

Ontology-based Web Services for Semantic Applications

Patricia L. Whetzel*, Nigam H. Shah, Natalya F. Noy, Clement Jonquet, Adrien Coulet, Cherie Youn, Michael Dorf and Mark A. Musen

Stanford Center for Biomedical Informatics Research, Stanford University
Stanford, California 94305-5479

Researchers have turned to the Semantic Web to annotate, integrate, retrieve, and interpret disparate knowledge. Ontologies provide the domain knowledge to drive these processes and the successful creation of semantic applications in the health and life sciences require Web services that provide access to bioontologies. The National Center for Biomedical Ontology (NCBO), one of the seven National Centers for Biomedical Computing created under the NIH Roadmap, has developed BioPortal, which provides access to one of the largest repositories of biomedical ontologies available via both Web browsers and Web services (via REST services).¹ The BioPortal *Ontology* Web services allow programmatic access, download and traversal of ontologies in software applications, the *NCBO Annotator* Web service “tags” text automatically with terms from BioPortal ontologies, and the *NCBO Resource Index* provides an ontology-based index of public, online data resources.

The *Ontology* Web services provide access to ontologies, their metadata, ontology versions, downloads, navigation of the class hierarchy (parents, children, siblings) and details of each term. Developers can use these Web services in their software applications to access this functionality. For example, as used by the Microsoft Word Add-in for Ontology Recognition², to mark up a research article at the time of writing² or as used by ISAcreeator³, the ECG Gadget⁴, and Jinx⁵, to annotate experimental metadata.

The *NCBO Annotator* Web service processes text to recognize relevant biomedical ontology terms. Users can customize the Web service to limit results to a particular ontology (e.g. SNOMED CT) or to a certain UMLS semantic type (e.g. T017 for ‘Anatomical Structure’) as well as many other parameters documented at: http://www.bioontology.org/wiki/index.php/Annotator_Web_service. The entity recognition engine, MGREP, was developed by the National Center for Integrative Bioinformatics and is combined with the BioPortal *Ontology* Web services to create the *NCBO Annotator* Web service. The availability of this functionality makes the task of creating ontology-based annotations accessible for any biomedical

researcher by using the Web service directly through programmatic access or by using Add-ins such as the NCBO Annotator Excel Addin.

The *NCBO Annotator* Web service has also been used by NCBO to generate ontology-based indexes of several online biomedical data repositories (e.g., GEO, ClinicalTrials.gov, dbGaP, DrugBank, PharmGKB, and Reactome) resulting in the *NCBO Resource Index*. These indexes are generated based on using the textual metadata of the database record as input to the *NCBO Annotator* and the database records are linked together via shared ontology annotations. These linkages take advantage of the semantic relationships within the ontology, including subsumption relationships among ontology entities and mappings between entities in different ontologies. The *NCBO Resource Index* Web service provides biomedical investigators a mechanism for programmatic search of the index using ontology terms. For example, one can search for all experiments and clinical trials corresponding to ‘malignant melanoma’ from GEO and ClinicalTrials.gov.

In conclusion, the variety of ontology-based Web services provided by NCBO provides a “1-stop shop” for access to biomedical ontologies and use in semantic applications. More information on these Web services can be found at [www.bioontology.org/wiki/index.php/NCBO REST services](http://www.bioontology.org/wiki/index.php/NCBO_REST_services).

REFERENCES

- Noy NF, Shah NH, Whetzel PL, Dai B, Dorf M, Griffith N, Jonquet C, Rubin DL, Storey MA, Chute CG, Musen MA. (2009) BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.*, **37**, W170–3.
- Shah NH, Jonquet C, Chiang AP, Butte AJ, Chen R, Musen MA. (2009) Ontology-driven indexing of public datasets for translational bioinformatics. *BMC Bioinformatics*, **10**, Suppl 2:S1.

¹<http://bioportal.bioontology.org>

²Microsoft Word Add-in for Ontology Recognition
<http://ucsdbiolit.codeplex.com/>

³ISAcreeator <http://isatab.sourceforge.net/isacreeator.html>

⁴ECG Gadget http://wiki.cvrgrid.org/index.php/ECG_Gadget_User_Guide

⁵Jinx <http://ncmir.ucsd.edu/downloads/jinx.shtm>