Some Thoughts from TopQuadrant: Cloud and SI

Framing the Conversation: Ontologies within Semantic Interoperability Ecosystems
TopQuadrant Company

- **Our Mission:**
  Empower people
  —by making enterprise information meaningful

- **Our Foundation:**
  TopQuadrant was founded on and continues a strong commitment to Semantic Web standards

- **Our Evolution:**
  - Services Company (2002)
  - Business Solution Company (2010 - Today)
TopBraid Suite seamlessly fits into existing IT environments by providing customer-specific approaches to complete semantic application development.
TopBraid and Solutions

- We are both a platform and COTS vendor
- EVN is a Web-based collaborative ontology development environment
- RDM is a Web-based collaborative reference data management application
  - Enterprise Data Governance is an extension to RDM
- However, we are not a triplestore vendor, so are agnostic wrt the RDF Database
  - decide on case-by-case basis
  - “Free option” is Jena SDB over MySQL, for example
TopBraid and the Cloud

- **Goal**: Run TopBraid products in a virtualized environment for private and public cloud deployments
- **Method**: Use Docker to build self-sufficient image which will contain everything needed for our Web-based platform and solutions
  - Docker is a lightweight container virtualization platform
  - Build image, run in container (VM that runs one process)
- Deploying to Amazon Elastic Compute Cloud pulls images from private Docker registry (server storing images)
- Map domain and subdomain as task name to separate customers
- **Summary**: Can build new app images with little setup, and deliver on private cloud or AWS
Customer and Project Experiences

- Customers not yet rushing towards the Cloud for their Semantic Web/Linked Data apps
- We use github for ontology development in customer projects, we seldom deliver only ontologies
- How we see things might develop
  - Vocabularies/ontologies/reference data linking data already on the cloud or being moved to the cloud
  - Customers interested in using Big Data and Smart Data together. There is a subtext of cloud giving power to the data scientists, information analysts and ontologists. Think the mainframe->mini-computer change.
  - Cloud as silo-breaker. Workflows can reach to partners and the whole value chain whereas today, internal IT restricts who can get to what
  - Cloud removes the IT from equation and the info owners can directly purchase the services they need. Faster.
Cloud challenges we hear

- Access control because data that integrates systems has lots of diverse access
- The hidden challenge of culture change
- Organisational trust – this is a very large barrier mentioned in several customer situations
- "across domains": Domains > borders > boundaries (be it a network/security/country/etc) further complicate the discoverability, availability and accessibility of Services/APIs
- Policy starts being your first and foremost concern with Cloud Service interop
- Handling data volume is still an engineering challenge. Issues include evaluating a SPARQL query over a larger volume (trillions), running a reasoner over a large volume. Truth maintenance. Splitting the volume across a cluster. Handling a query or inference over a SPO when the S and O are on different nodes of the cluster
Cloud ontology services

- The front facing website (better search), synonyms, data navigation.
- Business reporting - answers across data silos.
- One view of the customer (a finance world favorite)
- All the cloud services you inherit in a merger/acquisition
- Potential wrt ontologies and cloud:
  - Ontology data stored in the cloud (schema and rdf dataset aspects) provided in a datastore that can replicate across cloud regions to help with performance on the user end. But this is no different than what happens with databases.
  - An ontology, with rules, and a reasoner, to help drive a system of cloud services. New service instances may be spun up in the cloud, or brought down, when certain conditions are met.
Semantic Interoperability in Engineering Industries
Application example 1: ERH

- We operator EPIM Reporting Hub in a secure network for Norwegian Oil and Gas operators
  - A sort of cloud, providing services to all operators and partners
  - Large OWL database at its core with data transformed from XML to ISO 15926-based OWL for storage

- Heavyweight Semantic Interoperability through deep semantic modelling and large scale data integration repository
Semantics and 15926

- ERH project is The Semantic Web meets ISO 15926
  - RDF/OWL, ISO 15926, SPARQL query, reasoning all mentioned in requirements
  - ISO 15926 upper ontology designed for data exchange with no domain content, that comes as reference data
  - Blending OWL and bits of ISO 15926 that are a modeling language is difficult
  - Existing reference data incomplete for ontology building (e.g. no reference data connection between what in OWL are Classes and related Properties)

- For ERH requirements, an approach closer to the structures in 15926-2 data model made more sense, and the vocabulary required is actually relatively small
  - Drilling : Activity
  - Production : Activity, Flow (Stream) and Product (Material)
  - and properties are about temporal parts of these
Approach

