Extracting Similar Flavors in the Spaghetti of Biomedical Ontologies

Maulik R. Kamdar, Tania Tudorache and Mark A. Musen

Stanford University

STANFORD SCHOOL OF MEDICINE
Stanford University Medical Center

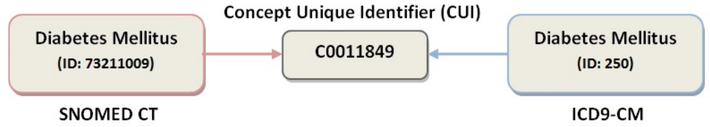
Motivation

Ontology Uses: Knowledge management, semantic search, data annotation, data integration, exchange, decision support and reasoning

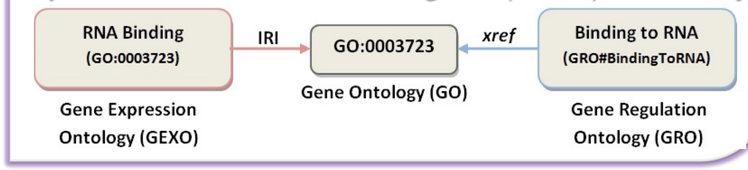
Advantages of Reuse

- Reduce engineering costs
- Avoid building existing ontology structures
- Enable semantic interoperability,
- Support use of terminologies in EHRs
- Query multiple, heterogeneous knowledge sources

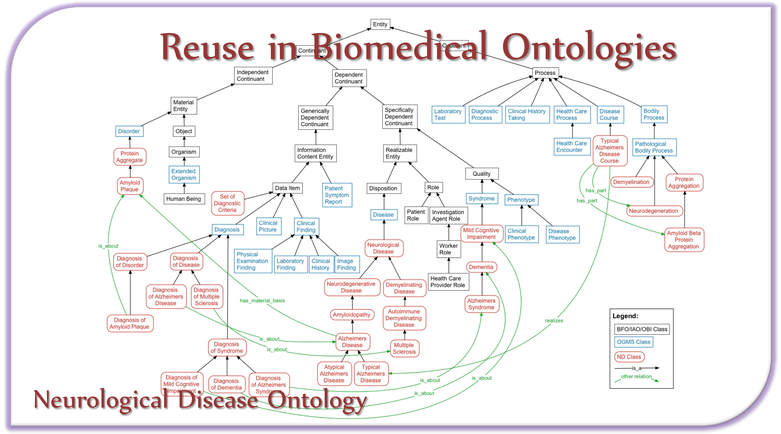
Unified Medical Language System (UMLS)



Open Biomedical Ontologies (OBO) Foundry

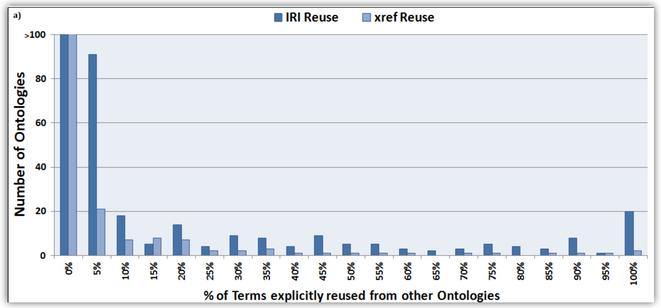


Reuse in Biomedical Ontologies

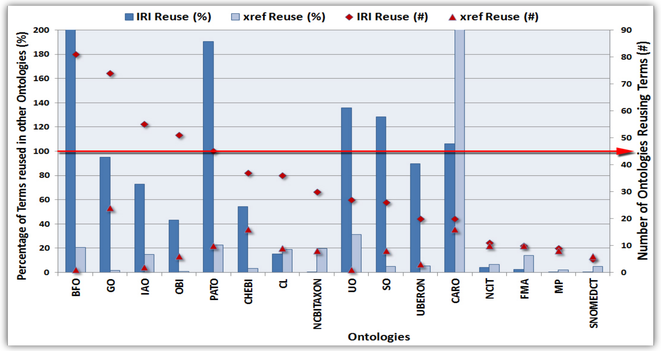


Results: Term Reuse

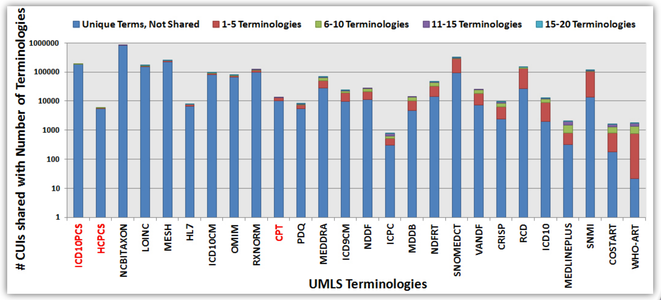
Total Term Reuse using all 3 constructs: 8.6%



