

# OntoSearch: a Semantic Web Service to Support the Reuse of Ontologies<sup>1</sup>

Edward Thomas, Yi Zhang, Joe Wright, Craig McKenzie, Alun Preece, Derek Sleeman  
Department of Computing Science, University of Aberdeen  
Aberdeen, AB24 3UE  
{ethomas, yzhang, jwright, cmckenzi, apreece, dsleeman} @  
csd.abdn.ac.uk

*This document provides an overview of the development of OntoSearch, an ontological search engine designed to help users find RDF based ontological information on the Semantic Web. It uses the Google API to search several million documents on the Semantic Web with supporting servlets to provide summary information & various visualizations of the documents found.*

## 1. Overview and Motivation

Finding a suitable ontology from the Internet is a hard task because of the difficulty of separating ontological data from the mass of instance data on the Semantic Web and quickly evaluating its suitability. There is still no good tool to handle this problem. Google offers a powerful web search engine. However, with regard to ontology searching, it has its own problems, such as a lack of visualisation facilities. Using the Google API<sup>2</sup> give us a chance to develop our own tool (OntoSearch/2) to search the relevant ontology files to meet the user's requirements.

## 2. Conceptual design of OntoSearch

The concept of OntoSearch is to facilitate knowledge reuse by allowing the huge number of ontology files available on the Semantic Web to be quickly and effectively searched using an online tool. The search results contain both ontologies and other SWDs to allow knowledge engineers to browse relevant documents to find supplementary information on a subject if a complete ontology is not available.

To allow fast development of a system it was decided to use the Google search engine through its API to search the available Ontologies in the RDF(S)<sup>3</sup>, OWL<sup>4</sup> and DAML (+OIL)<sup>5</sup> representational formalisms. Once the Google results have been returned, each document listed is examined and summary information identifying where the terms matched the returned documents and statistical data about the size of

---

<sup>1</sup> This work is part of the Advanced Knowledge Technology (AKT) [1] project, which is funded by EPSRC

<sup>2</sup> Google APIs: <http://www.google.com/apis>

<sup>3</sup> RDF(S): <http://www.w3.org/TR/rdf-schema/>

<sup>4</sup> |OWL: <http://www.w3.org/TR/owl-features/>

<sup>5</sup> DAML (+OIL): <http://www.w3.org/TR/daml+oil-reference>

the ontology is presented to the user to allow quick evaluation on the suitability of a large number of potential ontologies and other Semantic Web Documents (SWDs).

### **3. Outline of implementation**

OntoSearch is implemented as a number of specialised servlets which work together to provide the functionality of OntoSearch. This architecture allows additional features to be plugged into the basic system quickly and easily.

The SearchServlet is responsible for querying the Google database using the Google API object, it takes the users query either through a web form or as an HTTP GET request and returns a list of matching ontology files from Google, either as an HTML file or an RDF file (depending on a variable set in the user's request).

If an HTML file is returned, this contains several embedded iframe elements which each reference the DetailsServlet. This accesses each document and examines the ontology returned and displays a list of where the search terms were found in the ontology, gives general statistics about the size of the ontology and lists the namespaces used in the ontology.

Once a file has been found which appears to suit the purpose, a variety of visualisations are available through links on the results page, either in a format provided by Notation 3 (N3) or as a hyperbolic tree showing the structure of the ontology. The visualisation is particularly useful for examining an entire ontology quickly, browsing through multiple classes and linkages to quickly see the underlying structure.

### **4. Discussion**

The facilities provided by the current version of OntoSearch have been used by several people within AKT, and over 2000 searches have been performed. However there are several areas of the system which we intend to develop further.

To address these points, a new version of OntoSearch is currently being designed which will preserve the unique functionality of OntoSearch and to address the shortcomings identified to provide a more valuable tool. It will allow the data to be searched for a far wider set of parameters (including some aspects of structure), and the results returned will be more focused as an enhanced scoring mechanism will reflect the important aspects of the user's request.

The current HTTP API will be extended to allow the new features to be made available to other applications as a web service, and the basic RDF results which can be accessed at OntoSearch will be extended to include the full range of information available through the web interface.

### **References**

1. Advanced Knowledge Technology (AKT project) <http://www.aktors.org>
2. Zhang Y, Vasconcelos W, and Sleeman D. OntoSearch: An Ontology Search Engine (AI-2004). The Twenty-fourth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, Cambridge.