OBO Tutorial

James A. Overton   Christopher John Mungall

Abstract

The Open Biological and Biomedical Ontologies (OBO) provide a suite of interoperable ontologies for biology and medicine. Ontology users and developers benefit from close collaboration and shared best practises, and they benefit most when they are familiar with a range of shared tools and techniques. In this interactive tutorial we present a range of real-world examples focused on OBO Foundry and Library ontologies, covering the use and reuse of ontologies, processing data with ontologies, and contributing to ontology development.

Description

The Open Biological and Biomedical Ontologies (OBO) provide a suite of interoperable ontologies for biology and medicine. We propose a half-day OBO Tutorial at ICBO 2017, in which we will present a range of real-world examples focused on OBO projects, covering the use and reuse of ontologies, processing data with ontologies, and contributing to ontology development.

Every ICBO conference welcomes new members into the community of ontology users and developers. This interactive tutorial contributes to the field of healthcare and life science ontologies by introducing these new members to the fundamental theory and practise of ontology use and development, with a special focus on data sharing.

The tutorial content will be a thoroughly revised version of the material presented at previous ICBO conferences and freely available online at https://github.com/jamesaoverton/obo-tutorial. We will update the content with the latest improvements to our tools and best practises. By providing the material online, we help ICBO participants to prepare in advance, to refresh their memories later, and to continue learning about topics not covered in the half-day at ICBO 2017.

1. Names and Statements

The tutorial will begin with the fundamentals of Linked Data. Using a spreadsheet of ambiguous data as a concrete example, we explain the fundamental importance of unambiguous names for data sharing. Internationalized Resource Identifiers (IRIs) make excellent machine-readable globally-unique names. By combining IRIs and literal strings into subject-predicate-object triples, we can make unambiguous statements. In this way we introduce the basics of the Resource Description Framework (RDF) and its usefulness for data sharing.

2. Open Biological and Biomedical Ontologies

Ontologies address the problem of shared names for classes of things and the relations between them. We introduce the OBO library of ontologies, covering a wide range of biology and biomedicine, and the OBO principles that support interoperability between ontologies.

3. Using and Reusing Ontologies

Next we dive into some of the practicalities of working with ontologies. Our goal in this section is to build a small application ontology using terms from reference ontologies, to support the data in the example spreadsheet. The application ontology is based on the ontology-starter-kit. First, we discuss tools for finding the right reference ontology and the right terms within it. Second, we show how to get a local copy of that ontology and open it in an ontology editor. Third, we extract selected terms from reference ontologies into the application ontology.

4. Processing Data with Ontologies

Returning to our example spreadsheet, and using our new application ontology, we show how to work with instance data. This includes converting between tables and triples, querying instance data using SPARQL, using SPARQL to insert triples, and using reasoners to check data and add inferences.

5. Contributing to Ontology Development

The final part of the tutorial will show participants how they can submit term requests and bug reports to OBO projects, encouraging them to share their expertise by helping to develop open source community resources.

Requirements

This will be a hands-on tutorial and participants will be encouraged to bring their laptops, follow along, and try things out. During the tutorial an Internet connection will be helpful, but not required if the following are installed in advance:

1. the tutorial content, available on GitHub
2. Protégé for viewing OWL files
3. ROBOT for running examples
4. optionally the ontology-starter-kit

Items 1-3 are cross-platform (using Java), while item 4 requires Unix (Linux, macOS).

If this tutorial proposal is accepted for IBCO 2017, we would appreciate the organizers' help in contacting the participants in advance, so that we can provide the link to the online version of the tutorial and instructions for downloading required software. A room with tables and outlets for laptops would be preferable for the tutorial session.

We will also require a data projector that the organizers can use with their laptops.

Organizers

- James A. Overton james@overton.ca, Knocean Inc.
- Christopher John Mungall cjmungall@lbl.gov, Lawrence Berkeley National Laboratory

Dr. Overton organized OBO Tutorials at ICBO 2016, 2015, and 2014, and Dr. Mungall has organized multiple tutorials and workshops at previous ICBO conferences. Both are members of the OBO Foundry Technical Working Group, and have extensive experience teaching ontology use and development.

Discussion

We believe that introductory tutorials such as this one attract more people to ICBO, and are especially helpful for attendees from the local university and community. OBO Tutorials have been part of ICBO for many years, and our sessions have been consistently well-attended. Tutorials in 2014, 2015, and 2016 each filled the available space, with approximately 50, 30, and 30 participants respectively.

Large attendance makes it a challenge to provide a hands-on, interactive experience, and previous tutorials have not been as hands-on as we had planned. If this proposal is accepted for ICBO 2017, we plan to address that challenge in several ways:

1. contact participants well in advance with instructions on how they can prepare
2. recruit colleagues at ICBO 2017 to help with technical support and individual questions during the tutorial
3. update the content with greater focus on examples and exercises
4. tailor the content covered to the timeslot provided

We will advertise the tutorial on the relevant OBO and ontology mailing lists.