TopQuadrant™

TopBraid EDG™
Enterprise Data Governance

Semantic Knowledge Graphs are the Governance Architecture of the Future

November 7, 2019 11:30am ET

Today's Program

Semantic Knowledge Graphs are the Governance Architecture of the Future

- INTRODUCTION
  - **Brief Overview of Data Governance**
  - **Discuss Data Governance Challenges**
  - **Discuss Knowledge Graphs and Review Use Cases**
  - **Briefly discuss the technology behind them**
- DEMOS (several short demos will be done throughout)
- WRAP-UP
- Q & A

Jesse Lambert

TOPQUADRANT COMPANY

FOUNDATION

- TopQuadrant was founded in 2001
- Strong commitment to standards-based approaches to data semantics

MISSION

- Empower people and drive results — by making enterprise information meaningful

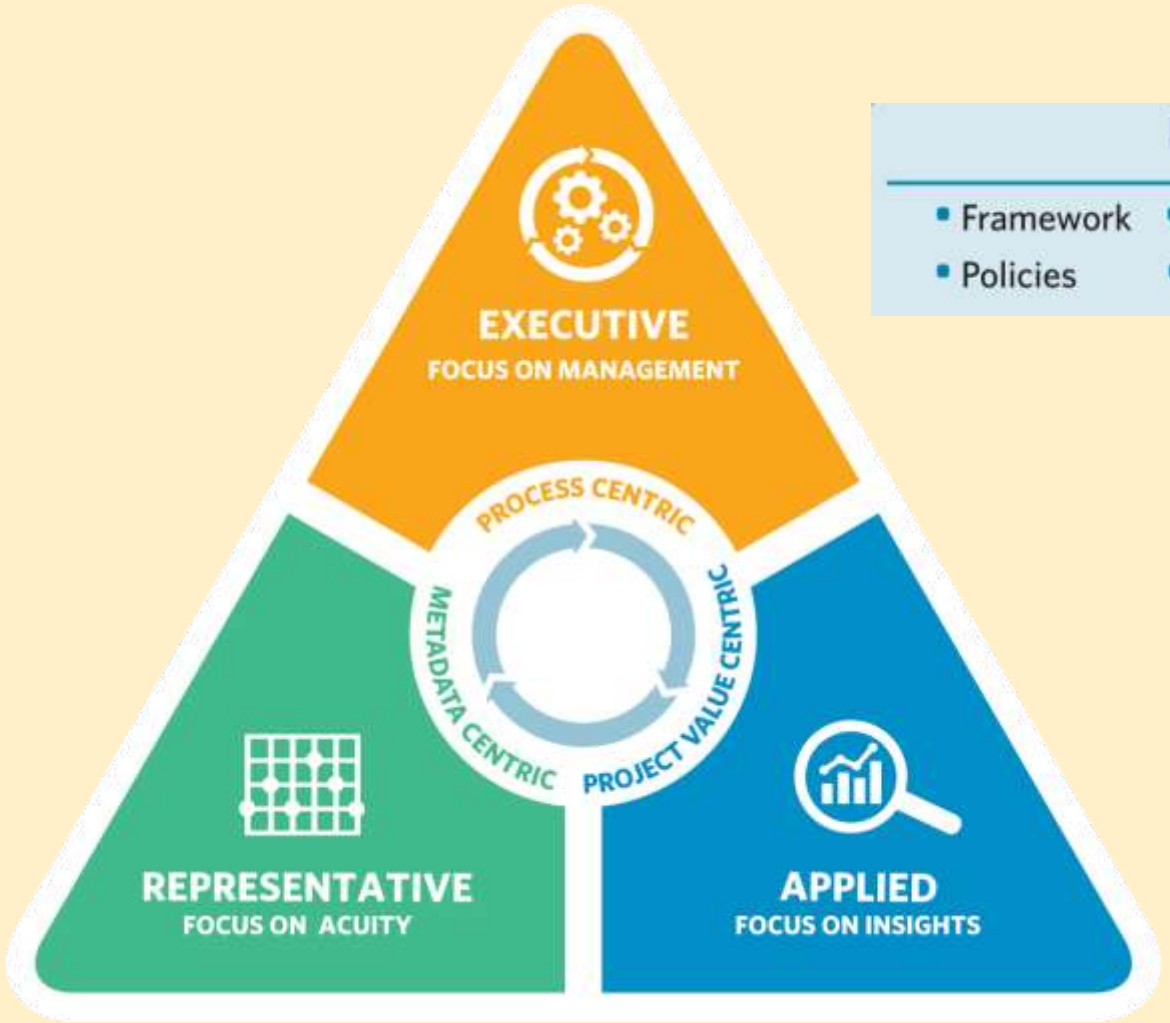
FOCUS

- Provide comprehensive data governance solutions

Data Governance Triangle

Common Primary Modes of DG Operation

- Process Centric
- Project Value Centric
- Metadata Centric



Executive Governance:

