQL/CLI)
- INCITS/ISO/IEC 9075-4 1999 Information Technology - Database Languages - SQL - Part 4: Persistent Stored Modules (SQL/PSM)
- INCITS/ISO/IEC 9075-5 1999 Information Technology - Database Languages - SQL - Part 5: Host Language Bindings (SQL/Bindings)

# SQL:2003

For SQL 2003:

- Part 5, host language bindings was re-integrated with Part 2, SQL/Foundation
- The Schema information tables have been pulled out into a separate part, Part 11, SQL/Schemata.
- In addition, there are four new parts:
  - Part 9 – SQL/MED – Management of External Data
  - Part 10 – SQL/OLB – Object Language Bindings – SQL embedded in Java
  - Part 13 – SQL/JRT – Java Routines and Types – Using Java methods and classes as stored procedures and data types in an SQL database.
  - Part 14 – SQL/XML – Using SQL and XML together.



# SQL:2003 & SQL:2008

With these changes, the complete list for SQL 2003 & SQL 2008 are:

- Part 1: Framework (SQL/Framework)
- Part 2: Foundation (SQL/Foundation)
- Part 3: Call-Level Interface (SQL/CLI)
- Part 4: Persistent Stored Modules (SQL/PSM)
- Part 9: Management of External Data (SQL/MED)
- Part 10: Object Language Bindings (SQL/OLB)
- Part 11: Information and Definition Schemas (SQL/Schemata)
- Part 13: SQL Routines and Types Using the Java™ Programming Language (SQL/JRT)
- Part 14: XML-Related Specifications (SQL/XML)



# SQL:2003 Technical Corrigendum

The Technical Corrigendum contains interpretations and bug fixes for all of the SQL:2003 parts:

- Part 1: Framework (SQL/Framework)
- Part 2: Foundation (SQL/Foundation)
- Part 3: Call-Level Interface (SQL/CLI)
- Part 4: Persistent Stored Modules (SQL/PSM)
- Part 9: Management of External Data (SQL/MED)
- Part 10: Object Language Bindings (SQL/OLB)
- Part 11: Information and Definition Schemas (SQL/Schemata)
- Part 13: SQL Routines and Types Using the Java™ Programming Language (SQL/JRT)
- Part 14: XML-Related Specifications (SQL/XML)

Corrigendum 2 was completed 2006-09, published 2007-04.

# SQL/XML:2006

- The SQL committees revised SQL/XML-2003 to align with the W3C (World Wide Web Consortium) version of XQuery.
- SQL/XML 2006 references the XQuery specification.
- Published in June, 2006.

# SQL:2008

The SQL committees completed eight of the nine parts at the July, 2007 meeting in New York.

- The ninth part, SQL/XML, was completed in March, 2008.
- Aligns with W3C XML/Update
- Some other new features
- Bug fixes
- All 9 parts published 17 July 2008

# SQL/Foundation:2008 New Features

- The following slides briefly discuss some of the new features in SQL/Foundation:2008
  - Regular Expressions
  - Order By Enhancements
  - Fetch First
- This is neither a complete list nor a complete discussion.

# Regular Expressions

SQL:2003 Similar Predicate was a hybrid of PERL, XML, etc. regular expressions.

- SQL:2008 REGEX predicates reference XML Regular Expression syntax
- Allows regular expression code from other sources to be reused
  - F841 LIKE\_REGEX predicate
  - F842 OCCURRENCES\_REGEX function
  - F843 POSITION\_REGEX function
  - F844 SUBSTRING\_REGEX function
  - F845 TRANSLATE\_REGEX function
  - F846 Octet support in regular expression operators
  - F847 Nonconstant regular expressions

# Order By Enhancements

SQL:2003 (and prior) did not support Order By in views and subqueries.

- SQL:2008 supports:

- F850 Top-level <order by clause> in <query expression>
- F851 <order by clause> in subqueries
- F852 Top-level <order by clause> in views
- F855 Nested <order by clause> in <query expression>



# Fetch First

- Most implementations already support some sort of fetch first or limit to n rows clause.
- SQL:2008 supports:
  - F856 Nested <fetch first clause> in <query expression>
  - F857 Top-level <fetch first clause> in <query expression>
  - F858 <fetch first clause> in subqueries
  - F859 Top-level <fetch first clause> in views

# Page Count Comparison

<b>Part</b>	<b>SQL 1992</b>	<b>SQL 1999</b>	<b>SQL 2003</b>	<b>SQL 2008</b>
Part 1 – SQL/Framework		85	81	93
Part 2 – SQL/Foundation	628	1,147	1,267	1,361
Part 3 – SQL/CLI (1995)	236	421	405	403
Part 4 – SQL/PSM (1996)	256	170	184	190
Part 5 – SQL/Bindings		261		
Part 9 – SQL/MED			498	485
Part 10 – SQL/OLB			405	414
Part 11 – SQL/Schemata			296	298
Part 13 – SQL/JRT			204	208
Part 14 – SQL/XML			266	444
<b>Total</b>	<b>1,120</b>	<b>2,084</b>	<b>3,606</b>	<b>3,896</b>



