

A New Ontology Lookup Service at EMBL-EBI

Simon Jupp, Tony Burdett, James Malone, Catherine Leroy, Matt Pearce, Julie McMurry, Helen Parkinson

European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBI),
The Genome Campus, Hinxton, Cambridge, CB10 1SD

Abstract. The Ontology Lookup Service (OLS) hosted at the EMBL European Bioinformatics Institute has been providing ontology search and visualisation services for over ten years. In this time the range and diversity of ontologies has changed dramatically. One of the major shifts has been the increasing use of the W3C Web Ontology Language (OWL) for representing biomedical ontologies. The OLS has been reengineered in order to accommodate the OWL standard and provide a wider range of ontology-based services to the community.

1 Introduction

The use of bio-medical ontologies for the annotation, integration and analysis of biological data is now well established in bioinformatics. The range and diversity of ontologies has increased dramatically over the last ten years and community efforts such as the OBO foundry have been instrumental in coordinating this activity. The demand for unified mechanisms for accessing large collections of ontologies has led to a number of dedicated ontology repositories becoming available to the community. The EMBL-EBI Ontology Lookup Service (OLS)¹ was one of the earliest public ontology browsers, offering searching and browsing services along with Web services for programmatic access to ontologies. Similar services are offered by other ontology repositories, such as BioPortal² and OntoBee³, that are notable for their support for the Web Ontology Language (OWL). Support for OWL in OLS has not been available until now.

The prevalence of bio-medical ontologies published in OWL along with changes to the OBO format (to the point where it is now considered a subset of OWL) has necessitated a reengineering of the OLS system to support this established standard. This has been coupled with an increasing demand for ontology-based services within EMBL-EBI as the role of ontologies becomes integral to the interoperability of data across the institute. In order to meet this demand the OLS has been redesigned and aims to better connect the ontologies to the data they describe.

¹ <http://www.ebi.ac.uk/ontology-lookup/>

² <http://bioportal.bioontology.org>

³ <http://www.ontobee.org>

2 Design and Implementation

This first iteration of the new OLS provides the following functionality: 1) An ontology crawler that detects when external ontologies have changed 2) A search engine for querying ontology term meta-data 3) Web services for querying the structure of the ontology 4) A user interface for exploring and visualising the ontologies 5) A versioning system to track ontology evolution.

The EU funded DIACHRON project⁴ has been developing a framework for managing the evolution of data on the Semantic Web. The move to OWL means that OLS can utilise DIACHRON technology to monitor ontologies for changes so that the system can be updated when new ontologies become available. DIACHRON is also providing tooling to detect changes at the level of individual terms between ontology releases. OLS has adopted the OBO foundry YAML-LD format for registering meta-data about how to access ontologies.

The OLS search index has been developed as part of the BioSolr⁵ project that is exploring optimal ways to index ontologies in open source search engines such as Apache-Solr. OSL provides a generic ontology indexing application for Solr that is also available as a standalone Solr application. To complement the search index we provide a RESTful API to ontology and entity (class, property and instance) meta-data. The ontology is persisted in a Neo4J database and the REST API was built using the Spring Data framework. This API supports the generation of Hypertext Application Language (HAL) JSON documents that contain links to related document for programmatically navigating the ontology structures. All of the ontologies in OLS will also be loaded into the EBI RDF platform⁶ and made available via a SPARQL endpoint.

Finally a new Website has been developed that includes Javascript based ontology widgets for searching and visualising ontologies. These widgets can be easily integrated into third-party websites.

2.1 Availability

The OLS beta can be viewed at <http://www.ebi.ac.uk/ols/beta>. The source code is available from <https://github.com/EBISPOT/OLS>.

3 Funding

We acknowledge funds EMBL and the DIACHON, BioMedBridges and CORBEL projects funded by the European Commission (grant agreement numbers 601043, 284209, 654248). BioSolr is funded by the BBSRC (BB/M013146/1).

⁴ <http://www.diachron-fp7.eu>

⁵ <http://www.flax.co.uk/blog/2014/06/11/biosolr-building-better-search-for-bioinformatics/>

⁶ <http://www.ebi.ac.uk/rdf/>