- Framework
- Methodology
- Stewardship
- Process
- Policies
- Metrics
- Maturity Assessment

Representative Governance:

- Glossaries
- Taxonomies
- Ontologies
- Data
- Technical and Enterprise Assets
- Reference Data

Applied Governance:

- Lineage
- Impact Analysis
- Metrics
- Data Quality
- Compliance Reporting

Challenges of Data Governance

NEW ROLES AND MORE PERSONNEL

In 2013, just 50 companies employed a Chief Data Officer. In 2015, Forrester found that 45% of companies (out of 3,000 surveyed) had created and filled the position.

A survey of Fortune 1000 senior leadership found that more than half of companies report that variety – not volume or velocity – represents the biggest challenge when it comes to managing their data.

CONSISTENT MEANING

Internal teams need to be able to speak the same language-and with the inability to handle data variety, meaning can easily be lost or misinterpreted.

Challenges of Data Governance

TopQuadrant™

Security

Weaknesses

Vulnerabilities

Infrastructure

Design

Plan

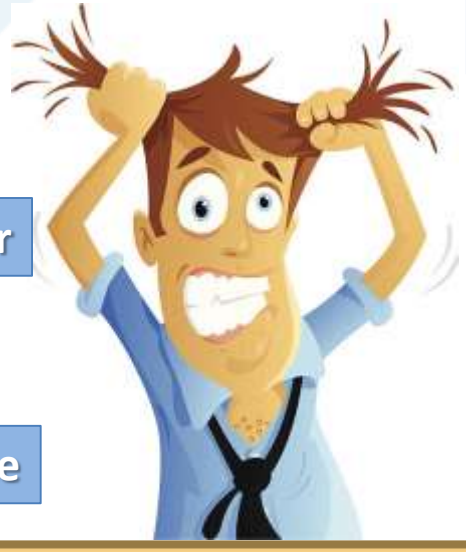
Monitor

Test

Code

Operate

Build



Deploy

Release

IT/DevOps

NEW ROLES AND MORE PERSONNEL

In 2013, just 50 companies employed a Chief Data Officer. In 2015, Forrester found that 45% of companies (out of 3,000 surveyed) had created and filled the position.

NEW INFRASTRUCTURE

Legacy systems can become overwhelmed with the increase in data volume, causing relationships to be lost, necessitating that organizations integrate existing systems and technologies with new ones to become more flexible and robust.

NEW SOURCES

As organizations collect data from a growing number of both internal and external sources, legacy data management solutions can fail to keep up. Instead of benefitting from data richness, poor management and understanding of data can result.

INCREASING VARIETY

A survey of Fortune 1000 senior leadership found that more than half of companies report that variety – not volume or velocity – represents the biggest challenge when it comes to managing their data.

CONSISTENT MEANING

Internal teams need to be able to speak the same language-and with the inability to handle data variety, meaning can easily be lost or misinterpreted.

Compliance

Should…?

Regulation

Can…?

Is it secret?
Is it safe?

