Why Ontologies?  
A Perspective from Military Intelligence

Ontologies for Information Integration Workshop, SUNY-Buffalo  
Thursday, April 18, 2013

Dr. Leo Obrst  
Information Semantics
Information Discovery & Understanding  
Command & Control Center  
MITRE
Towards Solving the Problem

- With the increasing complexity of our systems and our analytical needs, we need to go to **human level interaction**
- We need to **maximize** the amount of **Semantics** we can utilize
- From data and information level, we need to go to **human semantic level interaction**
  
  **Semantic integration: Ontology is the key!**
The Solution

• We need to offload the very real, heavy cognitive interpretation burden from humans to our systems.

• We need to represent human semantics using machine-interpretable ontologies.
Advancing Along the Interpretation Continuum

**Human interpreted**
- Relatively unstructured
- Random
- Info retrieval
- Web search

**Computer interpreted**
- Very structured
- Logical
- Text summarization
- Content extraction
- Topic maps
- Reasoning services
- Ontology induction
- Automatically acquire concepts; evolve ontologies into domain theories; link to institution repositories (e.g., MII)
- Automatically span domain theories and institution repositories; inter-operate with fully interpreting computer

**Simple Metadata:**
- XML

**Richer Metadata:**
- RDF/S

**Very Rich Metadata:**
- OWL, CL, FOL

Amount of time analyst spends decyphering the data

Moving to the right depends on increasing automated semantic interpretation

Amount of time analyst spends analyzing information & knowledge
Semantic Integration Implies Semantic Composition

- signifies the composition operation

**Simple Procedure Integration & Composition**
Concatenation, alignment of calling Procedure with called procedure:

- Caller: Do_this (integer: 5, string: “sales”)
- Called: Do_this (integer: X, string: Y)

**Simple Syntactic Object Integration & Composition**
Alignment of embedded interface definition language statements mapping two CORBA, Javabean objects

**Simple Semantic Model, Knowledge Integration & Composition**
Unification of tree or graph structures, with reasoning, simple Semantic Web ontologies:

**Complex Semantic Model, Knowledge, System Integration & Composition**
Unification of complex networks of graph Structures, with complex reasoning, complex Semantic Web ontologies:

Copyright 2013, The MITRE Corporation, Leo Obrst
Dimensions of Interoperability & Integration

Our interest lies here

3 Kinds of Integration

Syntactic

Structural

Semantic

Interoperability Scale

0% 100%
Intelligence Needs

- Information Extraction, Machine Translation, & other NLP
- Metadata Tagging/Indexing
- Database Technology
- Specific COTS applications tailored to solve specific problems in proprietary ways
- Information Infrastructure
- Information Sharing Initiatives/Paradigms
- Ontologies can assist these technologies, applications, infrastructure, and emerging information sharing initiatives/paradigms
Intelligence Problems

- Command and Control, Situational Awareness, Information Fusion across all levels
- Preventing Incidents, Crises: terrorism, WMD, disease outbreaks
- Crisis Management, Containment/Amelioration
- Risk Assessment, Strategic Assessment, Tactical Assessment
- Social Networking and Link Analysis
- Gathering Evidence
- Requires Knowledge from Multiple Intelligence Disciplines
- Requires Knowledge from Multiple Domains:
  - Cultural, linguistic, geophysical, geopolitical, electronic and communication devices, logistics, medicine, chemicals, nuclear energy, biology, physics, sociology, metallurgy, construction, machinery, weather, public health, emergency response, reconnaissance, espionage, weapons and armaments, electromagnetic spectrum, etc.
  - What domains are NOT included?
  - Temporal Reasoning
  - Bill of Materials, Part Composition/Decomposition, Distribution Tracking
Why Ontology Matters for Intelligence

- Ontology provides the foundation for epistemology (belief, hypothesis, emerging knowledge)
- Data percolating up from multiple sources: prone to multiple kinds of errors (source reliability, accuracy, etc.)
- Analysts must semantically interpret data using their own complex mental models about the world *today with very little machine assistance*
- Analysts must hypothesize, assign strengths of belief to potential knowledge assertions
- Ontology can help!
- Example 1:
  - Francis Bifflestick is a person
  - Persons have only 1 birth date
  - Francis Bifflestick has 1 birth date
  - Francis Bifflestick has 3 birth dates from the evidence
  - However, only 1 of those 3 or some other date is his correct birth date
  - Ontology represents Person, property of HavingBirthDate with cardinality = 1
Why Ontologies Matter for Intelligence (2)

