

# WebProtégé: Supporting the Creation of ICD-11

Sean M. Falconer, Tania Tudorache, Csongor Nyulas, Natalya F. Noy, Mark A. Musen

Stanford Center for Biomedical Informatics Research, Stanford University, US  
{sfalc, tudorache, nyulas, noy, musen}@stanford.edu

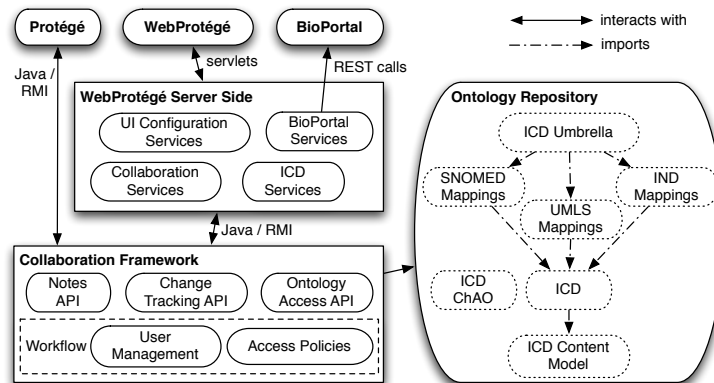
**Abstract.** WebProtégé is a highly customizable Web interface for browsing and editing ontologies, which provides support for collaboration. We have created a customized version to support the World Health Organization with the collaborative development of the 11th revision of the International Classification of Diseases (ICD-11). Our demo will present this customized version and focus on how content creation and collaboration is being supported in WebProtégé for the development of ICD-11.

## 1 Introduction

The International Classification of Diseases (ICD) is a public global standard that organizes and classifies information about diseases and related health problems [4]. Health officials use ICD in all United Nations member countries to compile basic health statistics, to monitor health-related spending, and to inform policy makers. In the United States, use of the ICD is also a requirement for all medical billing. ICD has therefore a major impact on many aspects of health care all over the world.

In 2007, the WHO initiated the 11th revision of ICD. Several ambitious goals were set for this version (details in [2]). One such goal is to allow the ICD to become a multi-purpose classification for a much larger number of usages. Previous versions of ICD were strictly classification hierarchies used for statistical purposes. To meet the new revision goals, ICD-11 will use OWL to create a rich formal representation. Another key difference between ICD-11 and previous versions is that the development process of ICD-11 will use a Web-based open process powered by collaboration and social features. That is, similar to Wikipedia, the WHO hopes that a large number of medical experts will contribute to the content of ICD-11.

Our group has been working closely with the WHO to provide the technical support for these ambitious goals. We have created a customized version of WebProtégé specifically designed to support the ICD authoring process. In [2], we discuss in detail the use of Semantic Web technologies for the revision of ICD. Our demo will showcase features of the customized WebProtégé such as content creation and collaboration. For the remainder of this paper, we present the architecture and highlight features of the user interface.



**Fig. 1.** An architecture diagram of WebProtégé used in the ICD context. WebProtégé front-end uses the services in the WebProtégé server side to display information to the user. The WebProtégé server side and the desktop client Protégé connect to the *Collaboration Framework* to access the ontology and the collaboration services. The *Ontology Repository* stores the ontologies available to the clients.

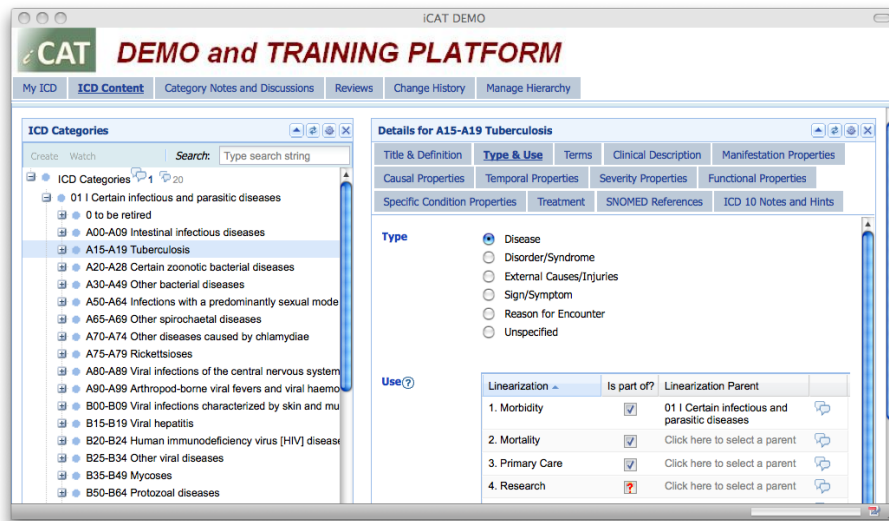
## 2 WebProtégé— Authoring Platform for ICD-11

WebProtégé is a pluggable and extensible platform for authoring ontologies. The tool was designed such that it could be customized for each project according to its own requirements. This is particularly important for the ICD revision project, as it is still in its infancy, many fundamental issues and requirements are undefined. Our tools need to be adaptable on the fly when changes are made to the content model, user-interface requirements, and workflow. In the following sections, we briefly describe the architecture and highlight some of the features of the user interface.

