

OBO Academy: Training materials for bio-ontologists

Nicole Vasilevsky¹, James Overton², Rebecca Jackson³, Sabrina Toro¹, Shawn Tan⁴, Bradley Varner⁴, David Osumi-Sutherland⁴, Nicolas Matentzoglou⁵

¹University of Colorado, Anschutz Medical Campus, Aurora, CO, USA

²Knocean, Inc, Toronto, Canada

³Bend Informatics LLC, Bend, Oregon, USA

⁴European Bioinformatics Institute, Hinxton, UK

⁵Semanticly, Athens, Greece

ABSTRACT

Motivation: Biomedical ontologies are widely available, with hundreds of ontologies under development, however, there is a lack of formal training on methods for ontology development, including best practices for how to create and edit ontologies, and the application of new tools and workflows. This presents a challenge for new and current ontologists to find and access training materials, and learn the methodology or hone existing skills. The OBO Academy provides open, online, self-paced training materials that aim to provide ongoing training for the ontology community on best practices in ontology development. The training materials cover a range of topics from basics like getting started in contributing to ontologies and editing in Protege, to more advanced materials that cover technical workflows such as using the Ontology Development Kit and ROBOT templates. The initial offering of materials is available online, and is under continuous development, and community feedback and contributions are welcomed (<https://github.com/OBOAcademy/oobook>).

1 INTRODUCTION

Hundreds of biomedical ontologies exist and are currently under development, however there is a distinct lack of training material and educational resources on how to create, develop, edit and maintain biomedical ontologies using current tools and best practices. Many ontology developers and curators are trained ad hoc and have not undergone any formal education on ontology development. The Open Biological and Biomedical (OBO) Foundry (<https://obofoundry.org/>) is a community of ontology developers that are committed to building open ontologies under a shared set of guiding principles and best practices. The OBO Foundry provides resources on ontology development (<https://obofoundry.org/resources>), offers in-person tutorials at the International Conference on Biomedical Ontologies (ICBO) and ad hoc training sessions at webinars and conferences, but, to date, there are no formal training programs. This presents a challenge for current ontologists and new ontologists entering the field to formally learn new skills and workflows. The OBO Academy is an effort to develop open, freely available training materials, provide formal documentation on ontology workflows and provide ongoing training for the OBO community. Below we describe our approach to developing these training materials and hosting online sessions for team members in the Monarch Initiative with the goal of training future and current bio-ontologists in current methods and best practices.

2 APPROACH

The offerings from the OBO Academy include freely available, open, online training materials (available here: <https://oboacademy.github.io/oobook/>). In addition, we have been holding biweekly training sessions with our instructor team for our team members in the Monarch Initiative. The online materials are intended to be self-paced learning materials and are organized into:

- **Tutorials:** learning-oriented material that contains exercises to help learners achieve basic competence in a specific area.
- **How-to guides:** task-oriented guides that function as directions for specific tasks.
- **Reference guides:** information-oriented documentation that describes a single topic in a succinct, technical and orderly way.
- **Explanations:** understanding-oriented documentation that clarifies, deepens and broadens the reader's understanding of a subject.

These are based on the Diátaxis framework (<https://diataxis.fr/>). A list of the currently available tutorials/lessons are outlined in **Table 1**. Additional tutorials and lessons are under development and we welcome input from the community to iterate on existing material or share new material. As part of the online materials, we are developing *pathways*, which are curated training materials that provide a linear path for learners working in a specific role. Two pathways will be available, one for Ontology Curators and one for Ontology Engineers/Developers. These will be accessible online (<https://oboacademy.github.io/oobook/pathways/pathways/>) and are intended to provide self-paced learning materials for a learner working in one of these two areas. We will offer

two sub-pathways for Ontology Curators: Gene Ontology (GO)-style curation, which is primarily based on editing in Protégé ontology editing software, and Ontology of Biomedical Investigations (OBI)-style, which is based on template-based ontology contributions.

Table 1: Self-paced tutorials available from the OBO Academy.

These lessons are initial offerings and intended to be continuously updated and community feedback is welcome.

Contributing to OBO ontologies
Ontology Pipelines with ROBOT
Analysing Linked Data
Automating Ontology Workflows
Ontology Design
Templates for OBO ontologies
Modelling with Object Properties
Ontology Term Use
Disease and Phenotype Ontologies

In addition to the open, online materials, we are currently offering biweekly, instructor-guided tutorials via zoom to members of the Monarch Initiative. Our initial online sessions followed the flipped classroom concept: organisers provided the participants with materials in advance from our repertoire of tutorials, and they were expected to work through the materials on their own in between sessions. The estimated preparation time for each lesson was about 3 - 7 hours. During the biweekly meetings, the instructors were available to answer questions and provide additional demonstrations where needed or to go into more depth for topics of interest. While the flipped classroom approach had its benefits, the students reported that they did not have sufficient time to adequately prepare and review all of the materials between sessions, so we shifted our approach to offer a combination of flipped classroom and a traditional learning environment, with a lecture or demonstration, followed by questions and discussion. Since the start of the new year, we've taken this approach and we post informal recordings of our sessions online (<https://oboacademy.github.io/oobook/courses/monarch-obo-training/>).

The materials are iteratively being developed, with new content being added upon request for the Monarch tutorials. We have a GitHub discussion board (<https://github.com/OBOAcademy/oobook/discussions>) used to drive the schedule of upcoming sessions where participants can suggest, request, and vote on tutorial topics.

The OBO Academy is intended to be community driven, where any of our community members can provide content for the online materials, and volunteer to teach a session for the Monarch tutorial. The materials are continuously updated and feedback and edits via GitHub pull requests are welcome from the community (<https://github.com/OBOAcademy/oobook>). A guide on how to edit the training materials is available on our site under OBOOK Editors and here: <https://oboacademy.github.io/oobook/contributing/>.

3 CONCLUSION

This resource is in the early stages of development, and the initial feedback from the online sessions and regarding the training materials has been very positive and well received. Our aim is to create a centralized resource with training materials from various sources, and iteratively improve the content.

For more information or to get involved, please create an issue on our GitHub tracker (<https://github.com/OBOAcademy/oobook/issues>) or view our contributing webpage (<https://oboacademy.github.io/oobook/contributing/>).

ACKNOWLEDGEMENTS

The Critical Path Institute commissioned the first version of this OBO course, contributing funding for the preparation and delivery of the materials and providing valuable feedback about the course contents and data for the practical exercises. This work is currently funded by the Phenomics First grant NIH-NHGRI # 1 RM1 HG010860-01.