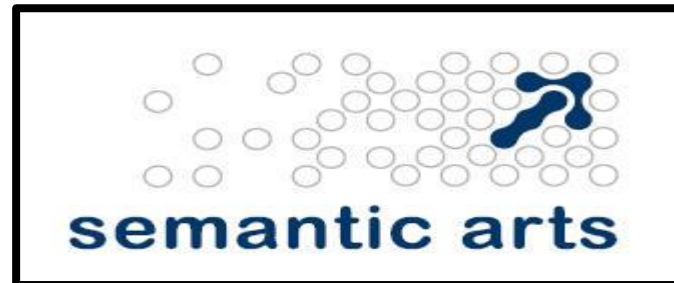# Knowledge Graphs in Industry: Examples and Lessons Learned

Wednesday April 15, 2020

Ontolog Summit Seminar
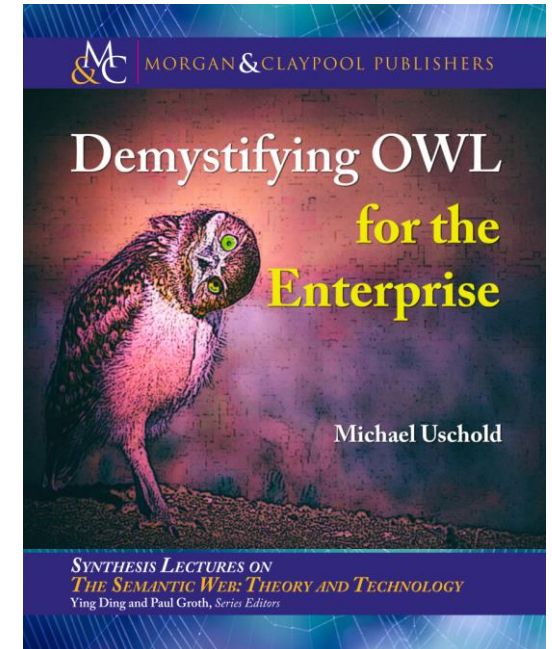
Knowledge Graphs
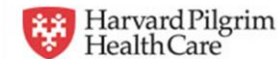
Michael Uschold

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
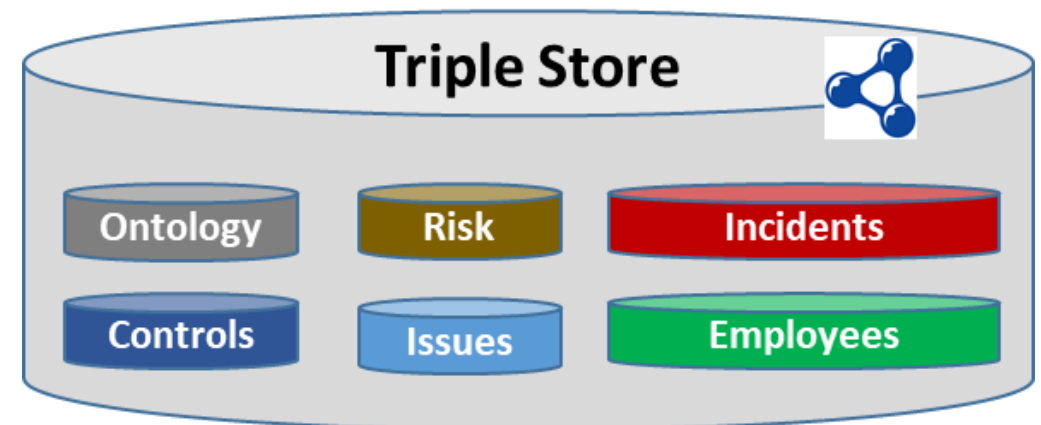
Who is Semantic Arts?

