Ontology Summit 2014:
Big Data and Semantic Web Meet Applied Ontology
Track A: Common Reusable Semantic Content

Session 8: 6 March 2014
Experiences in Knowledge Sharing: Lessons from research and experience in Big Data, Linked Data and Semantic Web Applications

Andrea Westerinen, Gary Berg-Cross and Mike Bennett

Co-Champions

Track A Mission Statement

- Semantic technologies such as ontologies and reasoning play a major role in the Semantic Web.
- One challenge in these efforts is to build and leverage common semantic content thus reducing the burden of new ontology creation while avoiding silos of different ontologies.
- However, crafting of whole or partial common semantic content has long presented challenges. Achieving commonality and reuse remain key ingredients for practical development of quality and interoperable ontologies.
- Despite development of such things as foundational top-level ontologies and the availability of broad domain models as starting points, the amount of reuse seems quite low in practice.

Progress

- Inputs:
  - Track A presentations from Jan 23 2014
  - Email dialogs
  - Co-Champion discussions & community input page
- Track A Goals:
  - Define/document:
    - Explicit conditions for and issues with reuse
    - Concepts/meta-ontology
    - Approaches to modularization and best practices
    - Specific design patterns and exemplary content
      - For content reuse in applied ontologies and semantic web/linked data, and for reasoning and big data
  - Expand tooling, such as OOR, to support defining and finding reusable content

Main Take-aways to Date

- Conditions for Re-use of ontologies
  - General conditions
    - Accessibility, trust, compatibility of microtheories etc.
  - Specific requirements by use case
    - Reasoning: are existing ontologies built with sufficient semantics to support reasoning applications
    - Conversely, do the constraints for reasoning hold for Linked Data? Big Data?
- Ontology Design Patterns
  - Exemplar: Event patterns (and how these fit to micro-theories)
- Tooling
- Governance
- Best Practices
  - Modularity
  - Annotation
- The Search for Primitives

Reuse 

- Reuse issues not unique to ontologies/schemas
  - Parallels and differences with software and data model reuse
- Capture and understand range of conditions, contexts and intended purposes for which an ontology/linked data is "safely" and productively reused
  - Confirm/track that reused content works “as expected” in new contexts
- Understand dimensions of variability and affects on modularity and reuse
  - Variability across contexts (for ex, concept or property present or absent in different contexts/uses)
  - Variability over time (evolution of a module and need to take current trends and future directions into account)
- Separate reuse of classes/concepts, from properties, from individuals and from axioms
  - Easier to target what is possible to reuse and reduces amount of transformation and cleaning

Seven Questions

1. How can we characterize or measure semantic content reuse, both between ontologies and by Big Data and Semantic Web communities?
2. What building blocks of common semantic content exists now to enable interoperability?
   - What additions are needed to move forward and how are these best achieved?
3. What is involved in reuse of Linked Data versus reuse of ontologies?
4. What is an example of a small set of semantic content that the community might propose for reuse?
   - Is there agreement on these or things like ODPs as building blocks?
5. What is an example of a large set that the community might propose for reuse?
6. Is it reasonable to expect reuse of an entire ontology like DOLCE and Semantic Sensor Network (SSN)?
   - If so under what conditions might this be reasonable?
   - Is it better to expect alignment rather than exact content reuse?
7. Is reuse about semantics alone or should it also address reasoning and data analytics?

Session Overview

- In this session we aim to bring together the lessons learned to date and new insights on the re-use of ontology content.

- We have three speakers lined up to share their perspectives on ontology re-use:
  - Historical perspective: sharing of knowledge among computers and humans
  - Linked Open Data (LOD): making sense of the data; formalizing for deductive reasoning
  - Strategies for reconciling heterogeneous data using shared ontologies

- These presentations will be followed by an open discussion:

  - Considerations and requirements for those wanting to identify and use semantic resources in each of these contexts

  - Things to think about when developing ontologies for re-usability

Session Agenda

- "Historical Perspectives on Projects for Knowledge Sharing"
  - Dr. John Sowa (VivoMind Intelligence)
- "Tactical Formalization of Linked Open Data“
  - Professor Michel Dumontier (Stanford BMIR)
- "Ontology Driven Data Virtualization“
  - Mr. Kingsley Idehen (OpenLink Software)
- Discussion