COBrA-CT: e-Science tools for ontology curation

Stuart Aitken

stuart@inf.ed.ac.uk

http://www.aiai.ed.ac.uk/project/cobra-ct

Bio-ontologies in OWL: editing, sharing and curating.

Bio-ontologies formalise our knowledge of molecular function, biological process, anatomy and cell type across the phyla. As with databases, bio-ontologies are created, published and updated over time. Ontology curation is typically a joint effort between subject matter experts, database curators and experimentalists. Hence tools for ontology curation should support sharing ontologies, annotating changes to them, and visualising those changes over time.
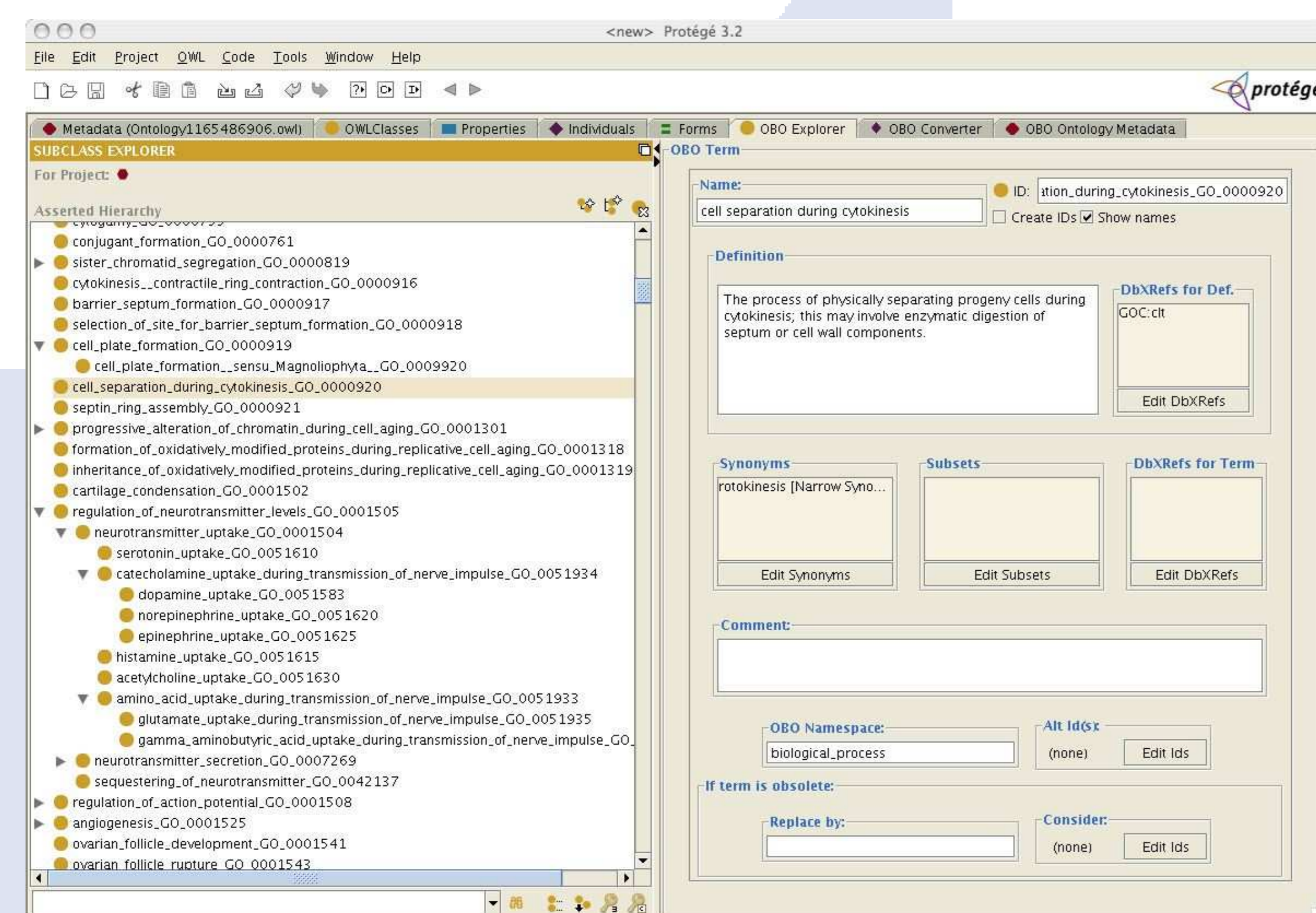

OBO Explorer tab in Protégé

The COBrA-CT project is developing tools for editing OBO ontologies (shown above) and tools for managing ontologies in the Web Ontology Language, OWL. The ontology manager has been implemented using Grid technology: an eXist database is connected with OGSA-DAI grid middleware to store OWL ontologies.