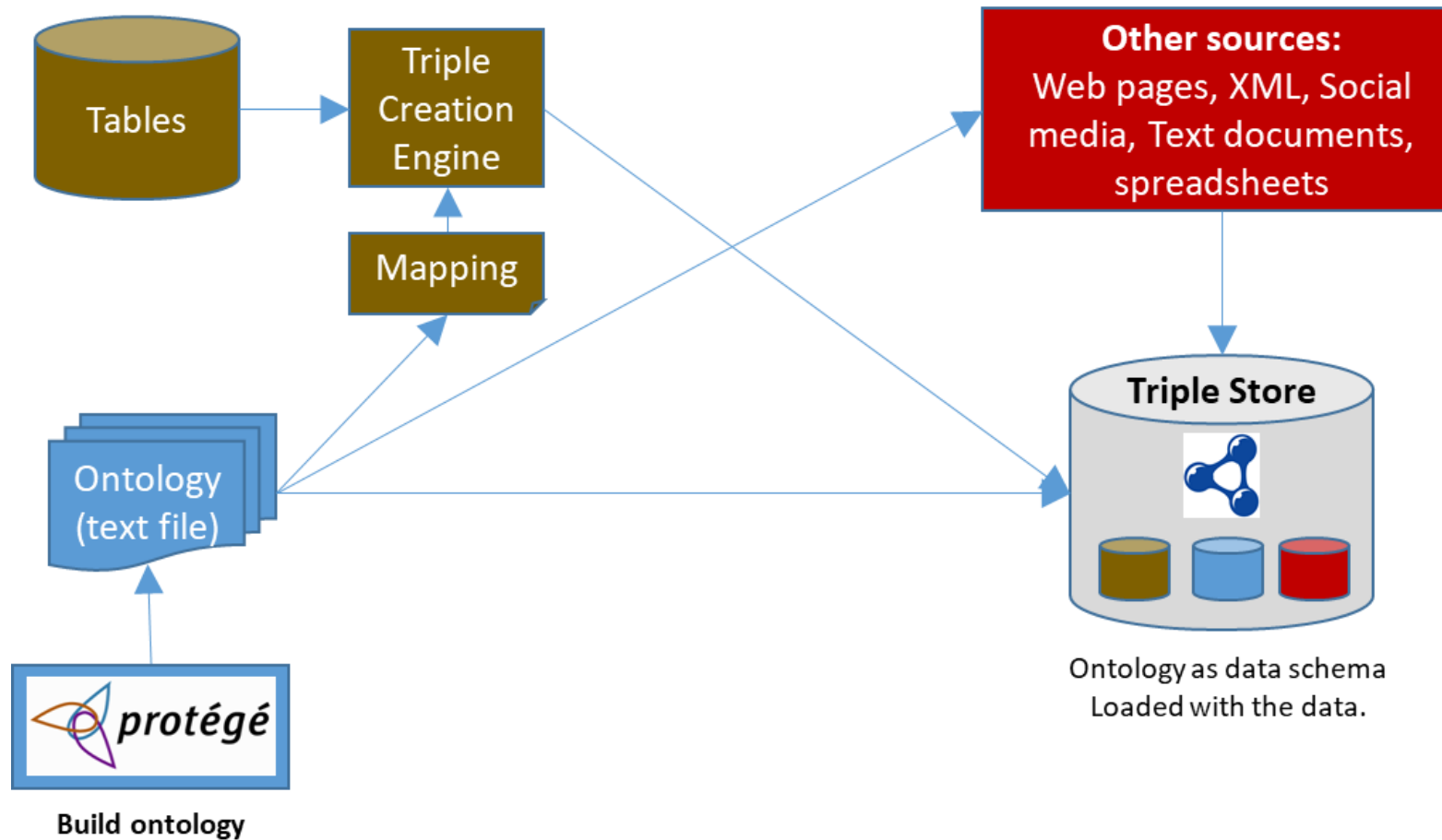# Example: Building and Using a Knowledge Graph for Risk

- Build ontology to cover the operational risk subject.

- Convert selected data to triples.

- Combine with enterprise-wide reference data on employees, organizations and places.

