

Ontology as a Service (OAAS): Rich Semantics for Data Analytics

Jane Greenberg, Yuan An, Tony Hu, and Yue Zhang, College of Computing and Informatics, Drexel University

NEED & INDUSTRIAL RELEVANCE

PROBLEM

- Financial, human, intellectual capital spent creating ontologies
- Access and use barriers exist; loss of ROI
- Ontologies remain hidden, 'dark assets'
- Moreover, limited use of vetted national ontologies registered @ NCBO (National Center for Biological Ontologies), and other places

SEMANTIC ONTOLOGIES ARE NEEDED to:

- Describe industry data assets
- Provide meaning, facilitate IR
- Support semantic-driven data analytics
- Enable transformative computational approaches, leverage semantics to attain new knowledge

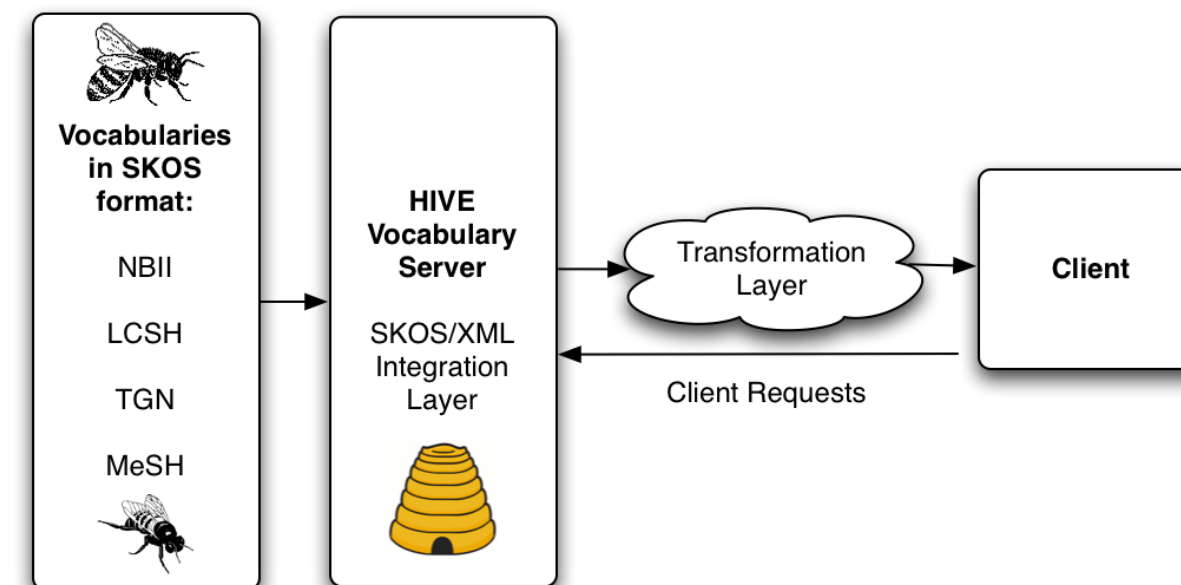
APPROACH (RESEARCH METHODS)

- Ontology registration, 3-step
 - Leverage linked-data feeds
 - Integrate the PROV-O life-cycle model
 - Add semantic similarity measures (SSM) (Lexic.Similarity (VS) + SyntacticSim. (SS))
- Service – OAAS development, extend the HIVE (Helping Interdisciplinary Vocabulary Engineering) infrastructure.
 - [HIVE-4-HEALTH](#)
 - [HIVE-4-BIOSCIENCE](#)
 - [HIVE-4-_____](#)
- Provide for plug-in algorithms, e.g. RAKE, Maui++
- Evaluate the OAAS performance and use



PROJECT GOALS

- Develop and advance an adaptable platform supporting rich semantic data analytics
- Improve ontology discovery and maintenance
- Enable plug-ins for selecting different algorithms, including machine-learning capacities
- Leverage semantic metadata + ontologies for daily operations and new insights

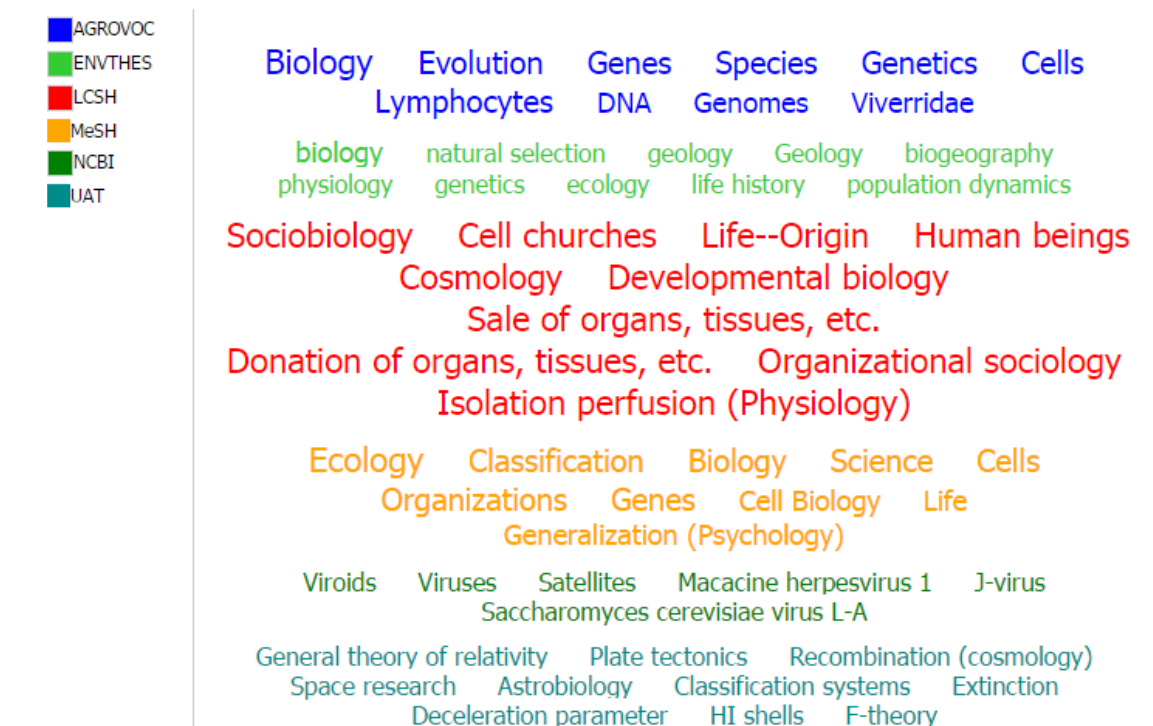


DELIVERABLES/OUTCOMES

- An innovative platform and adaptable ontology service
- Code supporting ontology registration, maintenance, life-cycle management, and automatic application
- Tested, confirmed use of similarity measures to monitor ontology currency
- Plug-in option enabling use of different indexing algorithms
- Framework to leverage broad spectrum of existing ontologies
- CVDI Ontology Roundtable

OBJECTIVES

- Advance ontology registration; support interoperability, cross-domain discover, and reuse of intellectual semantic assets
- Reduce semantic vocabulary silos
- Improve discovery and application of ontologies across the organization
- Increase organizational intelligence, leverage data and semantic (intellectual) assets



IMPACT

- An approach + platform for optimal application of rich semantic networks within/across industry
- Platform and service to leverage, reuse, and benefit from industry specific semantics assets + semantic resources dev. by government (NCBO) and others
- Enable intelligent, robust, transformative, semantically-driven data analytics, due to rich semantics and interoperability