### 2.1 Architecture

Figure 1 shows a high level architecture diagram and the interaction of the software components for the customized version of WebProtégé. The core functionality of the application is supported by the Protégé server, which provides access to the ontology content, such as retrieving and changing classes, properties and individuals in the ontology. The ontologies that the server accesses are stored in a database on the server side. The Collaboration Framework [3] provides the collaboration and ontology access services: *Ontology Access API*, *Notes and Discussions API*, and *Change Tracking API*.

Both WebProtégé and the “traditional” Protégé desktop client can connect through the Collaboration Framework to interact with the ontology. Support for importing terms from existing terminologies was one of the requirements for the ICD-11 revision. In order to search external biomedical terminologies and to import terms from these terminologies, WebProtégé accesses BioPortal, a repository of about 200 biomedical ontologies and terminologies [1]. BioPortal provides REST service access that enables search across different ontologies and access to information about specific terms.



**Fig. 2.** The ICD authoring tool using WebProtégé. Each tab contains one or more panels, called *portlets* that can be arranged by drag-n-drop. The left hand-side portlet shows the disease class hierarchy of the ICD ontology. The right panel shows the uses (linearizations) of the selected disease in the tree, in this case *Tuberculosis*.

## 2.2 The WebProtégé User Interface

WebProtégé is a web portal, inspired by other portals, such as myYahoo or iGoogle. Our vision is to enable users to build a custom user interface by combining existing components in a form that is appropriate for their project. The user interface is composed of tabs, either predefined ones or user-defined. A new tab is an empty container in which users can add and arrange by drag-n-dropping portlets. A portlet is a user interface component that provides some functionality. For example, the Class tree portlet displays the class hierarchy in an ontology and has support for class level operations (create and delete class, move class in hierarchy, etc.). WebProtégé is extensible and has a plugin infrastructure.

For the ICD-11 revision project, we had to take into consideration that most of the ICD editors are domain experts, mainly medical doctors, who do not have backgrounds in ontology engineering. We configured the customized version of WebProtégé to use simple Web-based entry forms with fields that correspond to the attributes of the ICD content model. We have tried to ensure that the user interface does not look like an ontology editing environment, but is a customized experience for the domain experts.

The customized WebProtégé user interface is shown in Figure 2. The user interface is organized in a series of tabs: MyICD, ICD Content, Category Notes and Discussion, Reviews, Change History and Manage Hierarchy tab. Each tab presents a certain piece of functionality to the user.

The main functionality of WebProtégé is the support for browsing and editing ontologies on the Web. Editing support is available in the ICD Content tab.

Medical experts from all over the world are using the system to edit ICD-11 simultaneously. Each change by an editor is immediately committed and visible to other editors of ICD-11.

Another key piece of functionality is collaborative support. This is critical in such a large distributed project. Editors can use WebProtégé to add notes to classes, properties, and individuals in the ontology. This allows authors to raise questions and discuss different issues that arise during editing. The WHO plans to use a peer review process to ensure quality of the ICD content model. In the current version, WebProtégé supports a prototypical implementation of this feature. A user with appropriate permissions can request the review of a disease description.

WebProtégé uses a declarative user interface where the user interface components and layout is specified in an XML configuration file. The configuration can be changed on the fly, allowing the interface to be updated without re-compiling or re-deploying the application. This feature provides great flexibility for customizing WebProtégé.

### 3 Summary

We have briefly presented a customization of WebProtégé, a web-based tool for distributed collaborative development of ontologies. The WHO is using WebProtégé as the primary development environment for ICD-11. Developers of other terminologies within the WHO Family of International Classifications (WHO-FIC) are beginning to use WebProtégé as well. In the demo session, we will demonstrate the main features and benefits of WebProtégé, specifically how to create and edit the ICD content model, how to create notes and discussion threads, and import terms from Bioportal.

### References

1. N. F. Noy, N. H. Shah, P. L. Whetzel, B. Dai, M. Dorf, N. Griffith, C. Jonquet, D. L. Rubin, M.-A. Storey, C. G. Chute, and M. A. Musen. Bioportal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Research*, 10.1093/nar/gkp440, 2009.
2. T. Tudorache, S. Falconer, C. Nyulas, and M. A. Musen. Will semantic web technologies work for the development of icd-11? In *International Semantic Web Conference (ISWC2010)*, 2010.
3. T. Tudorache, N. F. Noy, and M. A. Musen. Supporting collaborative ontology development in Protégé. In *Seventh International Semantic Web Conference, ISWC 2008*, Karlsruhe, Germany, 2008.
4. WHO. International classification of diseases, manual of the international statistical classification of diseases, injuries and causes of death: 10th revision. Technical report, 1993.