

Tutorial: Advances in Biomedical Ontologies

Barry Smith^{a,1} and William Hogan^b
^a*University at Buffalo, Buffalo, NY, USA*
^b*University of Florida, Gainesville, FL, USA*

Abstract. This tutorial will survey recent advances in biomedical ontologies. It will describe progress that has been made and outline emerging areas of ontology development. It will also indicate how the results of this work are being applied in advancing (1) interoperability of data and information systems in support of clinical research, and (2) new kinds of ontology-based experimentation in areas such as patient welfare and healthcare organizational structure.

Keywords. ontology, clinical and translational research

1. Topic and format

Proposed is a 3-hour tutorial for scheduling on the afternoon of April 23. As an alternative, the tutorial can be modified to create a 90 minute version for presentation during the main meeting.

The format in each case will be presentations and demonstrations by Drs Smith and Hogan describing ontology goals, content, and applications, together with extended question/answer and discussion periods.

2. Audience Level

The tutorial is intended for participants in the MIE event at all levels, including students. It will not presuppose any technical knowledge of ontologies.

3. Overview description of the contents

3.1. *Recent advances in biomedical ontologies (60 minutes)*

We will begin by describing what ontologies are and how they are used in biomedical informatics. This will be followed by a brief description of the Open Biomedical Ontologies (OBO) Foundry, and of recent additions to the Foundry, including the Basic

¹ Corresponding Author: Barry Smith, National Center for Ontological Research, 104 Park Hall, University at Buffalo, Buffalo, NY 14260, phsmith@buffalo.edu, <http://ontology.buffalo.edu/smith>

Formal Ontology (BFO), Ontology for Biomedical Investigations (OBI), Disease Ontology (DO), Cell Ontology (CL), and Protein Ontology (PRO).

We will describe how PRO is being re-used to represent (1) immune cells in CL, (2) drug ingredients and targets in the Drug Ontology (DrOn), and (3) protein targets of antibodies in the new ImmPort Antibody Ontology. We will then show the benefits for data interoperability that are yielded when ontologies are re-used in this way.

We will describe how DO is being reengineered to conform to the Ontology for General Medical Science (OGMS) and explore implications of this work for our understanding of the ontology underlying SNOMED.

3.2. Using ontologies to advance data exchange (60 minutes)

We will describe recent work in applying ontologies to the problem of exchanging and querying clinical data for research. Secondary use of EHR data in research typically involves transforming such data into what are called common relational data models (CDMs). The problem is that CDMs are hard to use properly without significant discussion and interaction with informatics personnel who are experts in the CDMs. This problem becomes even more intractable where multiple CDMs are used simultaneously, for example when data from multiple different hospitals needs to be aggregated for consortium clinical studies.

We will describe an alternative approach to reusing EHR data in clinical studies by transforming EHR data to an ontology-based format. We will then outline how this approach overcomes some of the problems with existing CDMs, for example (1) by allowing researchers with no specialist knowledge of CDMs to more easily write and to fine-tune their queries, and (2) by allowing more rapid accommodation of the data management environment to advances in clinical and translational knowledge. We hypothesize that these features can lead to a higher percentage of research questions being successfully addressed by the available data, and we will discuss our initial tests of this hypothesis, describing examples of the ontological resources employed, including the Ontology of Medically Relevant Social Entities (OMRSE), DO, DrOn, and OBI.

3.3. New avenues for biomedical ontology research (60 minutes)

Ontology of Organizations: We will describe recent work in applying biomedical ontologies to the description of organizational structures. For example, OMRSE is being used to represent various kinds of healthcare organizations and the facilities they operate to capture data about healthcare encounters. We will show how OMRSE is being used in tandem with the Document Acts (d-acts) ontology and the Ontology of Organizational Structures of Trauma Systems and Trauma Centers (OOSTT) to organize data about trauma systems in a major study that is designed to identify those features of trauma systems which are associated with improved patient outcomes.

Ontology of Capabilities: We will describe current work to develop an Ontology of Capabilities, relating both to the abilities (and disabilities) of human beings, to the skills, expertise and training of human beings, and also to the capabilities of organizations (such as trauma centers) and of associated systems, teams and equipment. We will show how a capabilities approach can be used as the basis for a new understanding of medical

professionalism and of patient welfare, and also how it can be used in the formalization of the ontology underlying the International Classification of Functioning, Disabilities and Health.

4. Instructor biographies

Barry Smith is a prominent contributor to both theoretical and applied research in biomedical ontology and he has organized some 50 tutorials and 80 workshops and conferences on ontology-related themes. The Basic Formal Ontology (BFO), which he founded in 2002, is now used as upper-level architecture by more than 300 groups of ontology researchers throughout the world. He plays a key role in the Protein Ontology, Infectious Disease Ontology, Ontology for General Medical Science (OGMS), and Environment Ontology (EnvO) projects, and also in the Open Biomedical Ontologies (OBO) Foundry initiative. His research has been funded by the National Institutes of Health, the US, Swiss and Austrian National Science Foundations, the US Department of Defense, the Volkswagen Foundation, and the European Union.

William Hogan is a prominent biomedical informatics leader who has longstanding research and applied expertise in EHRs, biomedical ontologies, and especially the intersection of these two domains. He has organized nearly 20 workshops and tutorials on ontology-related subjects. He founded and directs the ongoing development and maintenance of the Ontology of Medically Relevant Social Entities (OMRSE), the Drug Ontology (DrOn), and the Apollo-SV ontology for infectious disease epidemiology. He is a contributor to OGMS, the Information Artifact Ontology, and the Ontology of Biomedical Investigations and has also participated in BFO 2.0 development. He is the Director of Informatics for OneFlorida, a statewide data research network in Florida that houses EHR and administrative claims data on 15 million Floridians. His work involves application of his ontology expertise to improve OneFlorida data management and usage.

5. Expected learning outcomes

All participants will gain an understanding of basic ontology principles and knowledge of a series of important developments in current biomedical ontology research and applications. They will acquire an awareness of the structure and interrelations of some of the most important ontologies being used today in clinical and translational science and an appreciation of some of the benefits of ontology-based data management as compared to more traditional approaches.

Background reading

- [1] R. Arp, B. Smith, A. Spear, *Building Ontologies with Basic Formal Ontology*, MIT Press, Cambridge, MA, 2015.