PROPERTY GRAPH STANDARDS, PROCESS & TIMING

Keith W. Hare
Convenor, ISO/IEC JTC1 SC32 WG3 Database Languages

August 1, 2021
Introduction

• Database Language Standards Structure and Process
  • International Standards Hierarchy
  • ISO/IEC JTC1 Standardization Process
  • USA Standards Structure
  • Who participates?

• Property Graph Standards – SQL/PGQ & GQL
  • Brief History of SQL and GQL
  • SQL/PGQ & GQL expected dates

• Summary
Database Language Standards Working Group

ISO/IEC JTC1 SC32 WG3 Database Languages

- ISO – International Organization for Standardization
- IEC – International Electrotechnical Commission
- JTC 1 – Joint Technical Committee 1 – Information Technology standards
- SC 32 – Sub Committee 32 – Data Management and Interchange
- WG 3 – Working Group 3 – Database Languages
Standardization Steps and Acronyms

- New Work Item Proposal – NWIP
- Working Draft – WD
- Committee Draft – CD
- Draft International Standard – DIS
- Final Draft International Standard – FDIS
- International Standard – IS
ISO/IEC JTC1 Standardization Process

NWIP or Project Split

Development

Working Draft

Ready?

Yes

No

Yes

No

CD Ballot

Comment Resolution

Done?

Yes

No

DIS Ballot

Comment Resolution

Done?

Yes

No

FDIS Ballot

International Standard

Yes

No
Who participates – SC32 WG3?

Experts from the following national bodies participate in SC32 WG3:

1. China
2. Denmark
3. Finland
4. Germany
5. Japan
6. Korea
7. Netherlands
8. Sweden
9. United Kingdom
10. United States
Organizations Participating in National Bodies

- China
  - Ant Financial
  - Boray Data
  - CESI
  - Huawei
- Denmark
  - TF Informatik
- Finland
  - Profium
- Germany
  - 2nd Quadrant
  - Oracle
  - Rasdaman
- Japan
  - Hitachi
  - Tokyo Metropolitan University
- Korea
  - Bundang Hospital
  - CnTechSystems
- Netherlands
  - Cannan Consultancy
- Sweden
  - Neo4j
- United Kingdom
  - PR Brown
  - University of Edinburgh
- USA (see later slide)

Note: This list is probably incomplete
International Hierarchy mirrored in the US

ISO
International Organization for Standardization

IEC
International Electrotechnical Commission

JTC 1
Information Technology

SC 32
Data Management Interchange

WG 3
Database Languages

ANSI
American National Standards Institute

INCITS
InterNational Committee for Information Technology Standards

DM 32
Data Management & Interchange

DM 32 Expert Groups
specific task(s)
SQL/PGQ
GQL
Streaming SQL
Who participates – INCITS DM32?

**Mostly SQL**
- Actian Corporation
- IBM Corporation
- Intersystems Corporation
- Microsoft Corporation
- Oracle
- SAP
- Teradata

**Mostly GQL**
- ArangoDB Inc
- FairCom USA
- Google
- JCC Consulting Inc
- Katana Graph
- Neo4j Inc
- Optum Technology
- Redis Labs
- TigerGraph

**Mostly Streaming SQL**
- Alibaba Group
- Amazon Web Services
- Boray Data
- Confluent
- Hazelcast
- Materialize
- Snowflake
- SQLstream, a Thales Company

**Mostly Metadata**
- Farance Inc
- Institute for Defense Analyses
- William McCarthy
- National Cancer Institute
- Nurocor
- United States Dept of the Army - CERDEC-I2WD
Working In INCITS DM32

- Work done by interested parties in the Expert Groups groups
  - Property Graph Queries in SQL
  - GQL
  - PGQ & GQL have 2-hour web conferences on alternate weeks
- The Expert Groups agree on designs
- DM32 reviews papers from the Expert Groups
- Discussions in the Expert Groups is based on written papers
  - Concrete change proposals
  - Discussion papers
    - Basis for discussion of designs or alternatives
    - Discussion will eventually lead to a Change Proposal
Working In ISO/IEC JTC1 SC32 WG

• International Committee
  • Participants from national standards bodies
  • Participants operate as individual experts

• Monthly web conferences
  • Times offset by 8 hours every month
  • Was week-long meetings two to three times a year

• Written papers
  • Concrete change proposals
  • Discussion papers

• Final decisions are made in WG3
  • Decisions recorded in minutes
  • Editors apply approved papers to drafts
SC32 WG3 Formal Liaison Relationships