Because their use on the web will require ontologies to be represented in OWL (albeit at a level invisible to their users), we have participated in reaching a standard for transitioning Open Biomedical Ontologies (OBO) to OWL (including an agreed semantics for the transition and an agreed representation for synonyms, subsets and database cross-references). This has been carried out in collaboration with Chris Mungall (Howard Hughes Medical Institute/GO Consortium), Dilvan Moreira and Nigam Shah (Stanford) and others. These standards have been incorporated into the OBO Explorer editor.

The generic ontology editor Protégé is used for this work in order to have compatibility with, and reuse of, tools developed in the wider ontology research community. But because Protégé cannot handle OBO term annotations, our tools are essential for end-users to be able to work with bio-ontologies that have been reformatted in OWL.
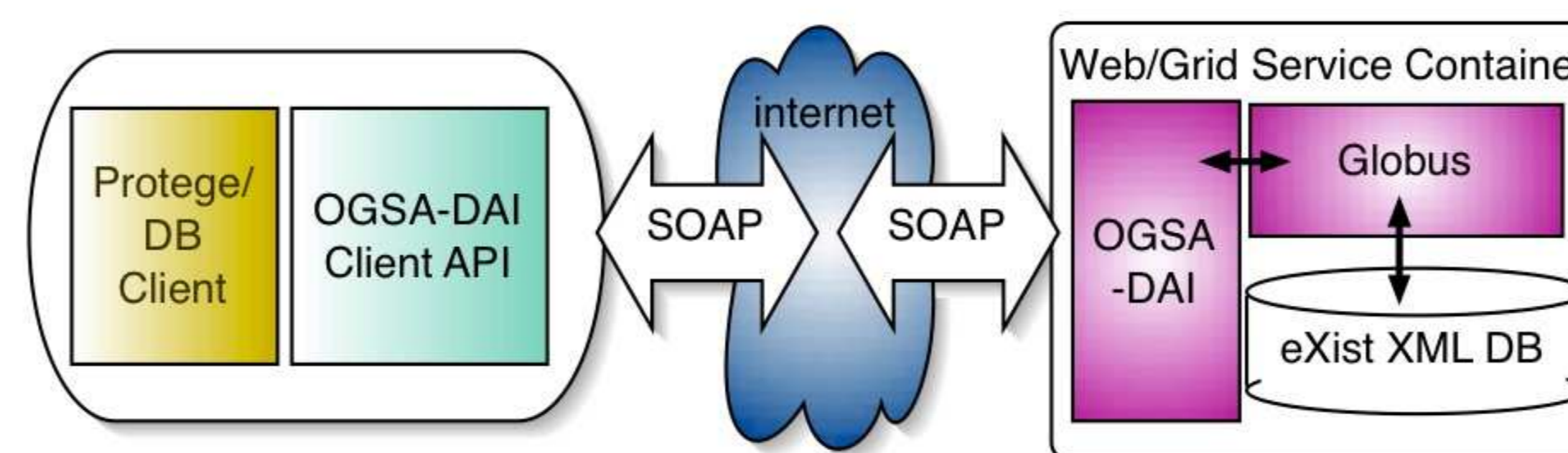
Client/Grid server architecture