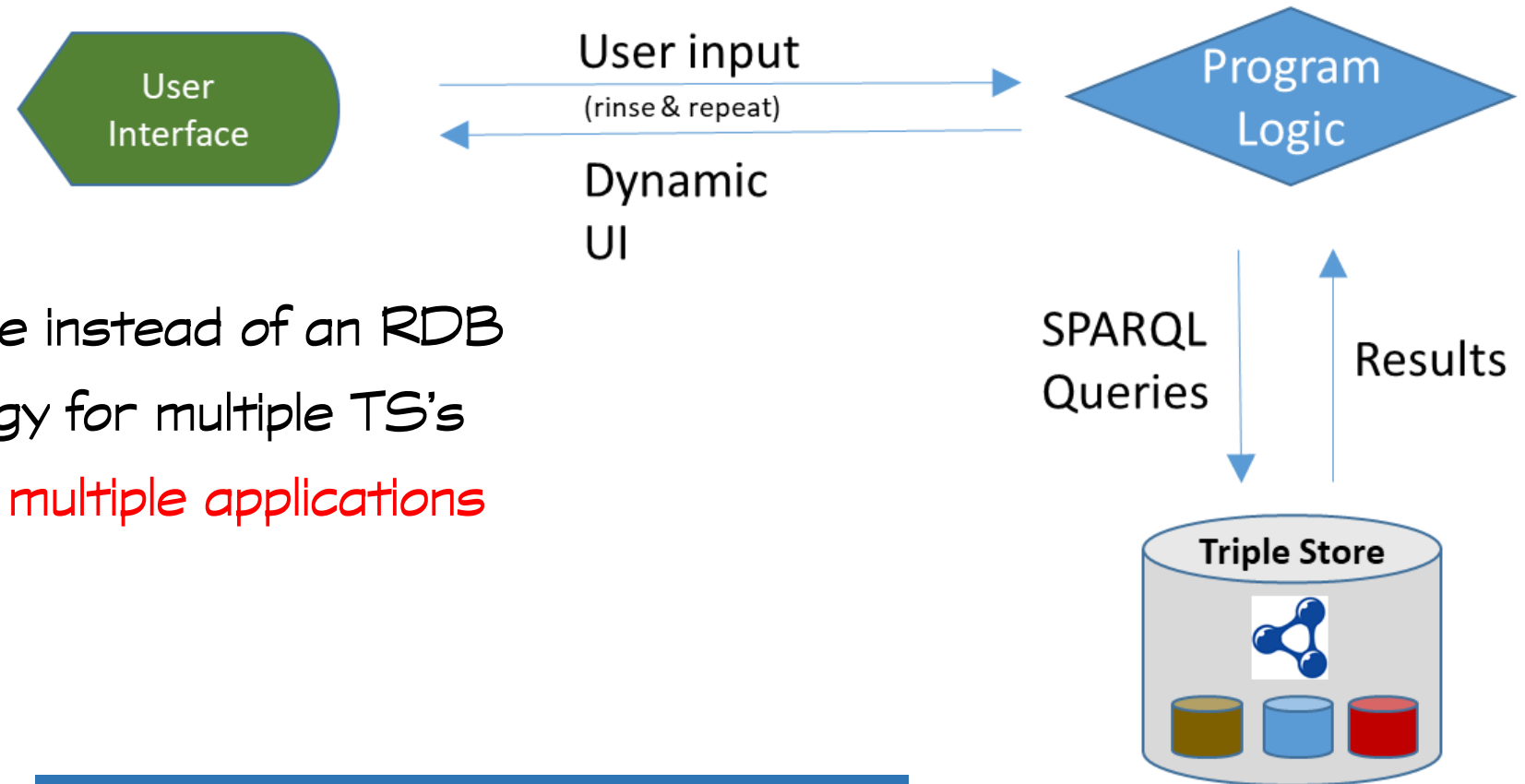- Build applications driven by the knowledge graph.

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

Ontology (text file)

protégé

Build ontology

Triple Store

Ontology as data schema
Loaded with the data.

semantic arts

Applications Driven by Knowledge Graphs (2/2)

User Interface

User input
(rinse & repeat)

