Ontology Summit 2014 Session 6
Synthesis I:
Approach to the Ontology Summit
2014 Communique and Proposed
Draft Outline
Thursday, February 20, 2014

General Chairs: Michael Gruninger (U. Toronto) Leo Obrst (MITRE)

# Draft Communique: Overall Structure & Questions

- We will generally follow the Track structure
- Summit Premise:
  - The Semantic Web and Big Data communities can bring a wide array of real problems and technologies (e.g. performance and scalability challenges, automated reasoning tools), while the Applied Ontology can bring a large body of content (i.e. ontologies) and ontological analysis techniques
- Each section will end with a set of research questions/problems which can serve to guide and direct future work. Examples include:
  - What are the obstacles to closer collaboration among the Semantic Web, Big Data, and Applied Ontology communities?
  - What ontologies are required by Semantic Web applications?
  - If these ontologies are not available in any current ontology repository, how can we engage the applied ontology community to develop them?
  - Can we apply existing analysis techniques to ontologies that are being widely used within the Semantic Web community?
  - What language expressiveness is required by the ontologies that are being developed by the applied ontology community?
  - What role is play by decidability and tractability in applications of ontologies?
  - How can Big Data leverage existing ontologies?
  - What requirements do problems encountered with Big Data impose on the design of ontologies?

### Draft Communique: Brief Outline

- 1. Introduction
- 2. Challenges
- 3. Sharable Reusable Content
- 4. Tools and Techniques for Ontologies on the Web
- 5. Bottlenecks in Ontology Engineering
- 6. Variety in Big Data, Linked Data, and Linked Open Data
- 7. Recommendations

References (a small number)

Endorsement

## Draft Communique: More Detailed Outline (1)

- Introduction
- 2. Challenges
- 3. (Track A) Sharable Reusable Content
  - A. Goals
  - B. Reuse
  - C. Conditions for Reuse
  - D. Ontology Design Patterns May be Useful
  - E. Big Data Landscape
  - F. Big Data Vocabularies Need Semantics
  - G. Reasoning is a Research Question?
  - H. Tooling
  - I. Best Practices

# Draft Communique: More Detailed Outline (2)

- 4. (Track B) Tools and Techniques for Ontologies on the Web
  - A. Opportunities and Challenges
  - B. Services:
    - 1. Increase pool of knowledge by tapping into the whole Web
    - 2. Example of Watson: search heterogeneous knowledge sources for evidence, hybrid/heterogeneous reasoning, generation of hypotheses, evidence retrieval (disambiguation) and scoring (machine learning, statistical models)
  - C. Techniques
  - D. Tools
  - E. Issues:
    - A. Variety, heterogeneity of information types, schemas, software, etc., and hybrid approaches
    - B. Ontologies for annotating Big Data or representing it?
    - C. Creation of ontologies from data: mining, machine-learning
    - D. Relations among formalisms not well understood: RDF/S, OWL, SPARQL, UML, Rules, First-Order Logic languages (e.g., Common Logic)
    - E. Tool limitations
    - F. Requirements: demonstration, human factors, exemplars

# Draft Communique: More Detailed Outline (3)

### 5. (Track C) Bottlenecks in Ontology Engineering

#### A. Bottlenecks

- Ontology engineering processes that are time consuming
- Social, cultural, and motivational issues
- Modeling axioms or knowledge representation language fragments that cause difficulties in terms of an increase in reasoning complexity or reducing the reuseability of ontologies
- 4. Identifying areas and applications that would most directly benefit from ontologies but have not yet considered their use and development

#### B. Potential Solutions

### C. Important Questions

### D. Important Findings

- Behavioral Abstraction
- 2. Entity-centric, frame-oriented data science
- 3. Improved data-driven techniques to scale the development of patterns and ontologies
- 4. KR choices have consequences for reasoning complexity, tool support, reusability
- 5. Unlimited patterns: discovered, not invented; bottom-up ontology development
- 6. Education & Buy-In: time, utility, ease of availability and use; support the requirements

## Draft Communique: More Detailed Outline (4)

- 6. (Track D) Variety in Big Data, Linked Data, and Linked Open Data
  - A. Introduction to Variety in Big Data, Linked Data, and Linked Open Data
  - B. Example Use Cases:
    - 1. Industrial Use of Ontologies for Big Data: OODA
    - 2. Using ontologies to Manage Biodiversity Data
    - 3. Using Ontologies to Manage Data About Ice

#### C. Issues:

- 1. Ontology Reuse
- 2. Automated Ontology Gap-Filling (Gaps in Ontologies)
- 3. Evolution: Dynamic Ontologies and Adaptation
- 4. Crowdsourcing Curation
- 5. Building Ontologies from Small Modules
- 6. Working with Existing Datatypes
- 7. Employing Multiple Languages
- 8. Data/Metadata Annotation and Semantic Tagging
- 9. Ontology Mapping
- 10. Adaptation to Existing Workflows of Domain Experts
- 11. Machine-learning Algorithms
- 12. Tool Incompatibility

### Draft Communique: More Detailed Outline (5)

#### 7. Recommendations

References (a small number)

**Endorsement** 

### **Emerging Common Themes & Issues**

- 1. Ontology Reuse
- 2. Automated Ontology Gap-Filling (Gaps in Ontologies)
- 3. Evolution: Dynamic Ontologies and Adaptation
- 4. Crowdsourcing Curation
- 5. Building Ontologies from Small Modules
- Working with Existing Datatypes
- 7. Employing Multiple Languages
- 8. Data/Metadata Annotation and Semantic Tagging
- 9. Ontology Mapping
- 10. Adaptation to Existing Workflows of Domain Experts
- 11. Machine-learning Algorithms
- 12. Tool Incompatibility
- 13. Ontology Design Patterns
- 14. Large-scale Reasoning
- 15. Time-consuming KR Processes
- 16. Education & Buy-in
- 17. Variety, Heterogeneity, and Hybrid Methods