# What happened to the missing parts?

There are several part numbers that have been retired:

- Part 5 - SQL/Bindings – integrated back into SQL/Foundation
- Part 6 - SQL/Transaction – withdrawn
- Part 7 - SQL/Temporal – withdrawn
- Part 8 - SQL/Objects – Extended Objects – integrated back into SQL/Foundation
- Part 12 - SQL/Replication – never got moving

# Part 7 - SQL/Temporal

Expansion of SQL Standard to deal with issues of time

- Transaction time – What was the data when it was recorded
- Valid Time – What should the data have been at a particular point in time
- Two competing academic viewpoints
  - Richard T. Snodgrass, Christian S. Jensen, “Developing Time-Oriented Database Applications in SQL”, Morgan Kaufmann, 1999. This book is out of print but can be downloaded from Rick’s web site at <http://www.cs.arizona.edu/people/rts/>
  - C. J. Date, Hugh Darwen, & Nikos Lorentzos, “Temporal Data & the Relational Model”, Morgan Kaufmann, 2002.
- Important issues, but technically challenging and complex

# Validation Test

NIST canceled the SQL Validation Test project in 1996

- No organization is testing SQL Standards Compliance
- Vendors more-or-less adhere to the SQL standard due to market pressure
- New features implemented when requested by users
- If users do not know about a standard feature, it is unlikely to be requested

# SQL/MM

The SQL Multimedia and Application Packages specifications are layered on top of the SQL Standard.

- SQL/MM Part 1: Framework
- SQL/MM Part 2: Full Text
- SQL/MM Part 3: Spatial
- SQL/MM Part 5: Still Image
- SQL/MM Part 6: Data Mining
- SQL/MM Part 7: History

# Directions

Over the decades, the SQL Standards process has incorporated new technology directions:

- Object oriented capabilities.
- Using Java with SQL.
- Most of the current SQL standards development is in the area of SQL and XML.
- We are not sure of the next directions.
  - System Versioned Tables
  - Additional expansions to SQL/XML
  - Row Pattern Recognition
  - SPARQL
  - Database security

# System Versioned Tables

Last year, WG3 adopted a proposal from the USA that added support for System Versioned Tables

- Essentially an audit trail of changes
- Syntax for executing a query as of a point in time
- Specify retention – Keep for  $n$  years
- First step towards full temporal support



# Row Pattern Recognition

New proposal for additional capabilities to support streaming data queries

- Patterns defined using Regular Expressions
- Regular Expression variables span subsequences of rows
- Defined using conditions on individual rows and their aggregates

# SPARQL

- Semantic Web Query Language for RDF
- Uses RDF triplets and OWL ontologies
- “Graph Query Language”
- Should be possible to represent SQL data and schema information tables as RDF triplets
- Apply reasoning engines to learn new things about existing data
- Lots of fuzzy definitions in this area



# SPARQL Abstract

- RDF is a directed, labeled graph data format for representing information in the Web. This specification defines the syntax and semantics of the SPARQL query language for RDF. SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions. SPARQL also supports extensible value testing and constraining queries by source RDF graph. The results of SPARQL queries can be results sets or RDF graphs

From “SPARQL Query Language for RDF”, W3C Candidate Recommendation 14 June 2007, <http://www.w3.org/TR/rdf-sparql-query/>

# SQL Security

Japanese have proposed work in two areas:

- Audit trails
  - Similar to Rdb audit capability
  - Not based on any security or auditing standard
- Ability to define limit for number of rows accessed by a query

# Other Possible Security Areas

- Separate create table privilege from read data and drop table privileges
- Column encryption
- Row tagging
- Authenticate against external authentication service, such as Kerberos
- Issues with standardizing security
  - Must integrate with security external to the database
  - No agreement on what security means

# Getting Copies of the Standards

Because of ANSI and ISO copyright restrictions, copies of the SQL standards specifications are available only for purchase.

- Electronic (PDF) or printed copies of the SQL standards are available from several sources.
- ISO – <http://www.iso.ch>, click on "ISO STORE", and search for 9075. Prices are in Swiss Francs. A CD of all 9 parts is CHF 356.00
- ANSI – <http://www.ansi.org/>, click on "eStandards Store" and search for SQL.
  - Make sure you choose the 2008 versions
  - ISO/IEC 9075-\*:2008

# Summary

- SQL Standards development is ongoing
- Visibility of SQL Standards development is not high
- New standards development is following new technologies
  - System Versioned Tables
  - XML
  - Row Pattern Recognition?
  - Security?
  - RDF & SPARQL?

# Questions?