Dynamic UI

Program Logic

SPARQL Queries

Results

Triple Store

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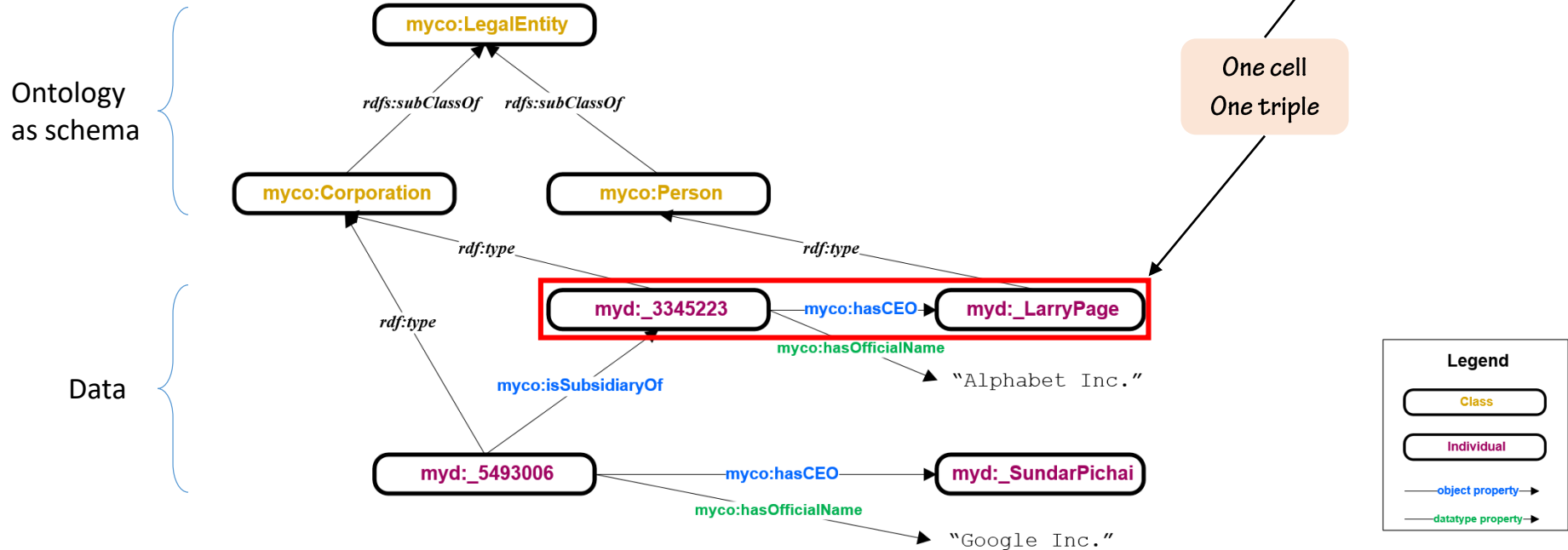- Use same TS for multiple applications
- More flexible
- Fewer silos

**How do we get triples from tables?**

semantic arts

Data and Ontology in One Graph

Corporation

CoorporationID	Subsidiary Of	Official Name	CEO
3345223	NULL	Alphabet Inc.	Larry Page
5493006	3345223	Google Inc.	Sundar Pichai

Ontology as schema

Data

myco:LegalEntity

rdfs:subClassOf *rdfs:subClassOf*

myco:Corporation   **myco:Person**

rdf:type   *rdf:type*

rdf:type

One cell
One triple

myd:_3345223 — **myco:hasCEO** → **myd:_LarryPage**

myco:hasOfficialName → "Alphabet Inc."

myco:isSubsidiaryOf

myd:_5493006 — **myco:hasCEO** → **myd:_SundarPichai**

myco:hasOfficialName → "Google Inc."

Legend

| Class |
| Individual |

→ object property →
→ datatype property →

- Scripts
- R2RML
- Home-grown tools
- TARQL
- More tools becoming available all the time.

