

A Search Interface for Researchers to Explore Affinities in a Linked Data Knowledge Base

Laurens De Vocht¹, Erik Mannens¹, Rik Van de Walle¹,
Selver Softic², and Martin Ebner³

¹ Ghent University - iMinds, Multimedialab
Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium

{laurens.devocht,erik.mannens,rik.vandewalle}@ugent.be

² Virtual Vehicle Research Center - Area Information and Process Management
Inffeldgasse 21a, 8010 Graz, Austria
selver.softic@tugraz.at

³ Graz University of Technology, IICM - Institute for Information Systems and
Computer Media
Inffeldgasse 16c, 8010 Graz, Austria
martin.ebner@tugraz.at

Abstract. Research information is widely available on the Web. Both as peer-reviewed research publications or as resources shared via (micro)blogging platforms or other Social Media. Usually the platforms supporting this information exchange have an API that allows access to the structured content. This opens a new way to search and explore research information. In this paper, we present an approach that visualizes interactively an aligned knowledge base of these resources. We show that visualizing resources, such as conferences, publications and proceedings, expose affinities between researchers and those resources. We characterize each affinity, between researchers and resources, by the amount of shared interests and other commonalities.

1 Introduction

Research 2.0 as adaptation of the Web 2.0 for researchers defines researchers as main consumers of information. Typically researchers define queries with a set of keywords when searching for information related to their work, for example using Google or digital archives such as PubMed. Linked Data technologies offer an entity based infrastructure to resolve the meanings of the keywords and the relations between them. Combining keyword resolution and resource expansion with Linked Data entities and filtering the results with personal preferences enhances the search precision. Currently many researchers have a Social Media account, such as on Twitter or Mendeley. We use these accounts to personalize the search. Our interface supports searching for scientific events, authors or groups of authors, as well as finding publications, and proceedings. The interface uses a search engine which relies on Linked Data knowledge base containing research related and personal information.

2 Real-time Keyword Disambiguation

We chose a real-time keyword disambiguation to guide the researchers in expressing their research needs. We do this by allowing users to select the correct meaning from a drop down menu that appears below the search box. Presenting candidate query expansion terms in real-time, as users typed their queries, can be useful during the early stages of the search [1]. In this case it is very important that the users understand meaning of the suggested terms. Therefore we use an as straightforward as possible representation of the keyword mappings as shown in Figure 1.

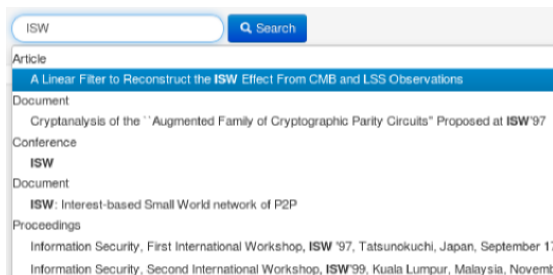


Fig. 1: Mapping of keywords to Linked Data entities.

3 Exploring Resources

Researchers can improve the definition of their “intended” search goal over several iterations. Each time a combination of various resources is visualized. The visualization suggests new queries: they are generally most useful for refining the system’s representation of the researcher’s need. In case they have no idea which entity to focus on or what topic to investigate next they get an overview of possible entities of interest, like points of interest on a street map. By profiling their activities and contributions on Social Media and other platforms such as their own research publications, the affinity with the proposed resources is enhanced. An iteration can consist of either one of two actions:

1. **Query Expansion:** The user expands the query space by clicking the results retrieved by initial keyword based search. The resolution of results happens based upon the properties of Linked Data like *rdf:label*, *owl:sameAs*, *rdf:seeAlso*, *dc:title* or *dc:description*.
2. **Additional Query Formulation:** Additional query expansion happens either through adding further keywords as well as through keyword combinations already entered where the back-end tries to deliver additional results based upon connection paths between the resources.

4 Visualizing Relations between Resources

We find relations between resources after matching the input given by the researcher in the knowledge base. With the delivery of first results, our engine expands the query and enhances the context. For this purpose we used a model and an implementation that builds upon on our earlier work on the “Everything is Connected” engine (EiCE) [2] and semantic profiling [3].

