Ontology Summit 2016: Healthcare Track Synthesis
April 14, 2016

Co-Champions:
Ram D. Sriram (NIST)
Leo Obrst (MITRE)
# Levels of Biological Information

## HEALTH CARE

- Ecologies
- Societies/Populations
- Individuals
- Organs

## BIOSCIENCES

- Tissues
- Cells
  - Protein and gene networks
  - Protein interaction networks
  - Protein
  - mRNA
  - DNA
Standards That Link to an EHR
Standards in HealthCare

- **Terminology**
  - Systematized Nomenclature for Medicine – Clinical Terms (SNOMED-CT), Logical Observation Identifiers Names and Codes (LOINC), RxNorm (normalized names for clinical drugs)

- **Classification Systems**

- **Devices**
  - IEEE 11073

- **EHR-Related**
  - Digital Imaging and Communications in Medicine (DICOM), Health Level Seven (HL7) Clinical Document Architecture (CDA)

- **Interoperability**
  - DICOM, HL7 Messaging, Health Insurance Portability and Accountability Act (HIPAA) Transactions, National Council for Prescription Drug Programs

- **Language Formats**
  - XML, X12

- **Internet Protocols**
  - HTTP/HTTPS → TCP/IP

- **De Facto Ontology Standard**: Open Biological & Biomedical Ontologies (OBO)
Biomedical Ontologies: Categories of Use

• Knowledge Management
  – Indexing and retrieval of data and information, access to information, mapping among ontologies

• Data Integration
  – Exchange and semantic interoperability

• Decision Support and Reasoning
  – Data selection and aggregation, decision support, natural language processing, knowledge discovery

Bodenreider, YBMI 2008
Semantic Interoperability/Integration Definition

- To *interoperate* is to participate in a *common purpose*
  - operation sets the context
  - purpose is the intention, the end to which activity is directed

- Semantics is fundamentally *interpretation*
  - within a particular context
  - from a particular point of view

- Semantic Interoperability/Integration is fundamentally driven by *communication of purpose*
  - participants determined by interpreting capacity to meet operational objectives
  - service obligations and responsibilities explicitly contracted
Towards Semantic Interoperability & Integration

• Semantic Integration is enabled through:
  – Establishing base semantic representation via ontologies (class level) and their knowledge bases (instance level)
  – Defining semantic mappings & transformations among ontologies (and treating these mappings as individual theories just like ontologies)
  – Defining algorithms that can determine semantic similarity and employing their output in a semantic mapping facility that uses ontologies

• The use of ontologies & semantic mapping software can reduce the loss of semantics (meaning) in information exchange among heterogeneous applications
Healthcare Track: **Goal and Mission**

- **Healthcare and biomedical domains:** a very large ecosystem of vocabularies and ontologies
- **Goal:** To provide insight into the nature of semantic interoperability issues and approaches in the Healthcare ecosystem
- **Mission:** Present issues of semantic interoperability and integration in domain of healthcare
  - Discuss approaches in semantic interoperability and integration in healthcare: precoordination, postcoordination, mapping, anchoring, foundational ontologies and architectures, hybrid approaches
  - Discuss approaches for achieving semantic interoperability of vocabularies and ontologies in healthcare
  - Discuss gaps in current approaches in semantic interoperability in healthcare
  - Discuss current/future challenges and prospects in semantic interoperability in healthcare
  - Propose best methods for achieving semantic interoperability in healthcare
Brief Summary of Talks

• Need ontologies from multiple sources for any practical application
• Current ontologies, such as SNOMED, have inconsistency issues
• Proposed methodologies like FHIR show some promise, but need further work
• Methodologies for ontology development were presented
• Several useful use cases presented
• Electronic Health Record (EHR): coherent, semantic representation sorely needed
Healthcare Track: **March 10**

- Kincho Law (Department of Civil and Environmental Engineering, Stanford University, USA):
  
  An Ontology-Based Approach for Facilitating Information Retrieval from Disparate Sources: Patent System as an Exemplar

- Olivier Bodenreider (US National Library of Medicine, USA):
  
  Ontologies Support Semantic Interoperability in Healthcare

- Eric Little (OSTHUS, Inc. USA):
  
  Integrating the Laboratory for Improved Analytics: How Semantics Is Driving an Evolution of 21st Century Labs

