

Ontorat: Automatic generation and editing of ontology terms

Yongqun He^{1*}, Jie Zheng², Yu Lin¹

¹ University of Michigan, Ann Arbor, Michigan, USA; ² University of Pennsylvania, Philadelphia, PA, USA

ABSTRACT

The web-based Ontorat (<http://ontorat.hegroup.org/>) program has been developed to automatically generate and edit ontology terms from a spreadsheet format input (Excel or tab-delimited file) based on an ontology design pattern (ODP). The Ontorat web interface is intuitive and suitable for users with basic background in Manchester ontology format scripting, the same language used in the Protégé OWL editor. Ontorat also collects ODPs and provides templates and sample data for future reuse. As a demonstration, Ontorat is applied to automatically generate assay terms with axioms and annotations added into the Ontology for Biomedical Investigations (OBI).

1 INTRODUCTION

Manually developing a new ontology can be very time-consuming. To reduce the time in ontology development, we can first develop an Ontology Design Pattern (ODP) (Noppens and Liebig, 2009), apply the ODP to generate an Excel or text file including ODP-oriented data, and then transform the structured data into an OWL format file. Ontorat is a web-based tool to perform such a task (Xiang et al., 2015).

Ontorat is developed based on the ODP strategy, and its development was inspired by the Quick Term Templates (QTT) procedure generated by the developers of the Ontology for Biomedical Investigations (OBI) (Rocca-Serra et al., 2011). Ontorat is able to convert a QTT template in a spreadsheet into an OWL file. Ontorat also includes other features as described elsewhere in this article.

In this software demo, we will introduce general Ontorat design and workflow, describe features of Ontorat, and provide use case demonstrations to show how Ontorat is used to facilitate new ontology term addition and existing ontology term editing.

2 ONTORAT SOFTWARE DESIGN

From the Ontorat web page, a user can enter setting options and upload the input data file via the Ontorat web input form. The input data file is generated by populating a predesigned template file guided by the ODP as mentioned above. The setting options specify the ontological meanings of the columns in the input data file and axioms between terms. After accepting the input data file and setting options from the user, the web server (via a PHP script) will be able to execute two operations: 1) generation of new ontology classes with logical axioms and annotations, or 2) adding new axioms to existing ontology terms. The Ontorat server will process the user's requests and generate either an

Ontorat settings file or an OWL output file. The Ontorat settings file can be stored and reused later (Fig. 1).

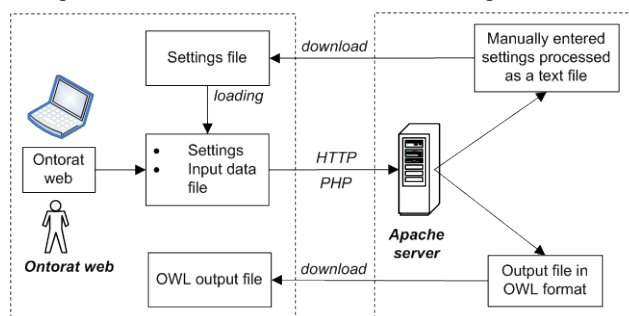


Fig. 1. Ontorat architecture and workflow.

3 ONTORAT WEB INTERFACE

Ontorat includes a user-friendly web form for adding setting options, uploading a data input file, clicking for execution, providing debugging information, and downloading the OWL output.

Input data file: Based on a pre-defined ODP, an Ontorat template file is generated to include all term and annotation types needed for defining a target term. The template file can then be filled up with specific terms and annotations for each type. Ontorat supports both Excel and tab-delimited text input files.

Ontorat settings: Ontology axioms are represented using Manchester OWL Syntax (Horridge et al., 2006). The axiom settings can be added one by one via the Ontorat web form or uploaded from an Ontorat setting text file in an Ontorat-specific setting file format. Ontorat can generate the setting file based on the setting inputs via the Ontorat web form.

Ontorat output file: The OWL output file can be visualized using the Protégé ontology editor (<http://protege.stanford.edu/>) and imported into the target ontology using the OWL import function.

4 ONTORAT NEW TERM GENERATION

One major function of Ontorat is to generate multiple new ontology terms based on an ODP for a specific ontology.

Fig. 2 provides a demo on how Ontorat can be used to generate a list of assays based on a design pattern (Fig. 2A) out of the Ontology for Biomedical Investigations (OBI) (Brinkman et al., 2010). Fig. 2B provides a portion of the Ontorat setting file. Like the Protégé OWL editor, Ontorat

* To whom correspondence should be addressed: yongqunh@umich.edu

also uses the Manchester OWL Syntax (Horridge et al., 2006) for scripting ontology axioms. For example, Ontorat uses the following Manchester syntax expression:

```
'has participant' some
<http://purl.obolibrary.org/obo/{columnM}>
```

to express the relation that a newly generated term has an axiom of 'has participant' some device (which is located in the column M of the Excel input file). Each row in the Excel file (Fig. 2C), starting from row 2, includes the information for generating a new ontology term. The output file can be displayed using Protégé OWL editor (Fig. 2D). The detailed files associated with this use case are available on the Ontorat web page: <http://ontorat.hegroup.org/designtemplates/assay/obi-assay.php>.

