

# WebProtégé: a collaborative Web-based platform for editing biomedical ontologies

Matthew Horridge\*, Tania Tudorache, Csongor Nuytas, Jennifer Vendetti, Natalya F. Noy and Mark A. Musen

Stanford Center for Biomedical Informatics Research, Stanford University, 1265 Welch Road, Stanford, CA 94305, USA

Associate Editor: John Hancock

## ABSTRACT

**Summary:** WebProtégé is an open-source Web application for editing OWL 2 ontologies. It contains several features to aid collaboration, including support for the discussion of issues, change notification and revision-based change tracking. WebProtégé also features a simple user interface, which is geared towards editing the kinds of class descriptions and annotations that are prevalent throughout biomedical ontologies. Moreover, it is possible to configure the user interface using views that are optimized for editing Open Biomedical Ontology (OBO) class descriptions and metadata. Some of these views are shown in the Supplementary Material and can be seen in WebProtégé itself by configuring the project as an OBO project.

**Availability and implementation:** WebProtégé is freely available for use on the Web at <http://webprotege.stanford.edu>. It is implemented in Java and JavaScript using the OWL API and the Google Web Toolkit. All major browsers are supported. For users who do not wish to host their ontologies on the Stanford servers, WebProtégé is available as a Web app that can be run locally using a Servlet container such as Tomcat. Binaries, source code and documentation are available under an open-source license at <http://protegewiki.stanford.edu/wiki/WebProtege>.

**Contact:** [matthew.horridge@stanford.edu](mailto:matthew.horridge@stanford.edu)

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

Received on November 22, 2013; revised on April 7, 2014; accepted on April 21, 2014

## 1 INTRODUCTION

Ontologies are important artefacts in bioinformatics and life sciences. Many biomedical ontologies, such as the ones made available in the BioPortal (Noy *et al.*, 2009), the Open Biomedical Ontology (OBO) Library (Smith *et al.*, 2007) and the OntoBee platform (Xiang *et al.*, 2011) are published and exchanged as OWL 2 ontology documents. Furthermore, biomedical ontologies are typically developed collaboratively by groups of experts who may receive incremental input from the end-user community. However, although there have been some domain-specific wiki-based tools for the collaborative editing of biomedical knowledge (Larson and Martone, 2013), Web-based tool support for community-based OWL ontology development is severely lacking. In this article, we present WebProtégé—a Web-based

tool for the collaborative development of OWL 2 ontologies. Although the latest version of this tool supports the complete OWL 2 language (Motik *et al.*, 2009), the default configuration of its customizable user interface (UI) is primarily geared towards lightweight ontology editing [it is possible to configure the user interface with OWL 2-specific views. Please see the WebProtégé user guide for more details (<http://tinyurl.com/webprotege-editing-owl>)]. In particular, WebProtégé provides out-of-the-box views that support a fragment of the language that is useful for capturing biomedical knowledge and is traditionally catered for by tools such as OBOEdit (Day-Richter *et al.*, 2007). Moreover, WebProtégé supplies user interface views that are optimized for editing the rich OBO-term metadata and definitions that are commonly associated with biomedical ontologies. We believe that these OBO-tailored UI views along with collaborative development functionality make WebProtégé an attractive tool for the community-based development of biomedical ontologies.

## 2 COLLABORATION FEATURES

WebProtégé is centred around the notion of a project—a collection of OWL ontologies, custom user interface configuration settings and collaboration settings. Projects can either be created from scratch or they can be initialized with existing ontologies. Any OWL-compatible ontology, including OBO format (Mungall *et al.*, 2011) ontologies, may be uploaded and imported. A key aspect of WebProtégé is that it supports multi-user real-time collaborative ontology development. Once a user has created a project, they can share it with their collaborators, adding them as viewers, commenters or editors. WebProtégé provides the following features to aid collaboration:

- **Threaded Discussions**—Threaded discussions allow issues to be tracked and associated with terms and change sets. Issues can be marked as being ‘open’ and can then be ‘resolved’ when they have been addressed.
- **Watches**—Users can keep an eye on changes to the ontology, changes that modify a description of a particular term or changes that modify hierarchy branches and descriptions of terms in branches. Notifications of changes are sent to users via email with direct links to the terms that were changed.
- **Change Tracking and Versioning**—WebProtégé is built on top of the OWL API, which means that it can take

\*To whom correspondence should be addressed.

advantage of a precisely defined notion of ontology changes and ontology change lists. Change lists, which correspond to operations in the user interface, are stored along with various metadata about which user enacted the changes, the timestamp of the changes and the fine-grained details of the changes. All project changes are tracked, and WebProtégé allows users to download any previous version or revision of an ontology, making it possible to compare different versions using tools such as Bubastis (Malone *et al.*, 2010) or Ecco (Gonçalves *et al.*, 2012).