- Parsa Mirhaji (Clinical Research Informatics, Montefiore Medical Center, Albert Einstein College of Medicine, USA):
  
  The Semantic Data Lake in Healthcare
Healthcare Track: April 7, 2016

• Keith Campbell (US Veterans Health Administration (VHA) Knowledge-Based Systems, USA):
  – SOLOR: Building a coherent foundation for clinical data by leveraging the SNOMED/LOINC Cooperative Agreement, and integrating RxNorm

• Barry Smith (U-Buffalo, NY, National Center for Ontological Research, USA):
  – An Ontology Ecosystem Approach to Electronic Health Record Interoperability

• Amit Sheth (Kno.e.sis, Wright State University, USA):
  – Ontology-enabled Healthcare Applications exploiting Physical-Cyber-Social Big Data

• Leo Obrst (MITRE, USA)
  – Semantic Interoperability in General and in Healthcare: Clinical Care
BACK UP SLIDES
Semantic Data Lake ➔ Parsa Mirhaji

- Platform for Big Data Analytics and Cognitive Computing
- Based on principles of HDFS/Spark/GraphX, and Semantic Web
- Knowledge Based (metadata, terminology, ontologies)
  - Enables linking Data -> Knowledge -> Linked Open Data
- URI level security, authorization, provenance
- Optimized for integration of heterogeneous health and well-ness data
  - Structured and non-structured data
- Sophisticated temporal and geospatial analytics, Machine Learning
- To support basic science, translational research, operational care coordination, population health management (ACO)
- A Montefiore, Franz, Intel, Cisco Collaboration
The Semantic Data Lake
SDL Knowledgebase
(UMLS Vocabularies and Semantic Network Linked to OMOP Terminology)
Cross Vocabulary Mapping
Ongoing projects supported by SDL at Montefiore
Patent Validity and Infringement/Enforcement Questions involves analysis of documents in various domains – Patents, USPTO File Wrappers, Court Documents, Scientific/Technical Publications, and Technical Product Literature

- Owned by disparate public (government) and private sectors
- The information is often available online, but siloed into several diverse information sources
- Today, the analysis is done manually and poorly by companies offering various patent research and strategy services
Summary: BIO-REGNET

Knowledge Source:

- Bio Ontology (Technical Domain)
- Patent System Ontology (Business/Legal Domain)

Siloed Patent System Information

- Issued Patents and Applications
- Court Cases
- Regulations and Laws
- Technical Publications
- File Wrappers

Integration

Biportal (bioportal.bioontology.org)
NCI Thesaurus

Scientific Publication

Patent Document

Court Case
Summary and Discussion

- IP informatics: from research/development, patent filings to infringement and IP protection
- Knowledge-Driven Ontology-Based Approach
  - Technological ontologies
  - Patent system ontology
- Generalization – Linking to other information sources – technical/scientific publications, product literature
- User Interface – Efficient presentation of relevant (semantic) information
- Comparative analysis of documents
- Scalability (Graph Database?)
- Experiment with more use cases in other technical domains outside of the biomedical domain
Integrating the Laboratory for Improved Analytics → Eric Little

• Discuss the current status of laboratories – data is siloed, machines are not integrated, numerous file types exist, and information is difficult to retrieve and utilize over time with any consistency

• Show how utilizing semantics can provide researchers with improved ways to query and understand their data from multiple perspectives by providing a common reference layer – show specific examples of how the Allotrope Foundation is solving this problem

• Discuss how semantic technologies such as those developed under Allotrope can be extended to provide advanced analytics capabilities that bring together both logical and mathematical approaches to graph theory, knowledge discovery and big data analytics

• Briefly discuss how this can impact the entire Life Sciences industry and drive real business value to a wide variety of customers
1. Information Exchange: Use of value sets for Meaningful Use

2. Analytics: Analysis of large distributed clinical data warehouses
Information exchange

• “Meaningful Use” incentive program
  – Use of certified electronic health record (EHR) systems
  – Requires use of select biomedical terminologies
    • For information exchange (e-prescribing, lab results)
    • For quality purposes (clinical quality measures)
      • Rely on reference value sets
Reference value sets