In the visualization we emphasize the affinities by showing, on a radial map [4], how the current focused entity is related to the other found entities. It is based on the concept of affinity that can be appropriately expressed in visual terms as a spatial relationship: proximity [5]. We additionally express the amount of unexpectedness as *novelty* of a resource in each particular search context. To further enhance the guidance of users during search we have used two other visual aids:

1. **Color:** Every entity has a type and associated unique color. For a certain result set the user gets an immediate impression of the nature of the found resources.
2. **Size:** We rank each entity according to novelty and relation to the context and enlarge those that should attract attention from the researcher first. A goal of the search is to explore information not seen before which makes it difficult to define an accurate search goal. Besides allowing to search specific entities, our visualization facilitates exploratory browsing. This is particularly useful when information seeking with unclear defined search targets [6].

Figure 2 shows how researchers can track the history of their search: the explored relations are marked red and clearly highlight the context of a search. Researchers can click on a list of resources they have searched to focus the visualization. A screencast of the search interface is available online⁴. In this screencast, we show how researchers interact with the search interface and the above described visualization.

5 Conclusions

We have developed an interface for personalized search in a social driven knowledge base for researchers. Combining the latest Linked Data technologies with an advanced indexing and path finding system, EiCE and Web 2.0 technologies (such as JQuery and Django). The result is a semantic search application providing both a technical demonstration and a visualization that could be applied in many other disciplines beyond Research 2.0. The main contribution of our work is, besides retrieving resources from Linked Data repositories, allowing users to interactively explore relations between the resources and find out the affinity with each resource.

⁴ http://semweb.mmlab.be/search_interface_for_researchers

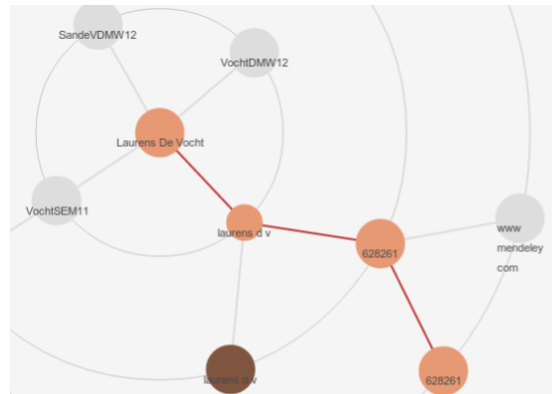


Fig. 2: A red line marks the explored relations in the visualized search context.

6 Acknowledgement

The research activities that have been described in this paper were funded by Ghent University, iMinds (Interdisciplinary institute for Technology) a research institute founded by the Flemish Government, Graz University of Technology, the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders (FWO-Flanders), and the European Union. The authors would like to acknowledge the financial support of the “COMET K2 - Competence Centres for Excellent Technologies Programme” of the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT), the Austrian Federal Ministry of Economy, Family and Youth (BMWFJ), the Austrian Research Promotion Agency (FFG), the Province of Styria and the Styrian Business Promotion Agency (SFG).

References

1. White, R.W., Marchionini, G.: Examining the effectiveness of real-time query expansion. *Information Processing & Management* **43**(3) (2007) 685–704
2. De Vocht, L., Coppens, S., Verborgh, R., Vander Sande, M., Mannens, E., Van de Walle, R.: Discovering meaningful connections between resources in the web of data. In: *Proceedings of the 6th Workshop on Linked Data on the Web (LDOW)*. (2013)
3. De Vocht, L., Van Deursen, D., Mannens, E., Van de Walle, R.: A semantic approach to cross-disciplinary research collaboration. *International Journal of Emerging Technologies in Learning (iJET)* **7**(S2) (2012) 22–30
4. Yee, K.P., Fisher, D., Dhamija, R., Hearst, M.: Animated exploration of dynamic graphs with radial layout. In: *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS)*. (2001)
5. Pintado, X.: The affinity browser. In: *Object-oriented software composition*. Prentice Hall (1995) 245–272
6. Pace, S.: A grounded theory of the flow experiences of web users. *International journal of human-computer studies* **60**(3) (2004) 327–363