semantic arts

Data and Ontology in Same Store
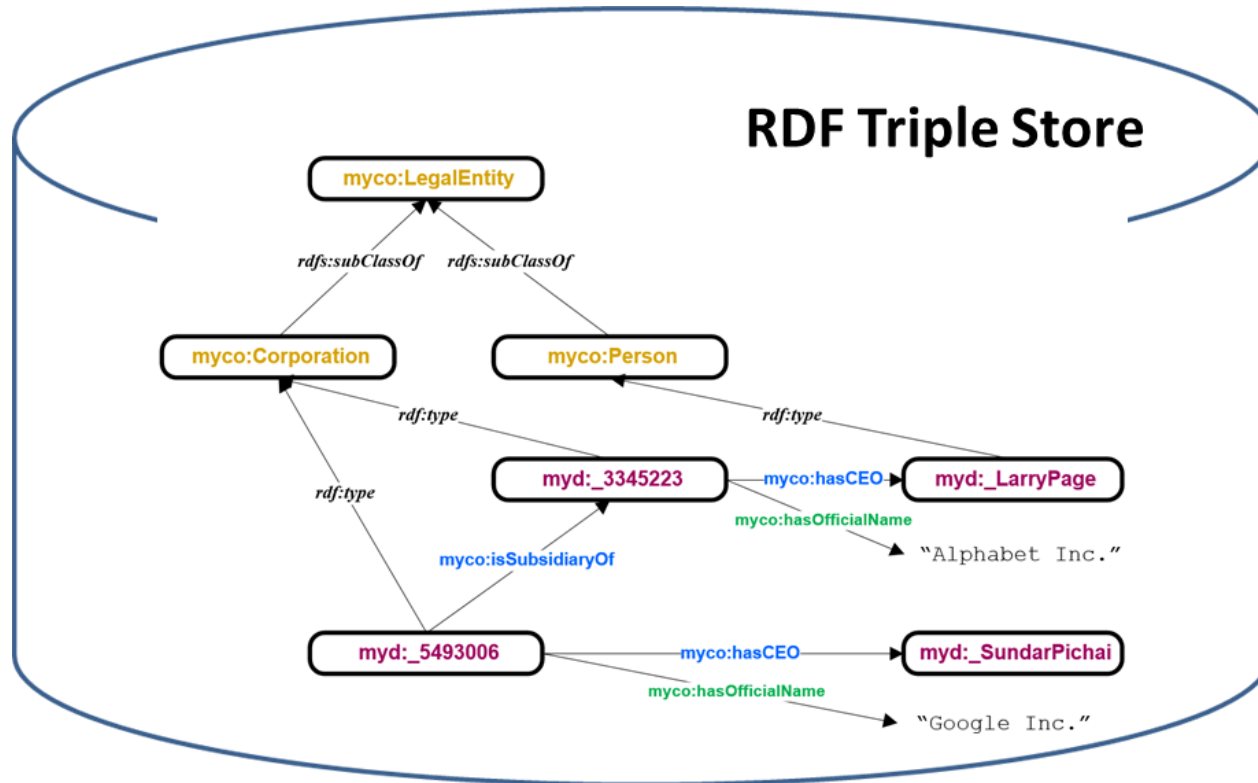
- Serves multiple applications
- Can be federated across multiple stores
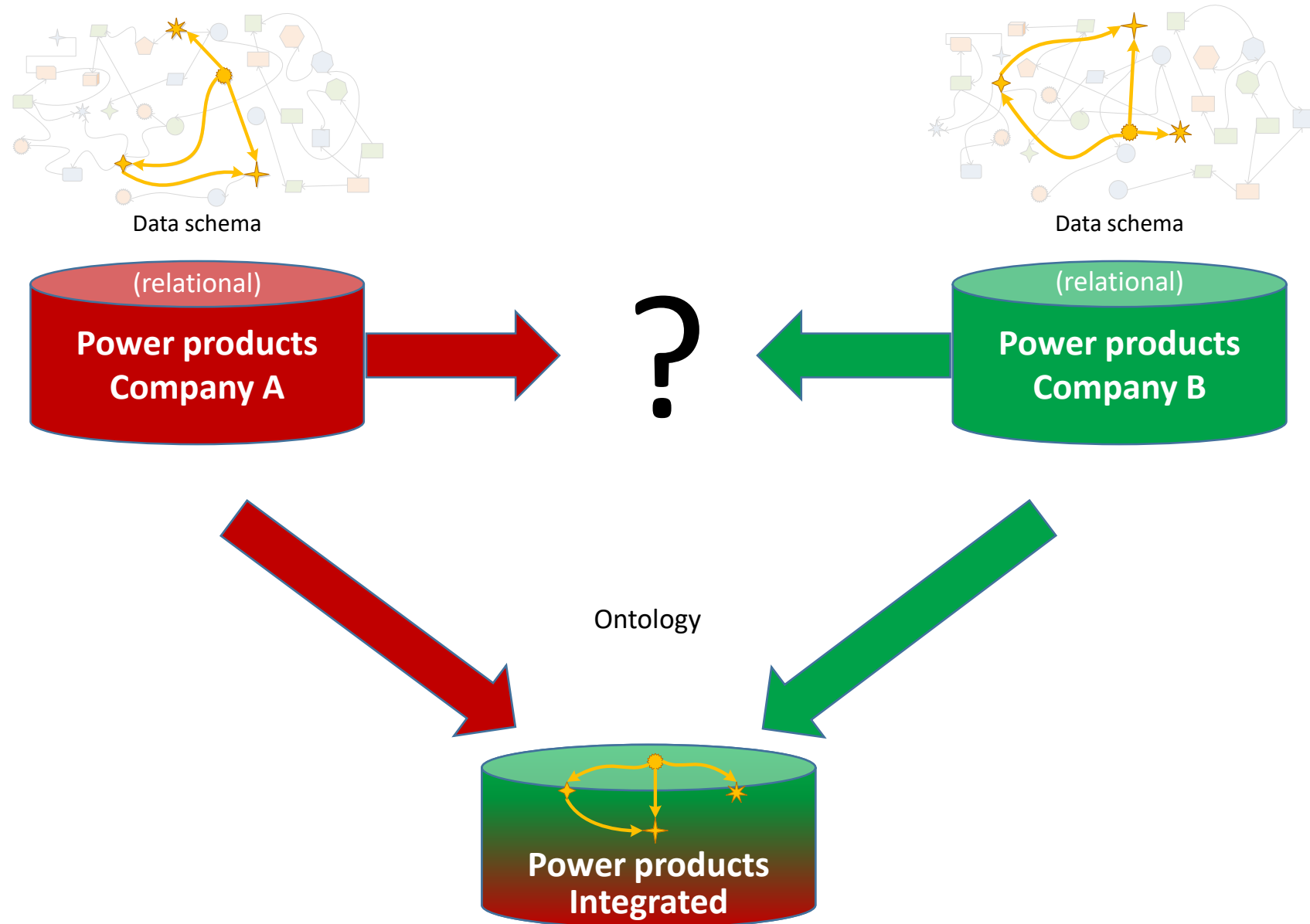- Data-centric rather than Application-centric