OWL ontologies are represented in XML. For archiving and sharing bio-ontologies in joint development efforts, we have developed ontology XML database and client tools. The screenshot (shown right) is of the stand-alone client application: the COBrA-CT Version Manager. A Protégé version is also planned. These tools allow a registered user to access ontologies (upload and download current and past versions) and share access to them: Ontologies can be either kept private, shared with named persons, or declared public and therefore shared will all users. The ontology server uses OGSA-DAI to access and query the eXist XML database, while the client uses OGSA-DAI activities.
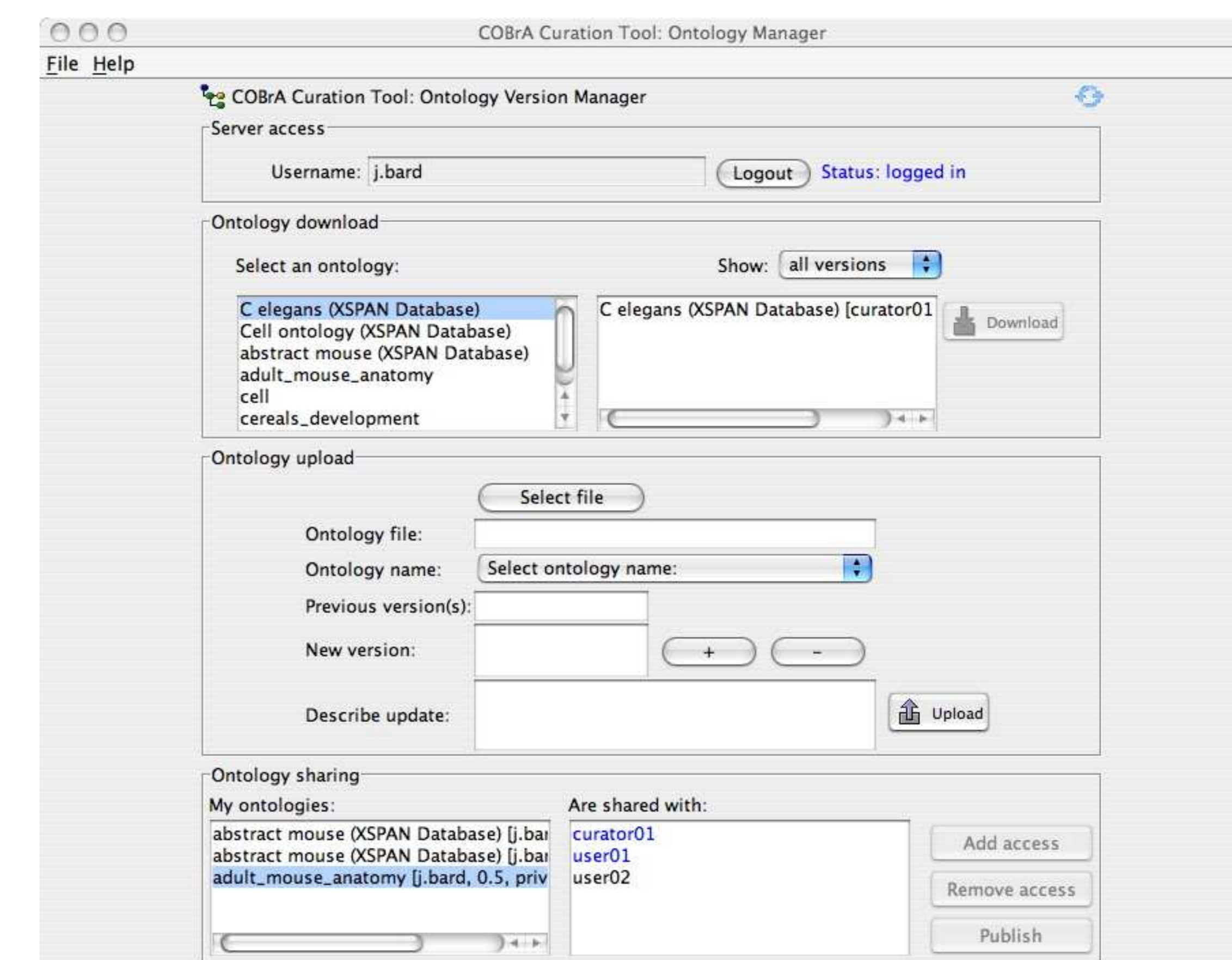
COBrA-CT Version Manager

Future work will develop improved visualisations of the part-of hierarchies that bio-ontologists are familiar with. These get translated into Description Logic definitions that are more complex to interpret. Additional future research goals are to explore visualisations of the differences between ontology versions, and to investigate more efficient storage of OWL ontologies as XML documents.

References

Aitken, S. and Chen, Y. COBrA and COBrA-CT: Ontology Engineering Tools, to appear in *Anatomy Ontologies for Bioinformatics: Principles and Practice* eds. Burger, A., Davidson, D., and Baldock, R.

Aitken, S., Korf, R., Webber, B. and Bard, J. COBrA: A Bio-Ontology Editor. *Bioinformatics* 2005 21(6):825-826.

Reference:	BB/D006473/1
PI:	Bonnie Webber
CO PIs:	Jonathan Bard, Wenfei Fan
Researchers:	Stuart Aitken, Yin Chen

School of informatics