• Ontology substrate:
  – What’s true of the world, best existing theories (scientific and commonsense)
  – These include classes (universals), instances (particulars), but also subclasses, arbitrary relations among subclasses, properties, value/property ranges, axioms and rules (can be considered meta-relations because they relate classes, relations, properties across the ontology), etc.
  – Machine interpretable semantics, i.e., software can use ontologies, make inferences comparable to what humans can make
  – Machines can assist humans at human conceptual level
• Evidence, belief, hypothesis, can be built on that ontology scaffolding
  – Build on what you know is generally true
  – Semantically describe instances from different perspectives
  – Example: a book can be coerced to be a weapon because it is also a physical object (not just an information bearing object)
Intelligence Concerned with Identity Management

- Who is that Person?
- What is that Thing?
- Where did that Person or Thing go?
- When did it occur?
- How do we know these facts?
- Are they facts?
- How do we get better information?
- Are these two persons, things the same?
Mainstream Information Technology cannot deal with these issues

• SOA, database technology, big data, cloud computing, programming cannot address

• Ontologies can address these issues
  – Formal ontological analysis:
  – Theory of Parts
  – Theory of Wholes
  – Theory of Essence and Identity
  – Theory of Dependence
  – Theory of Qualities
  – Theory of Composition and Constitution
Ontologies Increasingly Used

• More large-scale efforts taking place in government
  – COI vocabularies and ontologies to support information sharing
  – Many projects: develop coherent processes, procedures, education, modeling in OWL

• Example: Intelligence
  – The community realizes this is a large problem
  – Is strongly addressing problems of ambiguity, coreference
  – For people, can’t count on names!
  – Everyone tries to use names like object identifiers: you can’t do that!
  – Need to know about things too: material, organizations, systems, cargo, components, logistics, events, times, locations, features, relationships, attributes, qualities, etc.
Zooming in on the Search & Discovery Problem

• Today, automatically discovering relationships among people and other entities of interest in Web pages is very limited for free text search, enterprise search, and social networking applications
• Not enough patterns of discovery are available for finding relationships automatically
• There are a shortage of tools that can use those patterns
• Traditional discovery technologies require substantial processing, e.g., information extraction, question-answering
• Less precise search, discovery, and navigation cannot accurately discern:
  – Expertise and shared interest; communities & their dimensions
  – Geospatial and temporal information
  – Intelligence analytics: link analysis; identity and correlation management; participation in events; trust analysis
What Problems Do Ontologies Help Solve?

✓ Heterogeneous database problem
  – Different organizational units, Service Needers/Providers have radically different databases
  – Different **syntactically**: what’s the format?
  – Different **structurally**: how are they structured?
  – Different **semantically**: what do they mean?
  – They all speak different languages (access, description, schemas, meaning)
  – Integration: rather than $N^2$ problem, with single, adequate Ontology reduces to $N$

✓ Enterprise-wide system interoperability problem
  – Currently: system-of-systems, vertical stovepipes
  – Ontologies act as conceptual model representing enterprise consensus semantics

✓ Relevant document retrieval/question-answering problem
  – What is the meaning of your query?
  – What is the meaning of documents that would satisfy your query?
  – Can you obtain only meaningful, relevant documents?
Blind Men and the Elephant

Blind Men and the Elephant

* A Jain version of the story says that six blind men were asked to determine what an elephant looked like by feeling different parts of the elephant's body.

The blind man who feels a leg says the elephant is like a **pillar**; the one who feels the tail says the elephant is like a **rope**; the one who feels the trunk says the elephant is like a **tree branch**; the one who feels the ear says the elephant is like a **hand fan**; the one who feels the belly says the elephant is like a **wall**; and the one who feels the tusk says the elephant is like a **solid pipe**.

A wise man explains to them

**All of you are right.** The reason every one of you is telling it differently is because each one of you touched the different part of the elephant. So, actually the elephant has all the features you mentioned.

Thanks!