Knowledge Graphs in Industry: Examples and Lessons Learned

Wednesday April 15, 2020
Ontolog Summit Seminar
Knowledge Graphs

Michael Uschold
Who Am I?

- Michael Uschold
- 25 years experience as ontologist and trainer
- Built commercial ontologies in numerous industries
  - Finance, Service level agreements, Electrical products, Digital asset management, Manufacturing, Legal research, Healthcare, Consumer products and Corporation registration.
- Recent author
Who is Semantic Arts?
Example: Building and Using a Knowledge Graph for Risk

- Build ontology to cover the operational risk subject.
- Convert selected data to triples.
- Combine with enterprise-wide reference data on employees, organizations and places.
- Build applications driven by the knowledge graph.

In Addition:
- Equity Research
- Technology Assets
- Information Management
Applications Driven by Knowledge Graphs (1/2)

- Tables
- Triple Creation Engine
- Mapping
- Other sources: Web pages, XML, Social media, Text documents, spreadsheets
- Triple Store
  - Ontology as data schema
  - Loaded with the data.
- Ontology (text file)
- Build ontology
Applications Driven by Knowledge Graphs (2/2)

• Use a Triple Store instead of an RDB
• Use same Ontology for multiple TS’s
• Use same TS for multiple applications
• More flexible
• Fewer silos

How do we get triples from tables?
Data and Ontology in One Graph

- Scripts
- R2RML
- Home-grown tools
- TARQL
- More tools becoming available all the time.
Data and Ontology in Same Store

- Serves multiple applications
- Can be federated across multiple stores
- Data-centric rather than Application-centric
The Benefits

- Global IRIs: enhance data integration and reuse
- Cleaner data
- Existing applications more flexible
- **Meaning first approach**: enhances reuse
- Makes new things possible

You cannot reuse what you don’t understand.
Wide Variety of Projects at Semantic Arts

- Product data and configuration
- Information & Application Integration
- Ontology driven chat-bot.
  - Modeling conversations.
  - A taxonomy of ‘intents’
- Expertise modeling for a major consultancy
- Gaming industry: moving beyond taxonomies
Case Study: Database Integration

Power products Company A (relational) → ? → Power products Company B (relational)

Ontology

Power products Integrated

© 2020 Semantic Arts, Inc.
Information Providing Companies

- Strong metadata focus
- Retail products and services across many industries
- Commodity markets
- Faceted search is a common theme
- Industry Building Blocks to Capture the Global Economy
Modeling the Global Economy

**IBB Is An Information Resource Used to Analyze:**

1. **IBB analyzes companies** in terms of what each company provides. IBB also compares companies, side-by-side, @ the Line-of-Business level.

2. **IBB analyzes market areas** in terms of the IBB industries that define them.

3. **IBB analyzes the top 21,000 industries** in terms of competitors, buyer types, vendor types, substitutes, complements, trends, etc.
Putting the Global Economy into a Knowledge Graph

- Manual industry analysis entered into spreadsheets
- Most of this information is not available anywhere else.
- Creating a web application driven by a Knowledge Graph
- In just a few months (alpha release).

Benefits:
- Automation
- Data validation
- Flexibility
- Dramatically increase availability.
Experiences and Recommendations

- The Siren Call of Semantic Silos
- The role of SHACL
- Whither schema
- Change management
Semantic Silos: Beware of Paving the Cow Paths

Layering semantic technology over silos is like paving over the cow paths.

Get short term benefit but lose out on the bigger gains.

Solution: Build an Enterprise Ontology
Agile Creation of an Enterprise Ontology (1/2)

Phase 1:

- Identify questions you want answers to as initial requirements.
- Build the ontology and triple store to meet those requirements.
- Build out applications that use the data.
Phase 2: a second iteration

- Broaden scope by identifying another set of questions as requirements
- Extend the ontology to meet the requirements.
- Coordinate with other ontology authors in the enterprise
- Make data and ontology available as triples.
- Extend existing and/or build out additional applications
Modularity, Reuse & Federated Queries

- HR Department:
  
  ![Diagram of HR Department]

- IT Department:
  
  ![Diagram of IT Department]

- How to do a federated query to identify all Personnel?
Example: connect to generic schema

- HR Department:

- IT Department:

Identical concepts will collapse into one.
- Personnel
- Building
Example: as a single graph

KEY
- Green: general
- Yellow: HR
- Pink: IT

- Federated query to identify all Personnel becomes possible.
- Jurisdiction, City and Country are all GeoRegions

URIs for Classes and Properties Enable Schema Reuse
There’s No Free Lunch

- It’s a lot of work. It’s worth it.
  - Agreeing on terminology & minting patterns
  - Evolution and extension
- Managing impacts of ontology changes
- SHACL for data validation
The Purpose of SHACL

- SHACL was designed to do many things
- A key one is to enable separation of the meaning of the subject matter (represented in OWL) from the needs of a particular application.
- One OWL ontology can be the basis for many triple stores and applications by using different SHACL constraints
The Role of SHACL

- OWL and SHACL have a structural similarity
- An OWL restriction has an echo in SHACL
- Seems a bit redundant
- Some argue that you ONLY need SHACL
What About Using SHACL instead of OWL?

Although it can work for point solutions, it has major down sides

• Blurs the distinction between
  ▪ What is true in the real word, which is relatively stable.
  ▪ What is true for an application which can change frequently

• Limits reuse/sharing of an ontology for different applications.
• Undermines the core purpose of an ontology
• Encourages paving the cowpaths with Semantic Silos
Whither Schema

- A lot of people prefer to not have a schema
- One major graph tool vendor says: you cannot have one
- You can build systems this way, but you cannot express meaning
- Defeats the original purpose of an ontology-driven approach
- Undermines understanding and reuse and drives more silos

SCHEMA: always use one, the earlier the better.
Ontology Update Pipeline

- With traditional RDB-driven applications updating the schema is rarely an option. A major driver for rigidity.
- We can do this, but there is no free lunch.
- Must inform all downstream users of ontology changes
Ontology Update Pipeline

Semantic technology specialists must:

- Update all their ontologies
- Update all the triple-creating code (e.g. TARQL)
- Re-create all the triples
- Re-load all the triples into productions stores
- Update all the SPARQL that drives application functionality.
Summary & Conclusions

- Semantic technology is going mainstream
- Drive applications from a Knowledge Graph not an RDB
- Enterprise ontology silences the siren call of semantic silos
- Beware of using SHACL instead of OWL
- Ontology evolution pipeline requires care

Developing production systems driven by ontology and knowledge graphs is now repeatable and fairly predictable