### 3 AN OBO FRIENDLY UI

The WebProtégé user interface is customizable, and users can arrange individual views (known as ‘portlets’) to suit their needs. We provide several out-of-the-box views that are specifically designed for editing biomedical, OBO-Style, ontologies. Heavily inspired by OBOEdit widgets, many of these views shield the user from the verbose and intricate OWL representation of OBO format constructs. Some of these views are described below.

- **1 TermId View:** presents a user-friendly OBO Id and Name for the selected term and hides the under-the-hood representation of term Ids (IRIs).
- **2 Term Definition View:** obviates the need for the user to know what the correct ‘definition’ OWL annotation property (IAO\_0000115) is, and manages the database cross-reference annotations on the OWL annotation assertion that specifies the term definition.
- **3 Synonyms View:** manages the specification of synonyms, the scope of these synonyms and database cross references for them.
- **4 Relationships View:** simple OBO relationships correspond to OWL ObjectSomeValuesFrom restrictions that are superclass expressions of the subject term. This view manages the SubClassOf relationship and ObjectSomeValuesFrom restrictions under the hood, meaning that users do not need to know about these OWL constructs. Furthermore, auto-completion (9) assists the user in typing in correct property names and term names.
- **5 Cross-Products View:** similar to the Relationships View, this view manages EquivalentClasses axioms and the bundling of component relationships, represented by ObjectSomeValuesFrom restrictions, into intersection classes.

### 4 IMPLEMENTATION

WebProtégé is based on the OWL API (Horridge and Bechhofer, 2011). This API is widely used by people that need programmatic access to OWL ontologies, and it is the basis of other widely used

tools such as Protégé (Horridge *et al.*, 2006), services such as the module and diff analysis tools (Gonçalves *et al.*, 2012; Malone *et al.*, 2010) and OBO workflow tools such as the OBO Ontology Release Tool. For loading and processing OBO ontologies, we use the OBO format tools, which provide a ‘reference’ OBO parser among other things. The WebProtégé user interface and client/server communication is implemented using the Google Web Toolkit (<https://developers.google.com/web-toolkit/>) and JavaScript.

### 5 FUTURE DEVELOPMENT

WebProtégé is currently under heavy development. Some of the features that are relevant for the biomedical ontology community that we plan to add are as follows: support for exporting and publishing projects as OBO ontologies (via integrations with the OBO-Ontology Release Tool); tighter integration with third-party issue trackers that are typically used by biomedical ontology projects (such as GitHub or GitHub Issues); online reasoning and query support (along the lines of the functionality provided by Protégé 4); mixed subclass and non-subclass (e.g. partOf) hierarchy views; better support for visualization and support for subset views (slims).

*Funding:* The National Institute of General Medical Science (GM103316) of the US National Institutes of Health.

*Conflict of Interest:* none declared.

### REFERENCES

- Day-Richter, J. *et al.* (2007) OBO-Edit—an ontology editor for biologists. *Bioinformatics*, **23**, 2198–2200.
- Gonçalves, R.S. *et al.* (2012) Ecco: a hybrid diff tool for OWL 2 ontologies. In: Klinov, P. and Horridge, M. (eds) *Proceedings of OWL: Experiences and Directions Workshop 2012*. Vol. 849, CEUR Workshop Proceedings. <http://ceur-ws.org>.
- Horridge, M. *et al.* (2006) Supporting early adoption of OWL 1.1 with Protégé-OWL and FaCT++. In: Cuenca Grau, B. *et al.* (ed.) *OWL: Experiences and Directions (OWLED)*. Vol. 216, CEUR Workshop Proceedings.
- Horridge, M. and Bechhofer, S. (2011) The OWL API: a Java API for OWL ontologies. *Semant. Web*, **2**, 11–21.
- Larson, S.D. and Martone, M.E. (2013) NeuroLex.org: an online framework for neuroscience knowledge. *Front. Neuroinform.*, **7**, 18.
- Malone, J. *et al.* (2010) Modelling sample variables with and experimental factor ontology. *Bioinformatics*, **26**, 1112–1118.
- Motik, B. *et al.* (2009) *OWL 2 Web Ontology Language Structural Specification and Functional Style Syntax*. W3C Recommendation, W3C—World Wide Web Consortium. <http://www.w3.org/TR/owl2-syntax/> (May 2014, date last accessed).
- Mungall, C.J. *et al.* (2011) Obo flat file format 1.4 syntax and semantics [draft]. *Technical report*, Lawrence Berkeley National Laboratory.
- Noy, N.F. *et al.* (2009) BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.*, **37**, W170–W173.
- Smith, B. *et al.* (2007) The OBO Foundry: coordinated evolution of ontology to support biomedical data integration. *Nat. Biotechnol.*, **25**, 1251–1255.
- Xiang, Z. *et al.* (2011) Ontobee: a linked data server and browser for ontology terms. In: Smith, B. *et al.* (ed.) *ICBO International Conference on Biomedical Ontology*. Buffalo, New York, USA. [http://icbo.buffalo.edu/ICBO-2011\\_Proceedings.pdf](http://icbo.buffalo.edu/ICBO-2011_Proceedings.pdf).