

Khresmoi: semantics in medical search

H. Müller¹, A. Hanbury², A. Roberts³, K. Pentchev⁴, and V. Momtchev⁴

¹University of Applied Sciences Western Switzerland (HES-SO),

²Technical University of Vienna, Austria,

³University of Sheffield, United Kingdom,

⁴Ontotext, Bulgaria

Abstract. Khresmoi is an EU funded project on medical information analysis and search for three user groups: patients, general practitioners and radiologists. Semantics are used in the project backend based on a knowledge engine and support several aspects of the search engine.

Keywords: Medical information search, semantic search

1 The Khresmoi Search System

The Khresmoi¹ project aims to improve medical information access for three user groups with different information needs: patients, general practitioners and radiologists. Several prototypes are available online [1] and many aspects of medical information access have been addressed such as information trustability, readability of information, user interface design, ranking and the access to medical image data. An overview of the system is given in Figure 1. Semantic extraction and use of an underlying knowledge base are important techniques of Khresmoi



Fig. 1. Overview of the Khresmoi project.

Semantic Search Traditional full-text search does not address complex information seeking behaviour, such as exploratory search. The precision oriented approaches typically employed do not address the information needs of end-users engaged in exploration, where both recall and interactive retrieval capabilities are important. In this environment, semantic search is establishing itself as the next generation search paradigm, in that it better meets a wider range of information needs. Semantic search addresses the challenges of exploration by finding

¹ <http://khresmoi.eu/>

information that is based both on the presence of words, and crucially on their meaning. Khresmoi is indexing semantically annotated health and medical texts, in order to create an index to drive semantic search. Typical indexing solutions store plain text tokens, not semantic annotations. We are using Mimir² to store and retrieve semantic annotations. Mimir indexes plain text tokens using the MG4J search engine. Semantic annotations are stored in Ontotexts OWLIM. This means that we can query over full text, annotations, and constraints on those annotations in terms of ontologies in the Khresmoi knowledge base (KB).

Khresmoi Knowledge Base Data integration is the process of ensuring interoperability between data sources by providing a unified view of the information contained in them. In general, knowledge consistency is spread over four nested levels of heterogeneity: systematic, syntactic, structural and semantic. The Khresmoi KB contains a set of semantically integrated data sources suited for the domains of clinical radiologists, medical practitioners and the general public. The storage engine behind the KB is the proprietary Ontotext semantic repository OWLIM-SE. It persistently stores various types of information such as ontologies, instance data, unstructured text, semantic annotations (i.e. links between ontology instances and text annotations), image meta-data, multilingual semantic networks, user feedback and other system runtime data. The KB exposes this data via a standard SPARQL endpoint and the Sesame API. In addition, RESTful web services (WS) for resource disambiguation, semantic search and modality classification are provided. In order to make the operation of the services efficient and to more tightly integrate the knowledge from the different datasets in the KB, this was enriched with additional instance mappings and relations extracted from external sources and unstructured fields. Versioning of each of the sources in the KB is implemented.

2 Discussion and Conclusion

Khresmoi has developed a user-driven approach to medical information search for three user groups. One of the results is that content understanding is essential for many aspects and particularly for professional searchers. Semantics help to analyze and link the content. The evaluation of search quality of semantic search in user tests is foreseen in a next step.

Acknowledgments This work was supported by Khresmoi (257528).

References

1. Aswani, N., et al.: Khresmoi: Multimodal multilingual medical information search. In: Proceedings of Medical Informatics Europe (2012)

² <http://gate.ac.uk/mimir/>