InPhO: A System for Collaboratively Populating and Extending a Dynamic Ontology[*]

Mathias Niepert, Cameron Buckner, Jaimie Murdock and Colin Allen
Indiana University
Bloomington, Indiana, USA
{mniepert, cbuckner, jammurdo, colallen}@indiana.edu

ABSTRACT

InPhO is a system that combines statistical text processing, information extraction, human expert feedback, and logic programming to populate and extend a dynamic ontology for the field of philosophy. Integrated in the editorial workflow of the Stanford Encyclopedia of Philosophy (SEP), it will provide important metadata features such as automated generation of cross-references, semantic search, and ontology driven conceptual navigation.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; H.3.7 [**Information Storage and Retrieval**]: Digital Libraries—*Systems issues*

General Terms

Algorithms, Human Factors

Keywords

digital humanities, information extraction, formal ontology, dynamic ontology, answer set programming

1. INPHO: THE INDIANA PHILOSOPHY ONTOLOGY PROJECT

The scale, complexity, and dynamic nature of many digital reference works calls for sophisticated information management tools to capture and harness their metadata. Structured representations of their subject domains are required for scholarly functions such as semantic search, automated generation of cross-references and tables of contents, and ontology-driven conceptual navigation. For scholarly projects with limited resources, human-intensive methods for coding, populating, and maintaining ontologies are not viable. "Social tagging" approaches leading to so-called "folksonomies" will be useful for some purposes, but may not meet standards of academic review.

We present InPhO[1], a system that makes use of a dynamic reference work's most valuable informational resource: the domain experts that serve as their editors and contributors. Through statistical text processing and information extraction, taxonomic and non-taxonomic information is retrieved from both the documents of the SEP as well as external sources such as Wikipedia and academic genealogy datasets. Human expert feedback is used to evaluate the automated methods' recommendations without presuming knowledge of ontology design or placing undue demands on the contributors' time. The feedback from the authors provides overlapping pieces of expert knowledge which are ultimately put together by a nonmonotonic answer set program we have developed for this task [2].

A detailed description of the methodology can be found in [1] and [2]. Web interfaces are provided to the authors of the SEP to evaluate the suggestions generated by the automatic methods. For example, users can score the relatedness of two ideas in philosophy or the influence one philosopher has on another on a five point scale. Each of these evaluations is stored and a logic program is run to construct a global populated ontology from the local feedback facts. Figure 1 depicts a diagram of the system. The populated taxonomy can be browsed online[2].
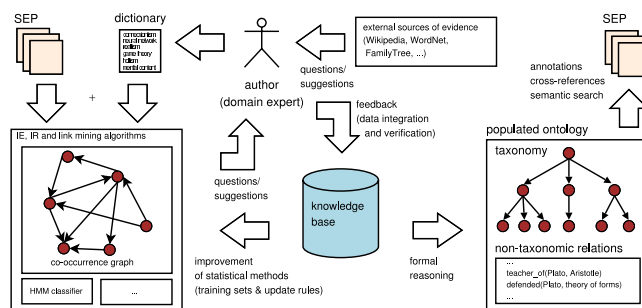

Figure 1: The metacontent generating engine.

2. REFERENCES

[1] M. Niepert, C. Buckner, and C. Allen. A dynamic ontology for a dynamic reference work. In *Proceedings of 7th JCDL*, pages 288–297. ACM Press, 2007.

[2] M. Niepert, C. Buckner, and C. Allen. Answer set programming on expert feedback to populate and extend dynamic ontologies. In *Proceedings of 21st FLAIRS*. AAAI Press, 2008.

[1]http://inpho.cogs.indiana.edu

[2]http://inpho.cogs.indiana.edu/taxonomy/