Most ontologies reuse <5% Terms

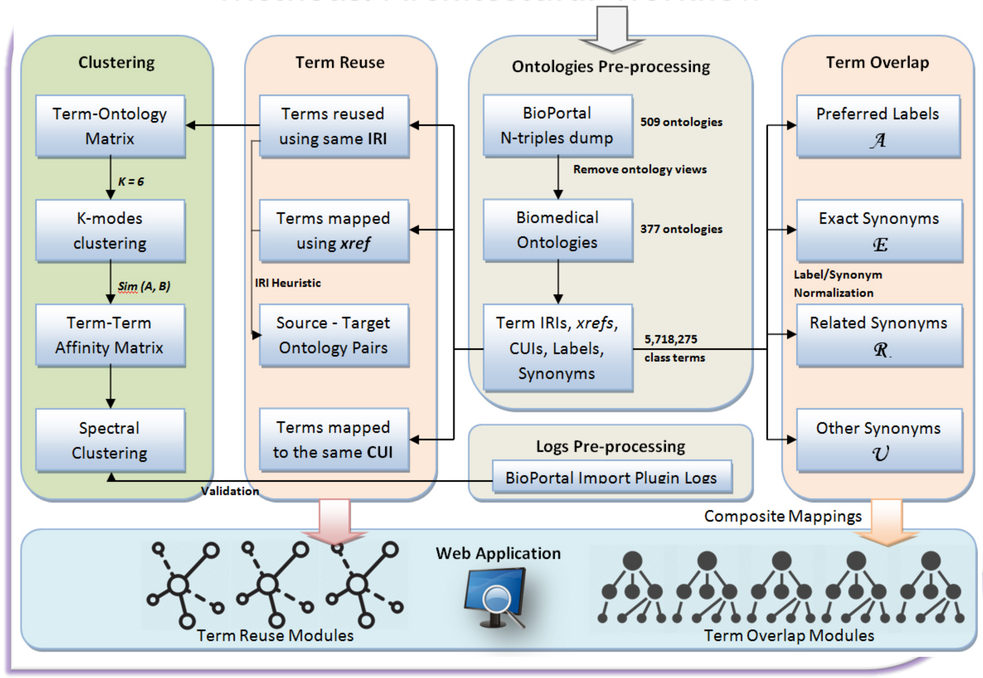


>100% Terms are reused from some ontologies

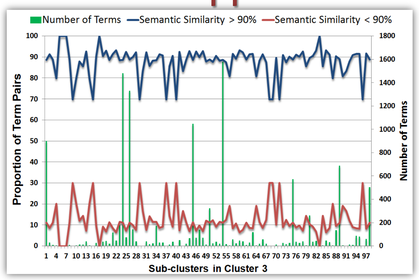


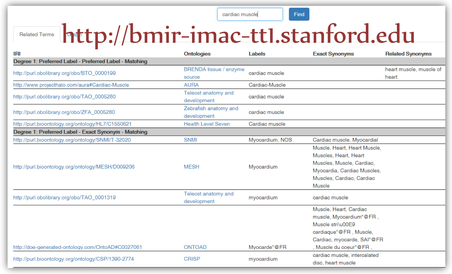
Most popular terminologies are composed of unique terms
Procedural terminologies do not share terms with each other

Methods: Architectural Workflow



Application: Term Recommender

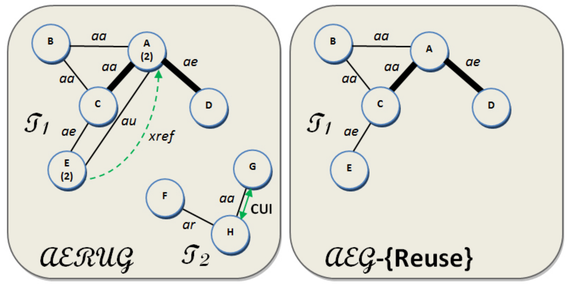


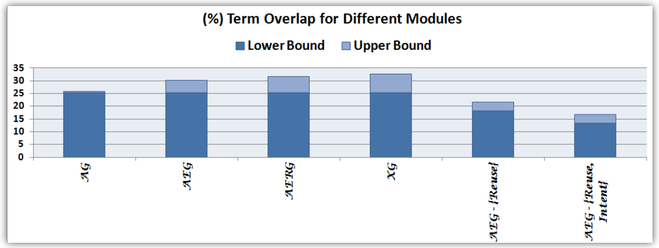


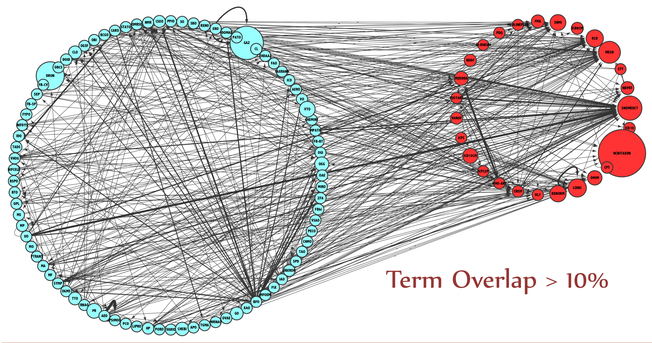
Acknowledgements: Musen Lab @ Stanford, US NIH Grants GM086587, GM103316

Contact: maulikrk@stanford.edu

Results: Term Overlap







Term Overlap > 10%

Intent for Reuse:

- **Different Versions:** NCIT:Cerebral_Vein/NCIT:C53037
- **Different Notations:** OBO:FMA_31396, OBO:owlapi/fma#FMA_31396, OBO:owl/FMA#FMA_31396, OBO:fma#Cartilage_of_inferior_surface_of_posterolateral_part
- **Different Namespaces:** http://ihtsdo.org/snomedct/, http://purl.bioontology.org/ontology/SNOMEDCT