NeuroLex.org – A semantic wiki for neuroinformatics based on the NIF Standard Ontology

Stephen D. Larson, Sarah M. Maynard, Fahim Imam, Maryann E Martone,

Department of Neurosciences, University of California, San Diego,
9500 Gilman Drive, La Jolla, CA 92093
slarson@ncmir.ucsd.edu, smaynard@ncmir.ucsd.edu, mimam@ucsd.edu,
maryann@ncmir.ucsd.edu

Abstract. Bridging the domain knowledge of a scientific community and the knowledge engineering skills of the ontology community is still an imperfect practice. Within the field of neuroscience, we have tried to close this gap by presenting an ontology through the medium of a wiki where each page corresponds to a class. By opening it to the World Wide Web, we have made the process of maintaining a ~20,000 concept neuroscience ontology (NIFSTD), more collaborative.

Keywords: Neuroscience, Ontology, Semantic wiki, NIF

1 Introduction

The neuroscience community needs its basic domain concepts organized into a coherent framework. Ontologies provide an important medium for reconciling knowledge into a portable and machine readable form. For many years we have been building community ontologies for neuroscience, first through the Biomedical Informatics Research Network and now through the Neuroscience Information Framework (NIF) projects (http://neuoinfo.org). These projects resulted in the construction of a large modular ontology, constructed by importing existing ontologies where possible, called NIFSTD[1]. One of the largest roadblocks that we encountered was the lack of tools for domain experts to view, edit and contribute their knowledge to NIFSTD. Existing editing tools were difficult to use or required expert knowledge to employ. By combining several open source technologies related to semantic wiki and NIFSTD, we have created NeuroLex.org, the first semantic wiki for neuroscience

2 Methods

NeuroLex.org is built on top of the open source Semantic Mediawiki platform[2]. This allows classes, properties, and instances to be represented within a wiki that is

easily editable by both registered an anonymous users and allows the content of that wiki to be exported as RDF/OWL. Semantic Mediawiki makes querying the ontology via properties or class hierarchy very straightforward. In addition, we have incorporated tools such as Semantic Forms, which allow the ontology classes to be edited as a form rather than as a free form wiki page with special text mark-up. Some of the fields in the form support autocomplete, which allows users to populate those fields with other classes from the ontology.

The content contributed to the NeuroLex is not directly added to the NIFSTD but is incorporated into the NIFSTD OWL file by a knowledge engineer after curation by the NIF ontology group. This ensures that the NeuroLex reflects the most up-to-date version of the lexicon content, but also allows the contents NIFSTD OWL file to be properly vetted before release. The curation process also involves some translation by the knowledge engineer to create more formal definitions (e.g. simple part-of relations translated into OWL restrictions) than are possible via the NeuroLex. Several practices have been put into place to keep the NeuroLex and the OWL file in sync, such as the ability to list all recent changes to all pages, and the ability to distinguish changes made by vetted scientists versus community members who are anonymous.

In order to import RDF/OWL into the Semantic Mediawiki, the python library known as pywikipediabot was used to upload each concept in the NIFSTD as its own page. Several pywikipediabot-based scripts were subsequently written to bulk edit the content on the Wiki. While we evaluated Ontowiki as a possible alternative to the Semantic Mediawiki which would simplify the RDF/OWL import / export issue, Semantic Mediawiki was a more robust and high-performance platform and therefore the cost of bridging the RDF/OWL gap ourselves was worth it.

3 Results

NeuroLex.org has evolved into a powerful platform for collaboratively maintaining and extending the NIFSTD ontology. We have been able to incorporate user feedback and create custom views of the ontology content with very rapid turnaround. Table 1 shows some key metrics that we have collected on its usage. They demonstrate that the NeuroLex is receiving a significant amount of traffic (~320 hits a day), has had a significant increase in hits since its launch, and has a reasonable number of contributing scientists making edits each day.

Table 1. Key metrics for usage of NeuroLex.org

Contributing neuroscientists	~12
Average edits per weekday	~25
Average hits per weekday	~320
% increase hits 01/09 – 10/09	815%
% hits from search engines	81%

4 Conclusion

We conclude that the Semantic Mediawiki is a good starting point for the collaborative maintenance of ontologies. Other groups are also using a similar approach, e.g. BioMedGT3. While we are still working through some issues, e.g. synchronizing the NIFSTD with the content on NeuroLex, exporting and importing OWL, and bulk uploading concepts, we believe that semantic wikis are a good tool for providing community contribution and feedback to projects like NIF.

Acknowledgments. With thanks to the Mediawiki project, the Semantic Mediawiki Project, Ontoprise, Yaron Koren, and Anders Larsson.

References

1. Bug, W.J., Ascoli, G.A. et al. (2008) The NIFSTD and BIRNLex vocabularies: building comprehensive ontologies for neuroscience. Neuroinformatics 6, 175-194 (2008)
2. Semantic Mediawiki, http://semantic-mediawiki.org
3. BiomedGT, http://biomedgt.nci.nih.gov