ANALYTICS AND REPRESENTATIONS IN SQL AND GQL

Keith W. Hare
SC32 WG3 Convenor
JCC Consulting, Inc.
November 13, 2019

Who am I?

- JCC Consulting, Inc. since 1985
  - President since August 2019
  - Senior Consultant 1985-2019
  - High performance database systems
  - Replicating data between database systems
- SQL Standards committees since 1988
  - Convenor, ISO/IEC JTC1 SC32 WG3 Database Languages (since 2005)
  - Vice Chair, ANSI INCITS DM32.2 Databases (since 2003)
  - Vice Chair, INCITS Big Data Technical Committee (2015-2017)
- Education
  - Muskingum College, 1980, BS in Biology and Computer Science
  - Ohio State, 1985, Masters in Computer & Information Science

Introduction

- SQL
  - Online Analytical Processing (OLAP) capabilities
  - Temporal
  - Multidimensional Arrays
- Property Graphs
  - Property Graph queries in SQL – SQL/PGQ
  - New project to specify a Declarative Property Graph Query Language – GQL
- Other topics under discussion
  - Streaming SQL
  - Additional statistical functions
- Summary

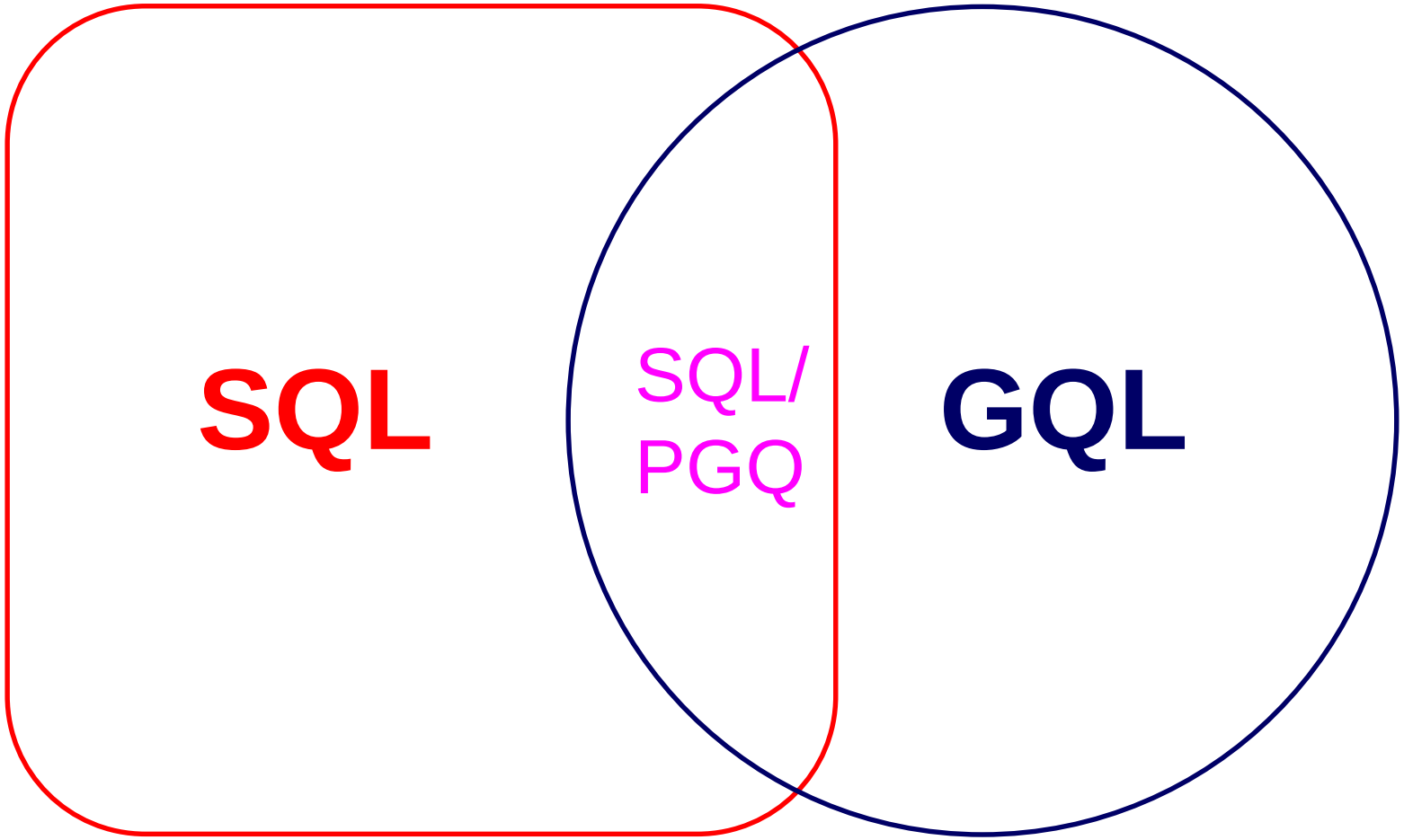
SQL

- Online Analytical Processing (OLAP) capabilities
  - Added in SQL-2003
  - Window Functions – Rank, Distribution, Aggregate, Ntile…
- Temporal
  - Added in SQL-2011
  - System Versioned Tables – what did the data look like at a particular point in time?
  - Application time period tables – what should the data have looked like?
- Multidimensional Arrays
  - SQL/MDA – Published in 2019
  - Multi-dimensional arrays as columns in a table
  - Operations on arrays

