What Users Want: Collaborative Development of Ontologies

Natasha Noy
Stanford University

In the Protégé group, we have been getting more and more requests recently to instrument Protégé to support collaborative development of ontologies. Sure enough, different users mean very different things by “collaborative ontology development.” Some users simply want to access and edit the same ontology from different distributed locations (something Protégé enabled users to do for a long time). Others request the ability to suggest changes, discuss modeling decisions, put out proposals for changes, and to vote on them (something we are working on right now).

In fact, as we get into specifics when discussing the requirements posed by users, we learn that there is a large continuum of workflows for collaborative ontology development. On the one end of the spectrum is the model where everyone accesses the same version of an ontology and changes are immediately visible to everyone (e.g., the client–server mode of Protégé and the prototype for collaborative development in Protégé\(^1\)). In another model, users can have their personal “sandbox” space and integrate their changes with the “master” version when they are comfortable with their privately developed model components, or when they get the permission to do so (e.g., schema editing in freebase\(^2\)). At the other end of the spectrum is the model where users always edit their local version, checking out an ontology module and integrating it back into the “master” version later, after resolving the inevitable conflicts (e.g., TDE from Apelon\(^3\)). Each of these models poses its own challenges in terms of collaboration support it requires. In addition, whatever the model, there are many questions we still don’t have answers to, simply because collaborative development of ontologies and the supporting tools are fairly nascent. For example, what is the level of expressive power for an ontology language that users can comfortably edit collaboratively? What are the workflows that support such editing: who can make changes and when, how do users reach consensus on the model. Do we need support only for shared development space, or for private spaces that link into the shared space as well?

There are also features that we do know users want, regardless of the mode of collaboration. These features include ability to annotate ontology component and changes; ability to create marginal notes and have discussions as an integral part of ontology development. Support for provenance of the data and annotations is critical: users must know who changed the concepts, when and why; have access to the history of each concept. For ontologies developed in large open groups, ways to establish trust and credibility are essential. Ability to create a view of the ontology that corresponds to the user’s personal view is also helpful. Finally, without explicit support in the tools for different workflow models for developing ontologies and vocabularies collaboratively, such tools are unlikely to be used in large projects.

Good as tools for collaborative ontology development may become, we still need infrastructure to ensure that users do not start developing an ontology from scratch any time they need one. Having well annotated and easily accessible ontology repositories is essential to ontology reuse. Such repositories should be not only conduits for open ontologies, but also enable users to provide peer reviews of ontologies, report on their experiences for specific ontologies when they used them for specific tasks, establish linkages between concepts in the ontologies and so on (see for example the BioPortal for the National Center for Biomedical Ontologies\(^4\)). Issues of trust and credibility, provenance of information, personalized views are just as pertinent for ontology repositories as they are for ontology-development tools.

\(^1\)http://smi-protege.stanford.edu/collab-protege/collab-protege.html
\(^2\)http://www.freebase.com
\(^3\)http://apelon.com/products/tde.htm
\(^4\)http://www.bioontology.org/ncbo/faces/index.xhtml