- LDBC (Linked Data Benchmark Council) — liaison since 2017
  - Industry/Academic consortium focused on graphs
  - Started with benchmarks & has evolved to model, language, and other topics
  - Working Groups of interest to SC32 WG3 focus on property graph language – PGQ & SQL
    - Existing Languages, Property Graph Schema, GQL Formal Semantics working groups
  - Support/strengthen WG3 standards
    - Review of WG3 documents
    - Contribution of papers to WG3 (critique/corrections, feature suggestions)
  - Property graph theory
    - ACM SIGMOD/PODS 2021 paper & presentation “PG-Keys: Keys for Property Graphs”
    - An evolving bi-directional process for collaboration

- OGC (Open Geospatial Consortium)
  - Requirements for supporting spatial data in GQL (v2 or later)
SQL Standards – a brief history

ISO/IEC 9075 Database Language SQL
- SQL-87 – Transactions, Create, Read, Update, Delete
- SQL-89 – Referential Integrity
- SQL-92 – Internationalization, etc.
- SQL:1999 – User Defined Types
- SQL:2003 – XML & OLAP
- SQL:2008 – Expansions and corrections
- SQL:2011 – Temporal
- **SQL:202x – SQL/PGQ, Property Graph Queries in SQL**
  - ISO/IEC 9075-16 *Information technology — Database languages SQL — Part 16: SQL Property Graph Queries (SQL/PGQ)*
  - SQL/PGQ project initiated September, 2017
GQL Standards – a brief history

ISO/IEC 39075 *Information Technology — Database Languages — GQL*

- Create, Read, Update, Delete, Transactions, Schema
- GQL project initiated September, 2019
Property Graphs – SQL/PGQ and GQL

**SQL/PGQ**
- Property Graph views of SQL tables
- **Graph Pattern Matching queries**
  - GRAPH_TABLE() in SQL FROM
  - Supports Reads
- Common foundation with SQL and graph query languages
- Does not support schema-flexible graphs

**GQL**
- Full DB language
  - DML – Create, Read, Update, Delete
  - DDL – Create Type, Create Graph
- **Graph Pattern Matching queries**
- Leverages common foundation from SQL and property graph languages
- Supports schema-fixed and schema-flexible variants
## Expected Dates – Detailed

<table>
<thead>
<tr>
<th>Date</th>
<th>SQL/Foundation (and four other parts)</th>
<th>SQL/PGQ</th>
<th>GQL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021-02-07</td>
<td>SQL/Foundation CD Ballot Ended</td>
<td>SQL CD Ballot Ended</td>
<td></td>
</tr>
<tr>
<td>2021-02-07</td>
<td>SQL/PGQ CD2 Starts (8 weeks)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021-10-03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021-11-27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021-11-28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021-11-28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021-11-28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022-02-19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022-03-04</td>
<td>SQL/Foundation DIS preprocessing starts</td>
<td>SQL/PGQ DIS Preprocessing Starts</td>
<td></td>
</tr>
<tr>
<td>2022-04-24</td>
<td>SQL/Foundation DIS Ballot starts</td>
<td>SQL/PGQ DIS Ballot Starts</td>
<td></td>
</tr>
<tr>
<td>2022-07-04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022-07-17</td>
<td>SQL/Foundation DIS Ballot ends</td>
<td>SQL/PGQ DIS Ballot Ends</td>
<td></td>
</tr>
<tr>
<td>2022-08-28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022-09-04</td>
<td>Final text to ISO</td>
<td>Final text to ISO</td>
<td></td>
</tr>
<tr>
<td><strong>2022-10-16</strong></td>
<td><strong>SQL/Foundation Published</strong></td>
<td><strong>SQL/PGQ Published</strong></td>
<td></td>
</tr>
<tr>
<td>2022-12-25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023-02-19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023-05-21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023-07-30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2023-09-10</strong></td>
<td><strong>SQL/PGQ Published</strong></td>
<td><strong>GQL Published</strong></td>
<td></td>
</tr>
</tbody>
</table>
Expected Dates – Summary

• SQL/PGQ
  • CD Ballot completed – February 2021
  • DIS Ballot starts – March 2022
  • DIS Ballot completes – July 2022
  • Published standard – October 2022

• GQL V1
  • CD Ballot completes – February 2022
  • DIS Ballot Starts – December 2022
  • DIS Ballot completes – May 2023
  • Published standard – September 2023

• Draft Standards are stable by DIS ballot start
Summary

• Standards Process
  • Iterative, collaborative process
  • Some amount of standards bureaucracy
  • Tedious at times
  • Results are pretty good

• Property Graph Standards
  • SQL/PGQ published October 2022
  • GQL published September 2023