semantic arts

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.

semantic arts

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

semantic arts

Case Study: Database Integration

Data schema

Data schema

(relational)

**Power products
Company A**

?

(relational)

**Power products
Company B**

Ontology

**Power products
Integrated**

semantic arts

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

semantic arts

Modeling the Global Economy

IBB Is An Information Resource Used to Analyze:
[1] Companies, [2] Market Areas and [3] Industries

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.

semantic arts

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.

semantic arts

Experiences and Recommendations

- The Siren Call of Semantic Silos
- The role of SHACL
- Whither schema
- Change management

semantic arts

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

semantic arts

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.

semantic arts

Agile Creation of an Enterprise Ontology (2/2)

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.

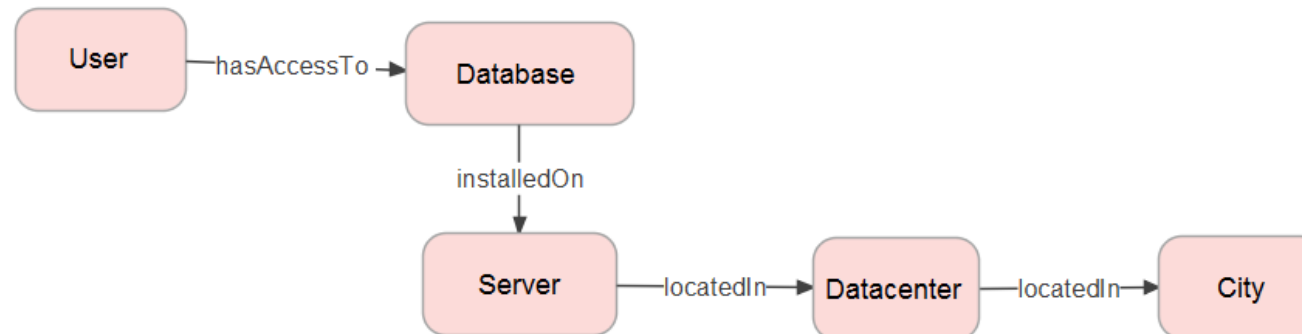- Extend existing and/or build out additional applications