- Provenance is handled by modeling the Report itself and linking all Activity/Flow instance to the Report
  - no need for reification of classification, composition
  - This was prior to the W3C Prov-O ontology existing

- ISO 15926 complexity -> preference for expressiveness in modeling capability over decidability
  - But keep interest in reasoning in future in mind
  - Ontology architecture and partitioning named graphs should mean we can satisfy reasoning requirements
  - If not, then SPIN/SPARQL rules or transform then DL reasoner (or both)

- SPARQL is the phase 1 query language

- Under the covers, the triples from a report are maintained using named graphs
  - This is also how access control is managed
Example ERH data

NPD Fact

Wellbore on a day is a temporalPartOf whole-life Wellbore

Daily activity happens on TemporalPartOfAWellbore
Application example 2: V-Con

- Virtual Construction for Roads (V-Con) is a project used by National Road Authorities (NRA) to manage the maintenance and deliver information about traffic and infrastructure on road networks.

- Netherlands and Sweden government ministries are the sponsors (i.e., the NRAs) and their key business processes are:
  - Asset management
  - Management of network usage

- A pre-commercial R&D project in 3 phases with competitors down-selected at each phase:
  - Phase 2 in-process now
V-Con Sys Eng Use Case

of the roles and transactions of the ‘systems engineering’ use case
3 Business Scope

3.1 Business objective, business process and business objects

The objectives of a National Road Authority (NRA) relate to the main road network and are summarised as:

- To enable smooth, safe and sustainable traffic on roads;
- To deliver reliable and useful information about the traffic and the infrastructure.

To realise the usage process objectives, an NRA provides the following services (functions) to users:

1) Safe and sufficient road space;
2) Monitoring of the current state of traffic and infrastructure and execution of static and dynamic traffic and infrastructural measures. These measures ensure that the current state of traffic and infrastructure better match the desired situation (control loop).

The aforementioned services are realised through a combination of business objects and business processes.

Within the physical living environment, the following (types of) business objects are important to an NRA:

- infrastructural systems:
  - for the movement of persons and goods
- structures:
  - buildings (traffic control centre, service buildings)
  - structures other than buildings (junction structures, other engineering structures)
- areas:
  - functional areas
  - administrative areas
4.3 Major features

The V-Con Solution shall address the following major features:

- Integration of different tools and standards in the area of Building Information Modelling (BIM), Geographical Information Systems (GIS), Systems Engineering (SE) and Product Life-Cycle Support (PLCS).
- Support for open standard-based information management and data exchange for different types of infrastructure projects during the entire life cycle for road infrastructure.
- Support for open standard-based data exchange with external stakeholders, such as construction companies, and internal stakeholders, such as the asset management department.
  - Use exchange standards wherever possible.
  - The standards to be used are defined by V-Con.
- Flexibility with regards to future development of NRA asset management system and relevant open standards for data exchange within this business area.
  - As these asset management systems and open standards are under development, the V-Con Solution should be flexible with regards to future developments.
- Support for integrating and mapping between standards and semantics defined on an international, country, company or project level.
- Support for the generic business model described in Chapter 5.
- Support for the business use cases described in the annexes.
V-Con Key Technical Challenges

- TC1: Support Ontology-based handling of datasets
- TC2: Support handling datasets in non-linked data formats
- TC3: Manage and store data structures and datasets
- TC4: Connect datasets from different domains
- TC5: View (connected) information
- TC6: Ensure system quality
- TC7: Ensure a future proof system
V-Con and SI

- Lightweight Semantic Interoperability through “context” and “common” ontologies and “linking rules sets” that relate them
- Datasets are left in context ontology form and only the scope that overlaps between contexts is modelled in common ontology
- Use reasoners, rule engines and humans to make linksets that specify links (e.g. that Bridge 1 in Dataset 1 based on Ontology 1 is actually the same as Bridge 2 in Dataset 2 based on Ontology 2)
Figure 4: A look inside the V-Con Solution
Cloud is one option we are considering for V-Con
  – We know some very large Building and Construction projects operate in the Cloud though
  – We’re told these NRAs are conservative organizations though, so acceptability is unclear

Clearly, the V-Con approach is a “style” of SI and we are researching how we might support that style more generally

Also considering: if ISO 15926 ontology (e.g. ERH) was the “Common Ontology”, then does the heavyweight vs lightweight approach just become another configuration option?
The End

Thanks for your attention 😊

dprice at topquadrant dot com