Quality

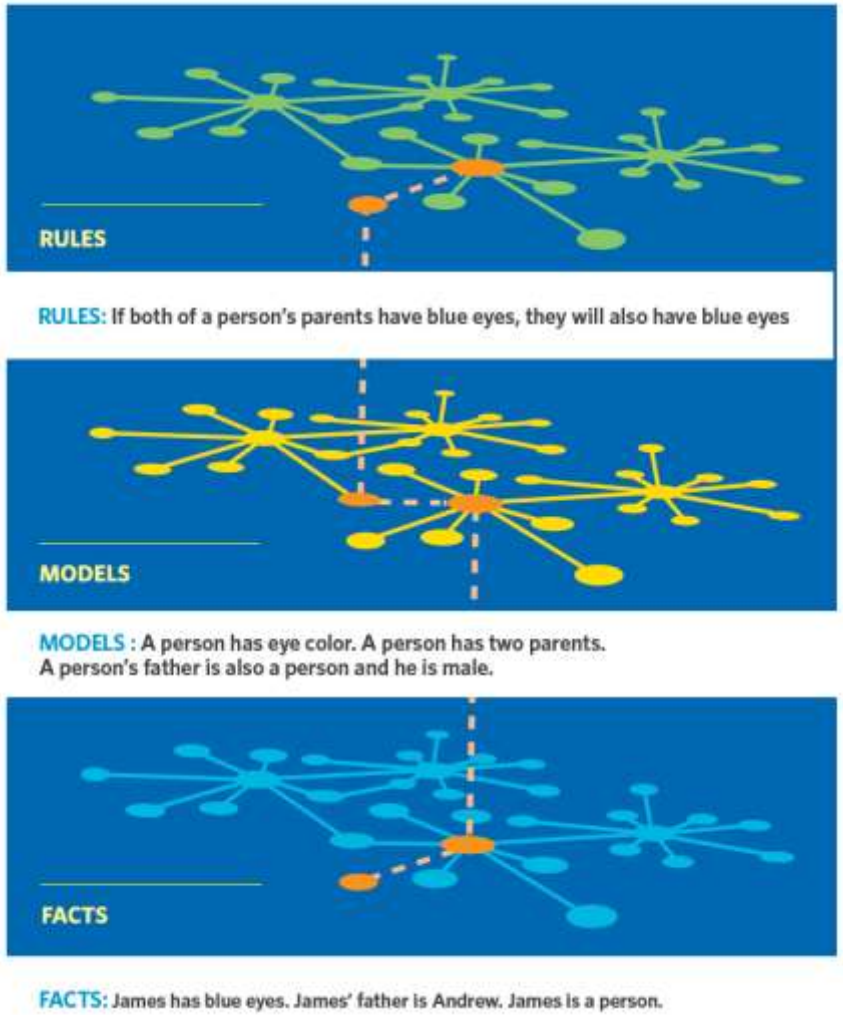
Quantity

Where…?

Data Governance Office

Enter Knowledge Graphs

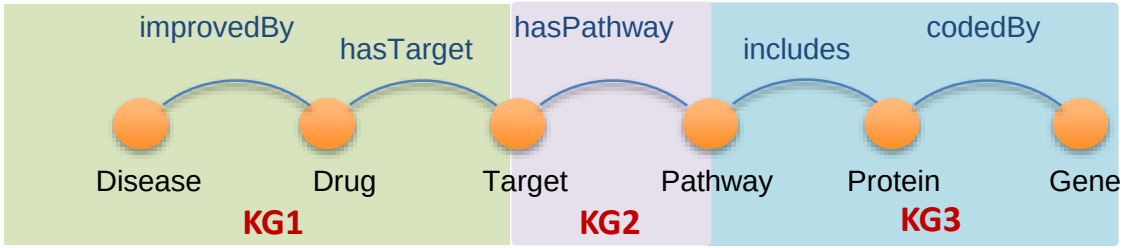
- A Knowledge Graph represents a knowledge domain
- It represents knowledge as a graph
  - A network of nodes and links
  - Not tables of rows and columns
- It represents facts (data) and models (metadata) in the same way
  - Rich rules and inferencing
- It is based on open standards, from top to bottom
  - Readily connects to knowledge in private and public clouds



There can be different types and instances of Knowledge Graphs …

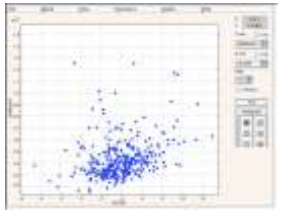
Some Use Cases for Knowledge Graphs

1. Graph Traversal



Traverse across connected graphs

2. Graph Analytics



Graph algorithms: Statistics, Centrality, Shortest Path, …

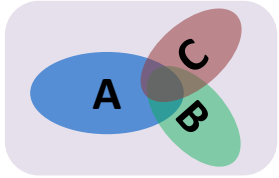
3. Data Integration

Query brokering across disparate systems by mappings to a unified model

4. Data Aggregation

360 Degree View across composable graphs

5. Information Insights



Find things that share common attributes or relationships

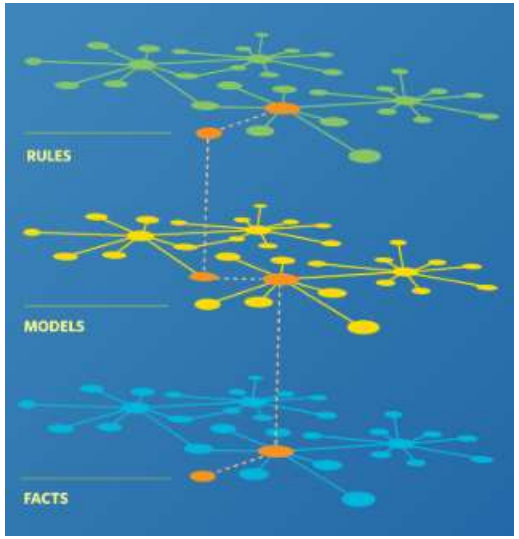
6. Lineage

Inferred dependencies across composable graphs and domains: enterprise, technical, data, governance