Keep on Iterating

semantic arts

Modularity, Reuse & Federated Queries
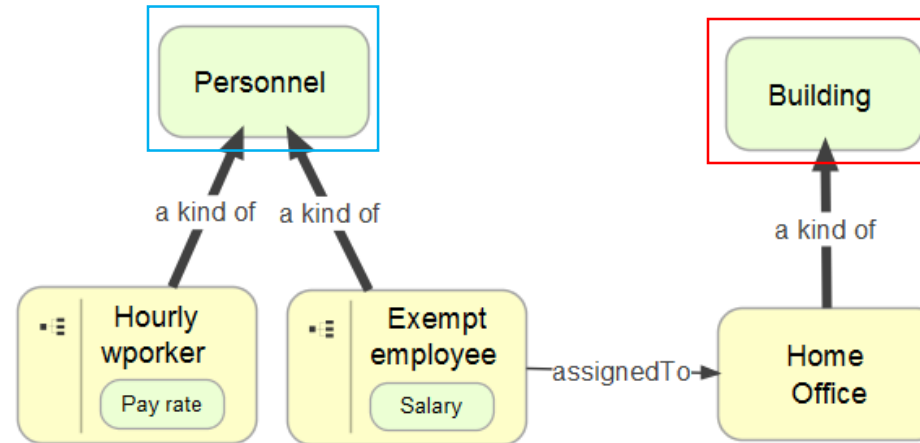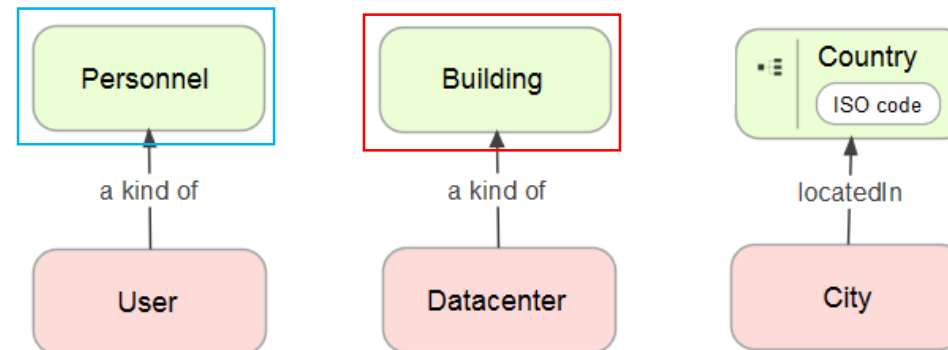
- HR Department:
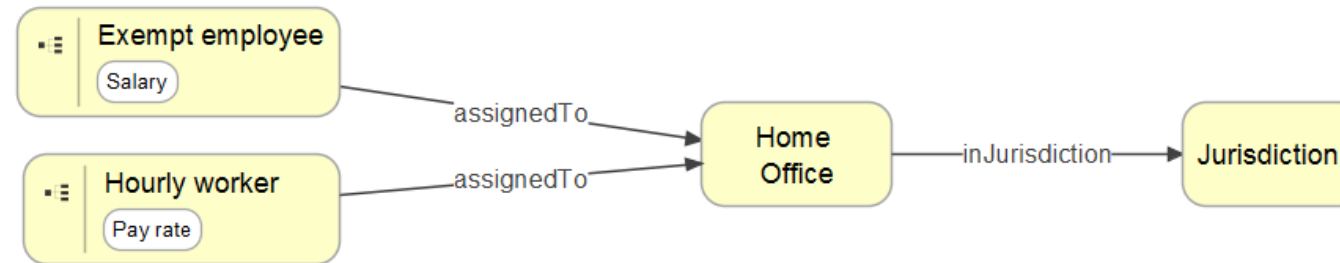
- IT Department:

