The OntoDev Suite of Ontology and Data Integration Tools

James A. Overton [1]

[1] *Knocean Inc., Toronto, Ontario, Canada*

Abstract

The OntoDev Suite (https://ontodev.com) of open source software brings together modular libraries and applications for ontology development and scientific data integration, with special emphasis on open science and the Open Biological and Biomedical Ontologies (OBO) community. The suite builds on the success of ROBOT to include data cleaning, ontology-driven validation, development and curation workflows, and more. We strive to make small, focused tools that work well with other best-in-class software, languages, and platforms.

The flagship of the OntoDev Suite is ROBOT (http://robot.obolibrary.org), a library and command-line tool for automating a wide range of common ontology development and application tasks, such as importing, converting, extracting, merging, reasoning, querying, template-based term building, quality control, and release processes.

ROBOT and the other OntoDev tools are often used together with a workflow execution tool such as GNU Make, and a version control system such as a GitHub repository. DROID (https://github.com/ontodev/droid) makes these powerful tools approachable to a wider audience by both hosting the tools and providing a simple, customizable web interface.

Most scientific ontology and data integration projects make use of spreadsheets at some point, but this can undercut the use of version control systems, which work best with text files. COGS (https://github.com/ontodev/cogs) and AXLE (https://github.com/ontodev/axle) make it easy to synchronize Google Sheets and Excel (respectively) with tab- or comma-separated value files (TSV or CSV) in a version control system. VALVE (https://github.com/ontodev/valve) is a table validation tool that also uses tables to configure the validation rules, with special emphasis on ontology-driven validation. By combining VALVE with COGS and AXLE you can see your data and your validation rules side-by-side, highlight and filter cells that fail validation, and see validation messages and suggested fixes as comments on those cells.

In addition to spreadsheets, many scientific projects also use relational databases and the Structured Query Language (SQL). It is often useful to include scientific ontologies and scientific data in the same SQL database. RDFTab (https://github.com/ontodev/rdftab.rs) and LDTab (https://github.com/ontodev/ldtab.clj) convert RDF/OWL into SQL tables, and Gizmos (https://github.com/ontodev/gizmos) and Gadget (https://github.com/ontodev/gadget) which are Python libraries for extracting, exporting, searching, and browsing ontologies via SQL (using RDFTab and LDTab, respectively).

OntoDev software is distributed under the BSD 3 Clause open source license. ROBOT is used by over one hundred ontology and ontology-related projects. Nine projects currently use DROID (https://droid.ontodev.com), and several of these use COGS, AXLE, and/or VALVE to work with and validate ontology term templates and other tables. OntoDev tools are used as part of several open scientific databases, including the Immune Epitope Database (http://www.iedb.org), Chemical Effects on Biological Systems (https://cebs.niehs.nih.gov/cebs/), the International Hundred+ Cohorts Consortium (https://ihccglobal.org), Computational Modelling of Immunology: Pertussis Boost (https://www.cmi-pb.org), and more.

Keywords

ontology, software, ontology development

International Conference on Biomedical Ontology (ICBO 2022),
September 25-28, 2022, Ann Arbor, Michigan, USA
EMAIL: james@overton.ca (A. 1)
ORCID: 0000-0001-5139-5557 (A. 1)