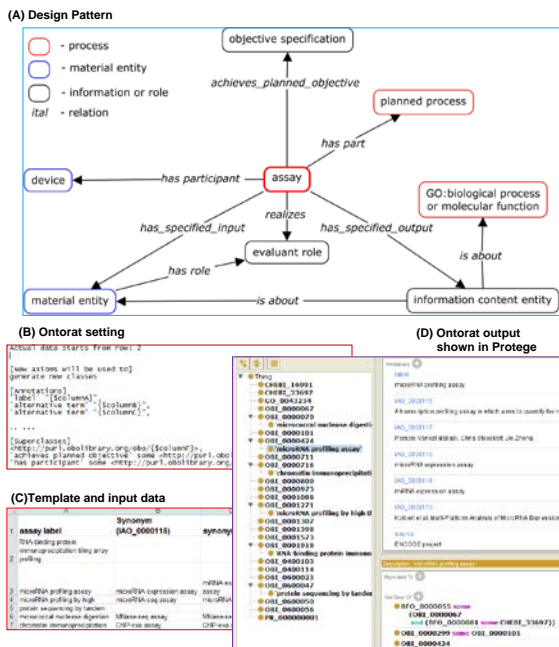


Fig. 2. Ontorat usage for enriching OBI assay terms. (A) Design pattern; (B) Ontorat setting in text format. This text format can be loaded to Ontorat web page directly. (C) Ontorat template together with input data; (D) The output OWL file displayed using the Protégé-OWL editor.

5 ONTORAT EXISTING TERM EDITING

Ontorat also supports editing existing terms by adding new axioms (e.g. annotations). For example, Ontorat was used to automatically add definition source and term editor annotations to over 50 ontology classes in the Biobank Ontology (<https://code.google.com/p/biobank-ontology/>). Such information was not included in the ontology at the early development stage. To add the annotations, the following settings were used in the Ontorat annotations input section:

```
'definition editor' "{columnC}",
'definition source' "{columnD}"
```

The Ontorat input and output OWL files for this use case are available on the Ontorat website: <http://ontorat.hegroup.org/designtemplates/biobank/index.php>.

6 COLLECTION OF DESIGN PATTERNS AND TEMPLATES

Ontology design patterns (ODPs) are reusable modeling solutions for ontology development. Ontorat has collected many ODPs and corresponding templates (<http://ontorat.hegroup.org/designtemplates/>). For each case, Ontorat provides an ODP diagram, an Excel template, a setting file, and an example with populated template data and output OWL file. These ODPs and templates can be reused to support fast and reproducible ontology development.

7 SOURCE CODE AND LICENSE

The Ontorat source code is openly available on GitHub: <https://github.com/ontoden/ontorat>. The Ontorat source code license is Apache License 2.0.

8 SUMMARY

With ever increasing needs of ontology development and applications, the web-based Ontorat program provides a timely platform for generating and annotating ontology terms based on design patterns.

ACKNOWLEDGEMENTS

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

REFERENCES

Brinkman, R.R., Courtot, M., Derom, D., Fostel, J.M., He, Y., Lord, P., Malone, J., Parkinson, H., Peters, B., Rocca-Serra, P., Ruttenberg, A., Sansone, S.A., Soldatova, L.N., Stoeckert, C.J., Jr., Turner, J.A., Zheng, J., and Consortium, O.B.I. (2010). Modeling biomedical experimental processes with OBI. *J Biomed Semantics* 1 Suppl 1, S7.

Horridge, M., Drummond, N., Goodwin, J., Rector, A.L., and Wang, H.H. (Year). "The Manchester OWL syntax", in: *OWL: Experiences and Directions (OWLED '06)*, eds. B.C. Grau, P. Hitzler, C. Shankey & E. Wallace: CEUR Workshop Proceedings), 10 pages.

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

Rocca-Serra, P., Ruttenberg, A., O'connor, M.J., Whetzel, P.L., Schober, D., Greenbaum, J., Courtot, M., R.R., B., S.A., S., R., S., Consortium, T.O., and Peters, B. (2011). Overcoming the ontology enrichment bottleneck with quick term templates. *Applied Ontology* 6, 13-22.

Xiang, Z., Zheng, J., Lin, Y., and He, Y. (2015). Ontorat: Automatic generation of new ontology terms, annotations, and axioms based on ontology design patterns. *Journal of Biomedical Semantics* 6, 4 (10 pages).