TopBraid EDG: An Enterprise Knowledge Graph Infrastructure for Data Governance

*TopBraid EDG supports integrated data governance across the ever growing numbers and types of data assets and governance needs – **because connections are important.***

- TopBraid EDG, is a rich set of interconnected knowledge graphs expressing knowledge about how data is used and managed in the enterprise ecosystem
- These integrated knowledge graphs are ready to be enriched with your enterprise specific knowledge
- After this enrichment, your enterprise is ready for implementing comprehensive data governance



Using Knowledge Graphs TopBraid EDG Delivers Data Governance 2.0 – Integrating Executive, Representative and Applied Governance Capabilities

Asset Management in TopBraid EDG

With TopBraid EDG, enterprises can:

COLLABORATE with all data stakeholders to improve the quality and consistency of data

CATALOG all your assets and create critical mappings

CONNECT all your data assets for increased visibility and agile responsiveness to new data requirements

CURATE your data and metadata to enhance access to all relevant information and use it more intelligently

TopBraid EDG includes an information architecture for different 'asset collections' that cover 100s of predefined asset types.

TopQuadrant™

DEMO

Flexibility – Glossary Example

Asset Example for Medical Enterprise

Lineage Asset Collections

Lineage Assets for Patient Discharge

Enterprise Asset Collections

Hospital Organizational Structures

Business Organizations

Activity and Process Structures

Information Assets

Technical Asset Collections

Hospital Information Systems

Data Flows

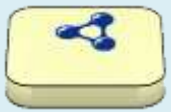
Data Asset Collections

Open MRS Data Assets

Datasets

Other Collections

Healthcare Glossary



Reference Data

Governance Assets

Business Areas

Data Subject Areas

Roles and Responsibilities

Workflows

Validation Rules

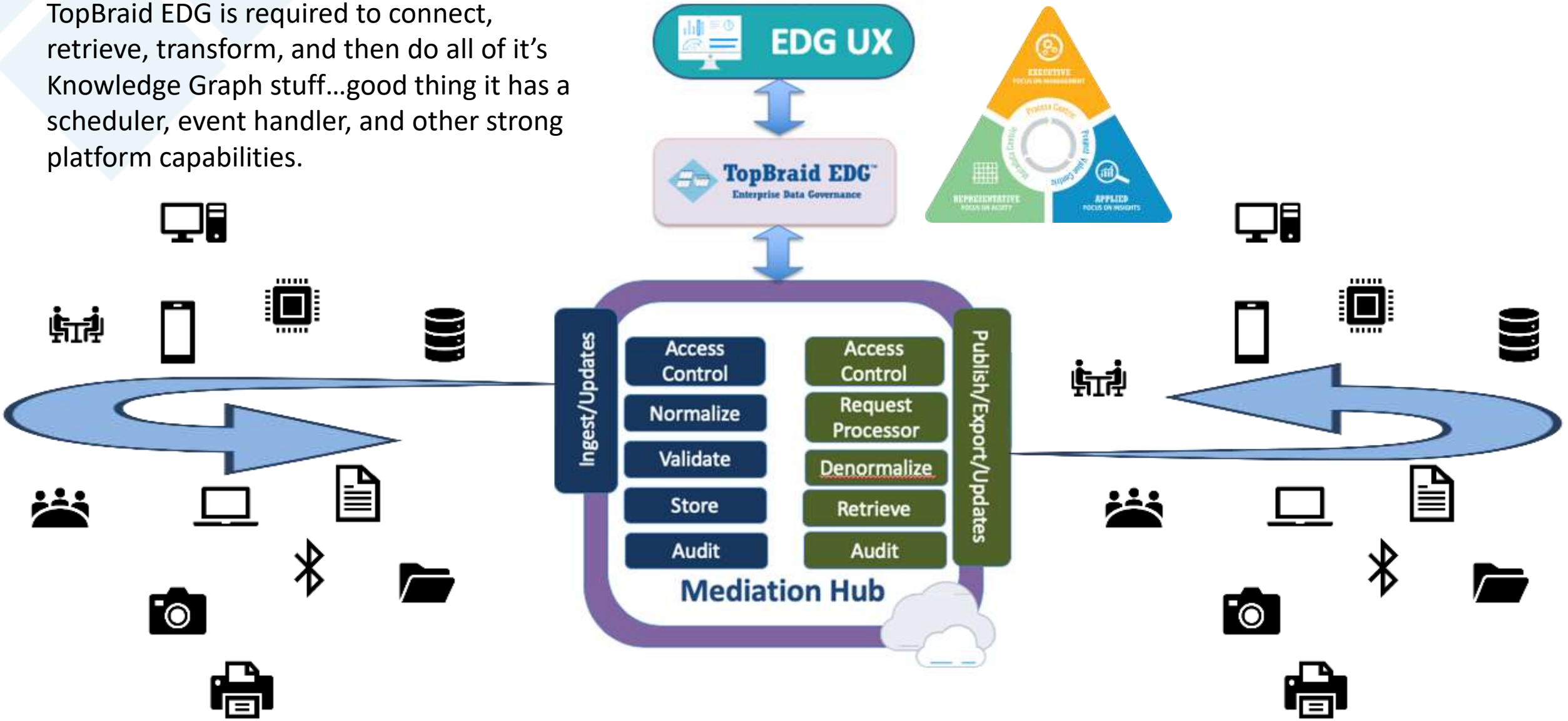
Inference Rules

Issues

Policies

Common Scenario