<table>
<thead>
<tr>
<th>eMeasure Title</th>
<th>Venous Thromboembolism Prophylaxis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>This measure assesses the number of patients who received VTE prophylaxis or have documentation why no VTE prophylaxis was given the day of or the day after hospital admission or surgery end date for surgeries that start the day of or the day after hospital admission.</td>
</tr>
</tbody>
</table>

- \( \text{$MedicationVTEProphylaxis} = \)  
  - Union of:  
    - "Medication, Administered: Low Dose Unfractionated Heparin for VTE Prophylaxis"  
    - "Medication, Administered: Low Molecular Weight Heparin for VTE Prophylaxis"  
    - "Medication, Administered: Injectable Factor Xa Inhibitor for VTE Prophylaxis"  
    - "Medication, Administered: Warfarin"

Value Set Details

<table>
<thead>
<tr>
<th>Name</th>
<th>Low Dose Unfractionated Heparin for VTE Prophylaxis</th>
</tr>
</thead>
<tbody>
<tr>
<td>OID</td>
<td>2.16.840.1.113762.1.4.1045.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code</th>
<th>Descriptor</th>
<th>Code System</th>
</tr>
</thead>
<tbody>
<tr>
<td>1361568</td>
<td>heparin sodium, porcine 2000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361574</td>
<td>heparin sodium, porcine 20000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361577</td>
<td>heparin sodium, porcine 2500 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361615</td>
<td>heparin sodium, porcine 5000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361853</td>
<td>0.5 ML heparin sodium, porcine 10000 UNT/ML Prefilled Syringe</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1362831</td>
<td>heparin sodium, porcine 10000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1362837</td>
<td>heparin sodium, porcine 12500 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
</tbody>
</table>

https://vsac.nlm.nih.gov/
Analytics

• Clinical data warehouses
  – Distinct from EHR systems

• “ETL” (extract – transform – load) processes
  – Data normalized to “standards”
  – Local data mapped to ontologies
    • Facilitated by ontology integrations systems (e.g., UMLS)

• Analysis leverages hierarchical and other relations
  – Transitive closures

• Facilitates analysis of large-scale data repositories
  – Including distributed repositories across institutions
Healthcare Track: April 7, 2016

- Keith Campbell (US Veterans Health Administration (VHA) Knowledge-Based Systems, USA):
  - SOLOR: Building a coherent foundation for clinical data by leveraging the SNOMED/LOINC Cooperative Agreement, and integrating RxNorm

- Barry Smith (U-Buffalo, NY, National Center for Ontological Research, USA):
  - An Ontology Ecosystem Approach to Electronic Health Record Interoperability

- Amit Sheth (Kno.e.sis, Wright State University, USA):
  - Ontology-enabled Healthcare Applications exploiting Physical-Cyber-Social Big Data

- Leo Obrst (MITRE, USA)
  - Semantic Interoperability in General and in Healthcare: Clinical Care
• Use description logic to integrate SNOMED, LOINC and RxNorm
• SNOMED/LOINC co-operative agreement is already achieving the integration of SNOMED and LOINC
• RxNORM raw data is mapped into OWL EL and then integrated into above.
• SNOMED + LOINC + RxNorm + post-coordination provides comprehensive coverage for typical clinical data representation requirements
• Even a total victory for EPIC (which is becoming the standard EHR in many organizations) will not solve the interoperability problem.

• MU Stage 2 requires use of SNOMED CT for value lists (see Olivier’s talk)
  – Human coding with SNOMED is unreliable and inconsistent
  – e.g. **Solitary leiomyoma (disorder)** has Concept ID: 254769006 and **Leiomyoma (morphologic abnormality)** has Concept ID: 4459800. Are these same or different?

• FHIR (Fast Healthcare Interoperability Resources) is proposed as a solution and still people are trying to figure this out
Personalized medicine needs large cohorts with rich phenotypic data conforming to common standards. How to get there?

THREE CHOICES
1. Everyone uses Epic
2. Government enforces common standards
3. Let’s start again from scratch, using the same approach we should have used from the beginning: rigorous testing-based development of EHRs by leading medical research institutions until we see what technologies will work
Building Ontologies with Basic Formal Ontology

By Robert Arp, Barry Smith and Andrew D. Spear

Overview
In the era of “big data,” science is increasingly information driven, and the potential for computers to store, manage, and integrate massive amounts of data has given rise to such new disciplinary fields as biomedical informatics. Applied ontology offers a strategy for the organization of scientific information in computer-tractable form, drawing on concepts not only from computer and information science but also from linguistics, logic, and philosophy. This book provides an introduction to the field of applied ontology that is of particular relevance to biomedicine, covering theoretical components of ontologies, best practices for ontology design, and examples of biomedical ontologies in use.

