

## **Intra- and interdisciplinary cross-concordances for information retrieval**

In the final phase of the project KoMoHe<sup>1</sup> at GESIS, a major evaluation effort to test and measure the effectiveness of the vocabulary mappings in a Knowledge Organization System (KOS) enhanced environment was conducted (Mayr & Petras, 2008). The paper reports on the evaluation results of different intra- and interdisciplinary cross-concordances.

The most important question is how effective and helpful the mappings are in an actual search. In a search portal like sowiport.de<sup>2</sup> which provides unified access to a variety of databases providing high quality content such as bibliographic metadata, full texts and data sets, the question becomes crucial whether cross-concordances can enable a distributed search with improved result sets. Can they bridge the differences in human language in order to facilitate a seamless search with the same query across different disciplinary databases? Our thorough Information Retrieval (IR) evaluation focuses on the quality of the associated search via cross-concordances.

Leveraging the cross-concordances should expand the search space, correct ambiguities and imprecision in the query formulation and therefore, in general, find more relevant documents for a given query. The retrieval results improve for all cross-concordances, however, interdisciplinary cross-concordances cause a significantly higher (positive) impact on the search results. For all cross-concordances in the test scenarios, more relevant documents were found compared to the query types without the use of cross-concordances; in particular cases, the retrieved set was even more precise (e.g. increase in precision as well).

In this paper we want to explore pros and cons of intra- and interdisciplinary cross-concordances<sup>3</sup>: Why are cross-concordances in one discipline (e.g. KOS in social sciences) less effective for IR than interdisciplinary mappings (e.g. social science KOS mapped to a psychology KOS)? Where and how can automatic mapping methods help in KOS mapping projects (Lauser et al., 2008)?

---

<sup>1</sup> The KoMoHe was funded by the German Federal Ministry for Education and Research, grant no. 01C5953. <http://www.gesis.org/forschung-lehre/programme-projekte/informationwissenschaften/projektuebersicht/komohe/>

<sup>2</sup> <http://sowiport.de/>

<sup>3</sup> Our current research on value-added service for digital libraries is funded by DFG, grant no. 658/6-1. <http://www.gesis.org/forschung-lehre/programme-projekte/informationwissenschaften/projektuebersicht/irm/>

## References

Lauser, Boris; Johannsen, Gudrun; Caracciolo, Caterina; Keizer, Johannes; Hage, Willem Robert van; Mayr, Philipp (2008): Comparing human and automatic thesaurus mapping approaches in the agricultural domain. pp. 43-53. In: Greenberg, Jane; Klas, Wolfgang (eds.): Metadata for semantic and social applications: Proceedings of the 8. International conference on Dublin Core and Metadata Applications. Berlin: Uni.-Verl. Göttingen. URL: <http://edoc.hu-berlin.de/conferences/dc-2008/lauser-boris-43/PDF/lauser.pdf>

Mayr, Philipp; Mutschke, Peter; Petras, Vivien (2008): Reducing semantic complexity in distributed digital libraries: Treatment of term vagueness and document re-ranking. In: Library Review 57, No. 3, pp. 213-224. URL: [http://www.ib.hu-berlin.de/~mayr/arbeiten/mayr-et al\\_LR08.pdf](http://www.ib.hu-berlin.de/~mayr/arbeiten/mayr-et al_LR08.pdf)

Mayr, Philipp; Petras, Vivien (2008): Cross-concordances: terminology mapping and its effectiveness for information retrieval. In: 74th IFLA World Library and Information Congress. Québec, Canada URL: [http://www.ifla.org/IV/ifla74/papers/129-Mayr\\_Petras-en.pdf](http://www.ifla.org/IV/ifla74/papers/129-Mayr_Petras-en.pdf)