



ASE: A Search Engine for Semantically Annotated Documents

Presenter: Dr Alejandro Metke-Jimenez, The Australian e-Health Research Centre, CSIRO

Audience

Health departments, Clinical bodies and Vendors that need to search large collections of medical records written in free-text.

Objectives

Attendees will learn about a new hybrid search model that allows combining traditional keyword-based search with semantic search to improve medical record retrieval.

Abstract

Ideally, medical records should contain structured data using a common, standardised model. This would allow issuing and solving very precise and unambiguous queries. However, a large number of medical records contain key information available only as free-text. Dealing with this unstructured information is a challenging but necessary task that has contributed to advances in the areas of Natural Language Processing and automated semantic annotation.

Existing tools are able to automatically annotate unstructured free-text with concepts from an ontology, such as SNOMED CT, but the quality of the results is far from perfect. Current search tools provide limited support for hybrid queries that refer to both the textual content and the semantic annotations. We introduce a new hybrid query language that addresses this gap by combining semantic queries with keyword-based queries.

This presentation will:

- demonstrate how the free-text sections of medical records can be automatically annotated with SNOMED CT concepts with tools such as MetaMap [1],
- discuss what are the limitations of these automated methods,
- introduce our proposed hybrid model and show examples of its usage,
- explain why this model can improve retrieval performance,
- showcase a demo of the Annotation Search Engine (ASE), our prototype implementation of the proposed hybrid search model. The prototype has been tested using the BLULab NLP repository, a collection of 81,617 de-identified clinical records from multiple U.S. hospitals during 2007 [2].

References

1. Aronson, A. R., & Lang, F. M. (2010). An overview of MetaMap: historical perspective and recent advances. *Journal of the American Medical Informatics Association*, 17(3), 229-236.
2. Koopman, B., Bruza, P., Sitbon, L., Lawley, M. (2011). Evaluation of medical information retrieval. *Poster proceedings of the 34th annual international ACM SIGIR conference on Research and development in information retrieval*.