

OntologyPlugin

[This text comes straight from the README.TXT written by John Xing]

Ontology plugin is a contribution from Michael J Pan. Currently it is used to do one kind of query refinement as implemented in refine-query-init.jsp and refine-query.jsp (both are called by search.jsp).

By default, ontology plugin is compiled, but query refinement based on it is ignored in search.jsp. To enable query refinement, do the following:

Download Jena 2.6.X here >>> <http://sourceforge.net/projects/jena/files/>

- Move Nutch-1.2.war to server /webapps and boot up server
- Copy ontology.jar from /WEB-INF/classes/plugins/ontology to /WEB-INF/lib
- Do the same with all jar dependencies within your clean Jena distribution
- Edit search.jsp by uncommenting so as to enable refine-query.jsp and refine-query-init.jsp
- Edit refine-query.jsp to the following

o Line 44 – String search URL = “../search.jsp?”+searchquery;

- Edit nutch-site.xml and add ontology plug-in to plugin.includes property

· specify absolute URI(s) of owl files to property extension.ontology.urls in ./conf/nutch-default.xml (or better, ./conf/nutch-site.xml).
N.B. your OWL files need to be in RDF/XML format for the parser implementation to succeed. It appears that files need to be hosted online as locally hosted files are not read properly by the plug-in.

· Finally ALL ontology files must be in RDF/XML format to adhere to the ontology parser specified in ontology.jar class files. It is important that OWL files are in this format.

Further to this, it is important to consider that this plug-in will probably not be supported in subsequent Nutch releases as both searching and indexing is being delegated to Solr. This is something you should consider if you plan to use this feature on a long term basis. It would be nice to have this ported as a Solr requestHandler plugin implementation ;0)

Possible issue:

In previous releases of Nutch <1.2

If search.jsp fails with this or similar error:

```
root cause

java.lang.NoSuchFieldError: actualValueType
    at
com.hp.hpl.jena.datatypes.xsd.XSDDatatype.convertValidatedDataValue(XSDDatatype.java:371)
```

It is because jena and tomcat are using conflicting versions of the same xerces library. To solve this, one needs to update tomcat's xerces library. Here's a reference:

<http://jena.sourceforge.net/jena-faq.html#general-1>

<<< [PluginCentral](#)