Ontorat web server for automatic generation and annotations of new ontology terms

Zuoshuang Xiang, Yu Lin, Yongqun He
University of Michigan Medical School, Ann Arbor, USA

Introduction

The Web Ontology Language (OWL) has been widely used for ontology development. However, development of a new ontology in OWL is often time consuming and requires specialized knowledge in ontology, OWL, and specific domain science. One frequently observed phenomenon during ontology development is that often we need to create a large number of new ontology terms that follow the same design patterns of logical definitions and axioms. Manual addition of these terms is time consuming, error prone, and often boring. Ontorat is developed to facilitate this process.

Ontorat is developed based on the Ontology Design Patterns (OCPs) in the field of ontology engineering. Ontorat uses the *class expressions with variables*, which is defined as "OWL class expressions but allowing variables at positions of class expressions. The range of a variable can either be a named class, class expression, or any subtype of it as produced by the corresponding rule in OWL." (Noppens and Liebig, 2009). Ontorat offers a web-based platform for writing up class expressions with variables with the aim to quickly generate a large number of new ontology classes or annotate existing classes in a specific target ontology.

The development of Ontorat is also inspired by the Ontology for Biomedical Investigations (OBI) project QTT (Quick Term Templates) (Rocca-Serra, et al., 2011). The OBI QTT procedure is based on the concept of a Quick Term Template (QTT), a template in a spreadsheet format that can be converted into an OWL file. The QTT procedure has been implemented in MappingMaster, a plugin program in the Protégé-OWL editor. Similar to MappingMaster, Ontorat uses an Excel or tab-delimited spreadsheet file as input. Ontorat also relies on the pattern program using the OWL Manchester Syntax, a new syntax designed for writing OWL class expressions (http://www.w3.org/TR/owl2-manchester-syntax/). Different from MappingMaster, Ontorat is implemented as a web-based application. The MappingMaster plugin works in Protege 3.4 but does not work in Protege 4.x. Ontorat provides a novel feature of automatic assignment of unique ontology identifiers. In addition, Ontorat, but not MappingMaster, can generate annotations for ontology terms.

Features and Usage

Ontorat (http://ontorat.hegroup.org) provides a user-friendly web form for data input. The Ontorat input include: (i) a target ontology in OWL format, (ii) the data file (an Excel file or tab-delimited text file), (iii) purpose of the new axiom generation, (iv) assignment of annotations using Manchester OWL Syntax, (v) assignment of equivalent classes, (vi) assignment of superclasses, (vii) definitions of ontology terms used in the above three assignments, (viii) prefix of term URI, (ix) information for automatically assigning term IDs, including prefix, number of digits, and start ID. The Ontorat output is an OWL file. Using the Protégé ontology editor (http://protege.stanford.edu/), an Ontorat output OWL file can be directly visualized and imported (or merged) in the target ontology (*e.g.*, VO) using the OWL import function.

As a demonstration, we will demonstrate how Ontorat is used to facilitate the development of the Vaccine Ontology (VO) (He, et al, 2009). VO is a community-based ontology in the domain of vaccine and vaccination. There are hundreds of licensed animal vaccines used in the USA. Ontorat is able to import these animal vaccines and their annotations into VO efficiently.

The Ontorat web server provides a robust and scalable approach for automatically generating new ontology terms and their annotations. Ontorat supports efficient ontology enrichment and expansion. The design patterns can be standardized and reused, allowing domain experts and data curators to contribute actively to the ontology development without knowing the specifics of OWL.

Acknowledgements: This research is supported by a NIH R01 grant (1R01AI081062).

References

Ozgur A, et al. Mining of vaccine-associated IFN-γ gene interaction networks using the Vaccine Ontology. *Journal of Biomedical Semantics*. 2011, 2(Suppl 2):S8.

Noppens, Y and Liebig, T. Ontology Patterns and Beyond Towards a Universal Pattern Language. *Proceedings of the Workshop on Ontology Patterns (WOP 2009)*, collocated with the 8th International Semantic Web Conference (ISWC-2009), Washington D.C., USA.

Rocca-Serra, P., et al. (2011) Overcoming the Ontology Enrichment Bottleneck with Quick Term Templates. *Journal of Applied Ontology*, 6, 13-22.