TopBraid EDG is required to connect, retrieve, transform, and then do all of it's Knowledge Graph stuff…good thing it has a scheduler, event handler, and other strong platform capabilities.



TopQuadrant™

DEMO

Adaptability – Organization Example

What is SHACL?

- **SH**apes **A**nd **C**onstraint **L**anguage
- W3C Standard since 2017
- Schema Language for RDF
- Rich Semantic Constraints
- Shapes for Classes and Properties
- Rules

What is GraphQL

- GraphQL is an application layer query language
  - designed to interpret a string from a server or client and return that data in an understandable, stable, and predictable format.
- Developed by Facebook for internal use
- Open sourced in 2016
- Rapidly growing in popularity, replacing REST style APIs

"Describe your data, ask for what you want, get predictable results."

https://graphql.org/

GraphQL Schema + SHACL ➔ Better Together (Semantic GraphQL)

- **GraphQL**
  - Lots of similarities to SHACL (types, fields, etc)
  - GraphQL has a huge user base
  - GraphQL has tool support and user-friendly syntax
- **RDF**
  - RDF is a flexible model for knowledge graphs
  - RDF offers URIs and subclasses
- **SHACL**
  - SHACL offers rich constraints
  - SHACL can represent UI metadata
- **RDF Knowledge Graphs can be turned into GraphQL query endpoints with SHACL**

The combination of GraphQL and SHACL ➔ "Semantic *GraphQL*"

TopQuadrant™

DEMO

Reference'ability – Country Example

Future Scenario

- Goal: As Data Governance (DG) becomes more and more a reality within enterprises, engineers and stewards need to be able to build it into their systems just as they do key aspects of security, access control, logging, and analytics.
- Requirements:
  - Model-driven JSON APIs (content/metadata/data)
  - Dynamic introspection (schema/meaning)
  - UIs can not be hard-wired (UI shouldn't change just because the backend did)
  - API driven integration VS "hundreds of connectors"
  - Connect structured and unstructured information intelligently

- This is what we call Knowledge Graph Driven Development (KGDD).

Governance built in instead of an afterthought.