After defining an ontology as a representation of the types of entities in a given domain, the book distinguishes between different kinds of ontologies and taxonomies, and shows how applied ontology bears some traditional ideas from metaphysics. It presents the confusions of the Basic Formal Ontology.
Ontology-enabled Healthcare Applications Exploiting Cyber-Physical-Social Data -> Amit Sheth

• Discussed several projects at Kno.e sis (See http://knoesis.org/amit/commercialization/index.php?page=2)

• Ontologies developed at Kno.e.sis
  – Drug Abuse Ontology
  – Depression Insight Ontology
  – Healthcare Ontology
  – Human Performance and Cognition Ontology (HPCO)
  – Ontology for Parasite Lifecycle (BioPortal)
  – Parasite Experiment Ontology (BioPortal)
  – Provenir Ontology (Provenir)
I was sent home with 5 x 2 mg Suboxones. I also got a bunch of phenobarbital (I took all 180 mg and it didn't do shit except make me a walking zombie for 2 days). I waited 24 hours after my last 2 mg dose of Suboxone and tried injecting 4 mg of the bupe. It gave me a bad headache, for hours, and I almost vomited. I could feel the bupe working but overall the experience sucked.

Of course, junkie that I am, I decided to repeat the experiment. Today, after waiting 48 hours after my last bunk 4 mg injection, I injected 2 mg. There wasn't really any rush to speak of, but after 5 minutes I started to feel pretty damn good. So I injected another 1 mg. That was about half an hour ago. I feel great now.
Smarter Data Generated with Ontologies

**Ontology**

**Lexicon**

**Lexico-ontology**

**Rule-based Grammar**

**ENTITIES TRIPLES**

Suboxone, Kratom, Herion, Suboxone-CAUSE-Cephalalgia

disgusted, amazed, irritated more than, a, few of I, me, mine, my Im glad, turn out bad, weird

ointment, tablet, pill, film smoke, inject, snort, sniff Itching, blisters, flushing, shaking hands, difficulty breathing

DOSAGE: <AMT><UNIT> (e.g. 5mg, 2-3 tabs) FREQ: <AMT><FREQ_IND><PERIOD> (e.g. 5 times a week) INTERVAL: <PERIOD_IND><PERIOD> (e.g. several years)
Objective:

– Explain the advantages and disadvantages of applying Semantic Web technology for enhanced visualization and clinical decision support in EHR systems
– Describe the Clinical Care Ontology, developed as part of this effort, and how it builds on existing clinical terminologies and ontologies
– Appraise the Clinical Care Ontology for its usefulness in EHR visualization using nine clinical use cases spanning diabetes, congestive heart failure, angina, and depression in inpatient and outpatient scenarios
– Describe findings to characterize the value of Semantic Web technology for enhanced EHR visualization and clinical decision support in general
Methodology

- Analyze current ontologies and terminologies
- Develop clinical use cases
- Create or adapt a Clinical Care Ontology
- Evaluate Clinical Care Ontology using clinical use cases
- Evaluate applicability of Semantic Web technology in context of clinical use cases

Goal: support patient-centered, care plan-centered visualization and reasoning
Clinical Care Ontology – Care Plan and Goals

CarePlan

Initial Patient State → Plan CarePlan → Schedule CarePlan → Execute CarePlan → Final Patient State

Condition

Assessment

Diagnose

Legend

- **Intervention**
- **State**
- **Encounter**
- **Observation**

Success!
Clinical Care Ontology – Care Plan and Goal

Diagram showing the relationships between Care Plan, Order, Condition, Intervention, Goal, Effect, Value, Quality, and Result. Arrows indicate the directional relationships among these elements, such as intendedToAddress, intendedToAchieve, intendedToEffect, affects, ameliorates, exacerbates, induces, affects, indicatedBy, indicatesDegreeOf, reflectsRiskFor, hasValue, hasQuality, evidencedBy, hasTrend, hasStartingQuality, hasCurrentQuality, hasTargetQuality.