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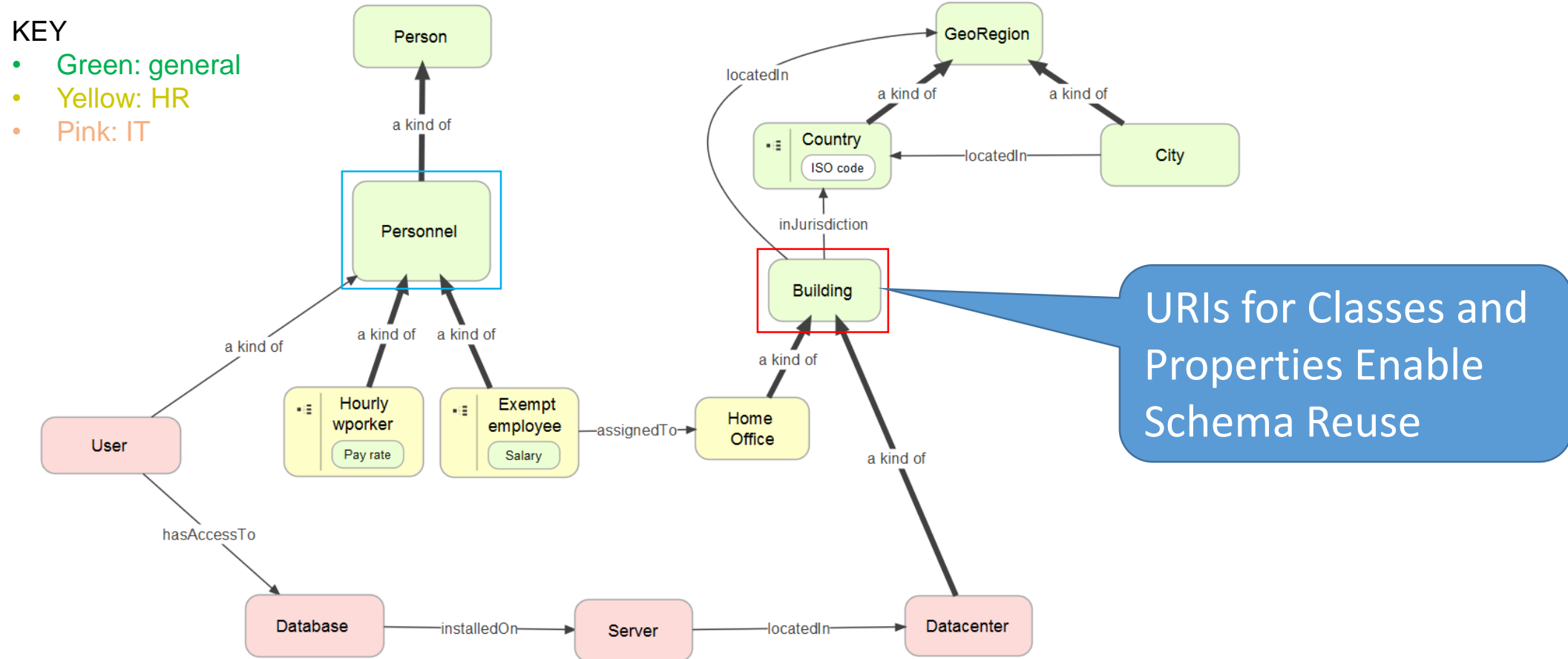
Example: connect to generic schema

- HR Department:

- IT Department:

Identical concepts will collapse into one.
- Personnel
- Building

semantic arts

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

semantic arts

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

semantic arts

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

semantic arts

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

semantic arts

What About Using SHACL <u>instead</u> 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

semantic arts

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.
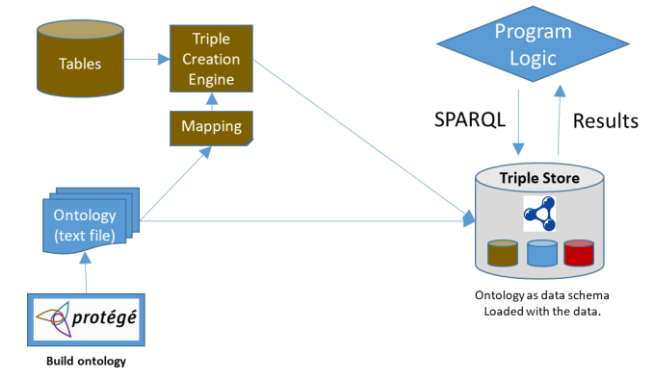
Summary & Conclusions

- 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

semantic arts