TopQuadrant™

DEMO

KGDD – Logical Model Example

TopBraid EDG – Composition

TopQuadrant™

Schema

Data

Enterprise Assets

Technical Assets

Big Data Assets

Data Assets

Glossaries

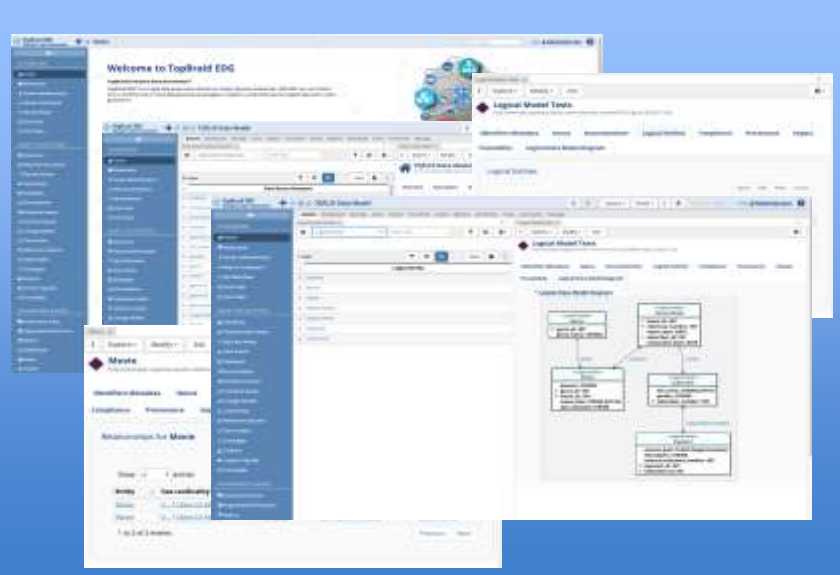
Terminologies

Governance Ontologies

Compose Data and Schema

TopBraid EDG™

Enterprise Data Governance



Compose UI Behavior, Queries, Rules and Constraints

Business Logic

UX/UI

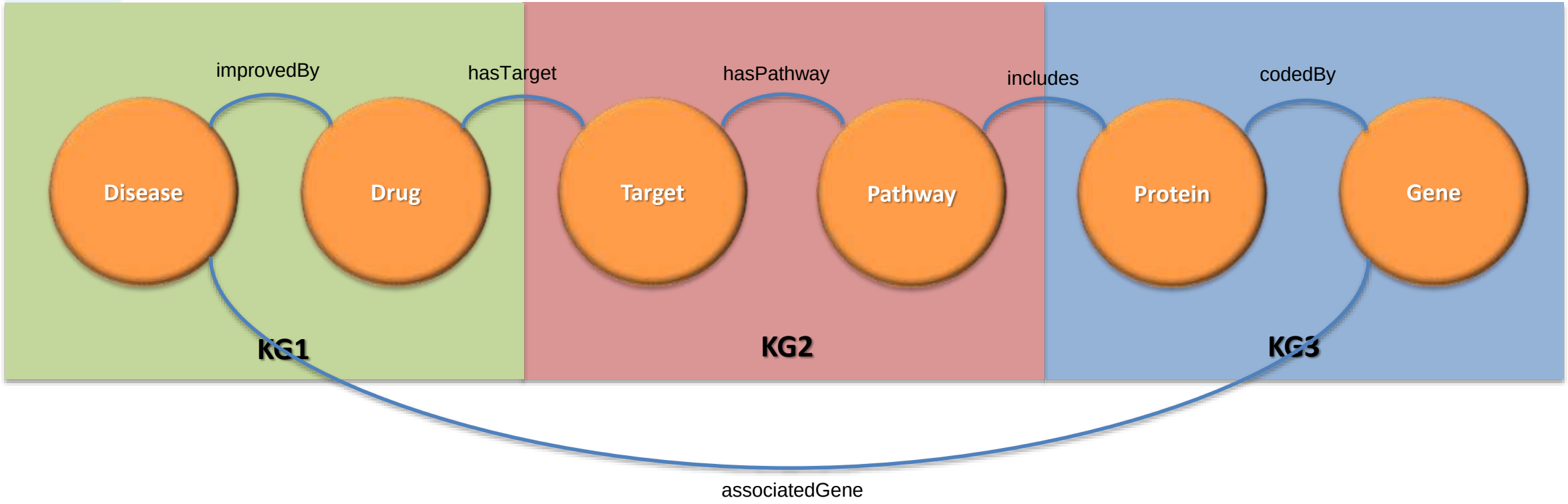
Access Control

+

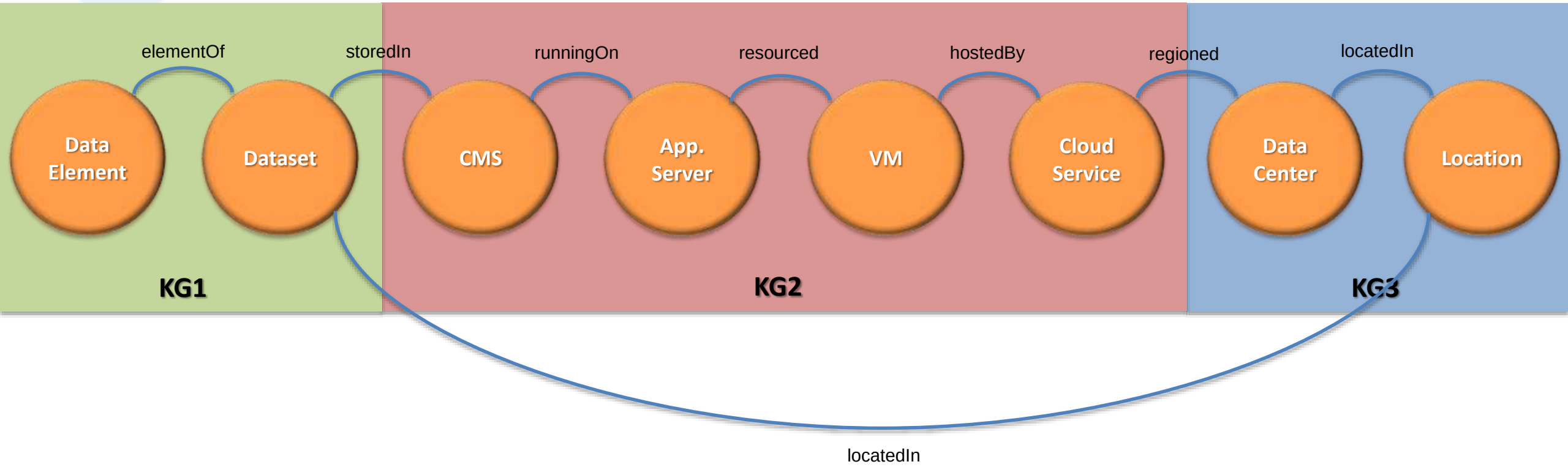
TopBraid Tagger and AutoClassifier™

** See Recent Webinar Recording: "Guided ML: Intelligent Autoclassification of Content Using Managed Vocabularies"*

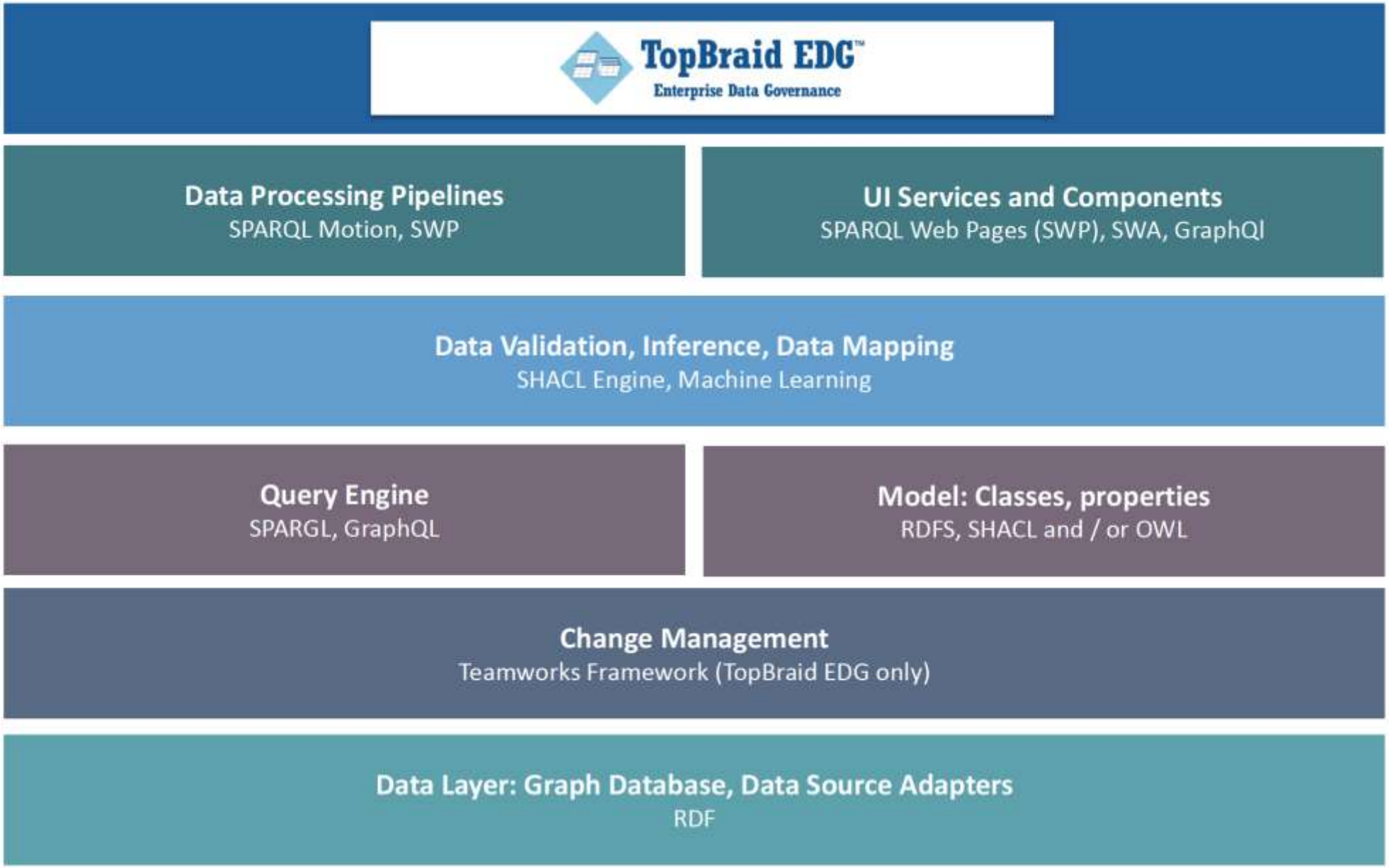
Self-Composing



Another Self-Composing Example



TopBraid EDG is Built on a Powerful Platform



TopQuadrant™

Conclusion

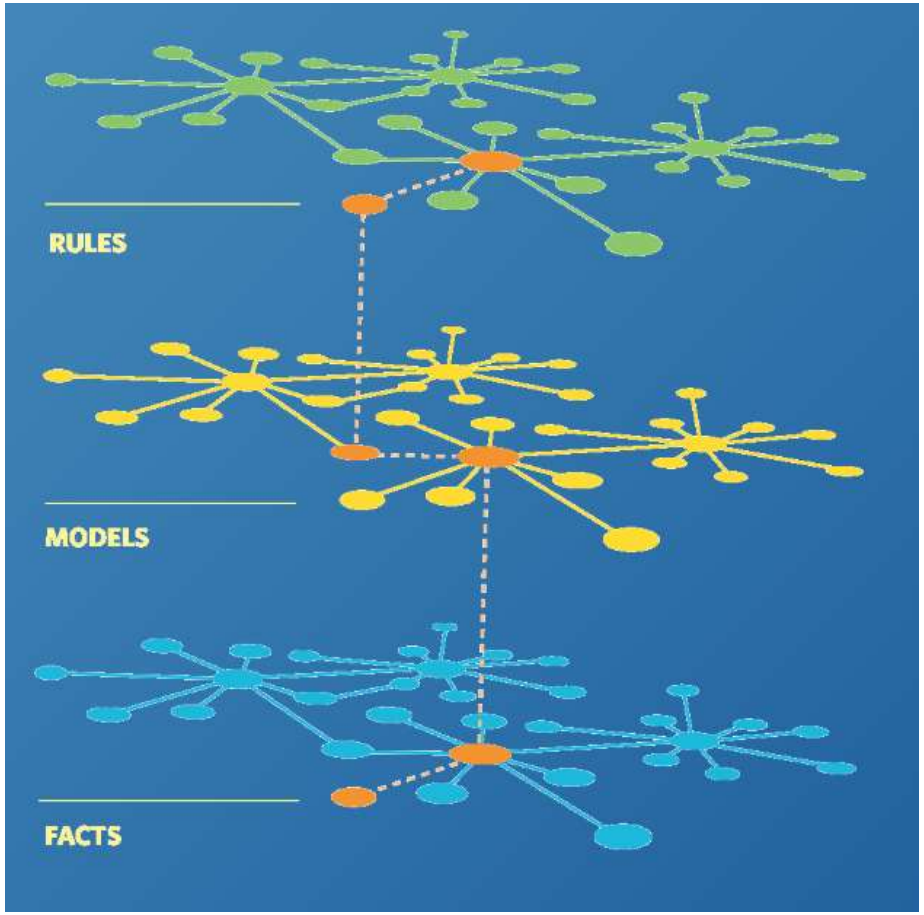
- Mature graph data standard – RDF
