

# iRDQL - Imprecise Queries Using Similarity Joins for Retrieval in Ontologies

Abraham Bernstein and Christoph Kiefer  
Department of Informatics, University of Zurich  
{bernstein,kiefer}@ifi.unizh.ch

Traditional semantic web query languages support a logic-based access to the semantic web. They offer a retrieval (or reasoning) of data based on facts. On the traditional web and in databases, however, exact querying often provides an incomplete answer as queries are over-specified or the mix of multiple ontologies/modelling differences requires “interpretational flexibility.” Therefore, similarity measures or ranking approaches are frequently used to extend the reach of a query.

This prototype demonstration extends this idea to the semantic web. It introduces iRDQL – a semantic web query language with support for similarity joins. It is an extension to RDQL, which is a SQL-like RDF data query language for RDF triples in Jena models, that enables the user to query for similar resources in ontologies. A similarity measure is used to determine the degree of similarity between two semantic web resources. Similar resources are ranked by their value of similarity and returned to the user. We show how iRDQL allows to extend the reach of a query by finding additional results.



?s1	?s2	sim(?s1, ?s2)
http://www.ontologyportal.org/translations/SUMO.owl.txt#Woman	http://www.ontologyportal.org/translations/SUMO.owl.txt#Man	1.0
http://www.ontologyportal.org/translations/SUMO.owl.txt#Woman	http://www.ontologyportal.org/translations/SUMO.owl.txt#Woman	1.0
http://www.ontologyportal.org/translations/SUMO.owl.txt#Man	http://www.ontologyportal.org/translations/SUMO.owl.txt#Man	1.0
http://www.ontologyportal.org/translations/SUMO.owl.txt#Man	http://www.ontologyportal.org/translations/SUMO.owl.txt#Woman	1.0
http://www.ontologyportal.org/translations/SUMO.owl.txt#Human	http://www.ontologyportal.org/translations/SUMO.owl.txt#Human	1.0
http://www.ontologyportal.org/translations/SUMO.owl.txt#Woman	http://www.ontologyportal.org/translations/SUMO.owl.txt#Plant	0.5
http://www.ontologyportal.org/translations/SUMO.owl.txt#Woman	http://www.ontologyportal.org/translations/SUMO.owl.txt#Microorganism	0.5

Figure 1: The figure shows the graphical user interface of our iRDQL query tool where the user can enter queries in a text field. In this particular example, every resource of type *human* is compared to every resource of type *organism*. The results are ranked by their similarity and displayed to the user in the table below.

In the demonstration we will show our current implementation of iRDQL. It includes a graphical user interface that allows the user to load several ontologies in which similar resources can be found. The user of the tool can enter iRDQL queries in a text field which subsequently get processed by the query engine. The results can either be presented as the textual query evaluation engine output or in the form of a table (as shown in Figure 1), where each of the query variables and the similarity measure’s output get entered into their own columns. The table can be sorted by each column by clicking into its header.

All similarity measures implemented in SimPack, our generic Java library of similarity measures for the use in ontologies can be used in the iRDQL queries. Thus, the demonstration will also give a more detailed explanation of SimPack. The tool is still under development and will, therefore, still change in appearance and have additional functionality.

Initial results of using iRDQL indicate that it is indeed useful for extending the reach of the query and that it is able to improve recall without overly sacrificing precision.