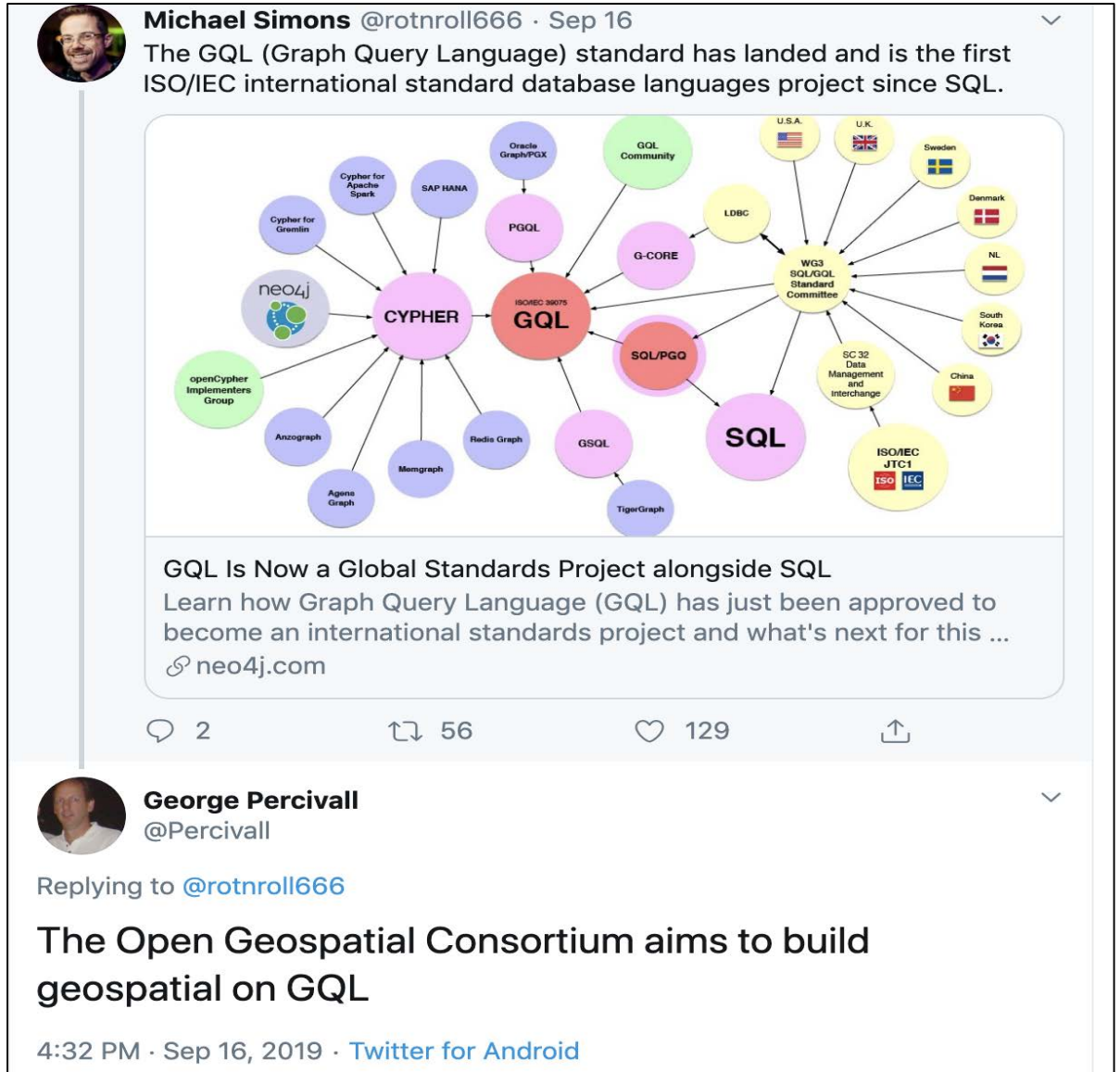
Property Graphs

- Property Graph queries in SQL – SQL/PGQ
  - Integrate property graph queries into SQL queries
  - Path language
  - Define property graph views over relational data
- New project to specify a Declarative Property Graph Query Language – GQL
  - Add, Modify, Query, Delete
  - Path language
  - Transactions
  - Use SQL specification where possible
  - Defined schema or schema-less

SQL, SQL/PGQ, and GQL



GQL Use Case – Geospatial



- OGC CTO & Chief Engineer
- Opengeospatial.org
- GeoSPARQL

Working with OGC

- Advantageous to have some geospatial types built in and supported by the language/platform - at least Coordinate Position
- OGC meeting in November in France:
  - WG3 representative: Tobias Lindaaker
  - “White paper” containing changes planned for GeoSPARQL as well as a requirements specification for geospatial on GQL

Other New Areas

- Streaming SQL
  - Query data before, or instead of, storing
  - Input from zero or more streams
  - Output to zero or more tables or streams
  - Additional windowing functions specific to streaming data
- Additional statistical functions
  - Identify useful statistical or analytical functions
  - Specify signature
  - Point at textbook or standard reference for algorithm

Summary

- SQL Standards have a long history of integrating new technologies
- Extensive analytical capabilities
  - OLAP
  - Temporal support
  - Multi-dimensional Arrays
- Property Graph Support
  - SQL/PGQ
  - GQL
- Future Enhancements
  - Streaming SQL
  - Additional functions