- Mature standard query language – SPARQL
- Mature, standard languages for expressing semantics and reasoning – SHACL and OWL
- Semantic meaning is part of data
- Many open standard ontologies
- Flexible & open to different models & viewpoints
- Vendor supported based on standards
- **TopBraid EDG is fit for a true enterprise solution**

… Questions?

Benefits of a Knowledge Graph based Platform for Data Governance 2.0

TopBraid Enterprise Data Governance (EDG):

- Is flexible and extensible, based on standards
- Integrates reasoning and machine learning
- Enables people (UI) and software (APIs/web services) to view, follow and query
- Bridges data and metadata “silos” for a seamless data governance
- Delivers Knowledge-driven data governance



As an enterprise knowledge graph infrastructure, TopBraid EDG supports Data Governance 2.0 and applications of AI / ML

TopQuadrant™

To Learn More about TopBraid EDG and Knowledge Graphs:

EDG Product Info:

- https://www.topquadrant.com/products/topbraid-enterprise-data-governance/

Contact us: at info@topquadrant.com to:

- Discuss vocabulary management solutions (glossaries, taxonomies, ontologies)
- Request a more targeted demo of TopBraid EDG – Vocabulary Management
- Ask for a free EDG evaluation account

More Resources ...

More Webinar Recordings, Slides, Q&A:

- https://www.topquadrant.com/knowledge-assets/topquadrant-webinars/

Short Videos:

- TopBraid EDG “Quick Grok” Videos
  https://www.topquadrant.com/knowledge-assets/videos/

- TopBraid EDG Animated Video
  http://www.topquadrant.com/EDGvideoemail/

Blog:

- https://www.topquadrant.com/the-semantic-ecosystems-journal/

Data Governance White Papers

- https://www.topquadrant.com/knowledge-assets/whitepapers/https://www.topquadrant.com/knowledge-assets/whitepapers/