Subword-based Semantic Retrieval of Clinical and Bibliographic Documents

P. Daumke[1,2], S. Schulz[2], M. L. Müller[3], W. Dzeyk[4], L. Prinzen[5], E. J. Pacheco[6,7], P. Secco Cancian[6], P. Nohama[6,7], K. Markó[1,2]

[1]AVERBIS GmbH, Freiburg, Germany

[2]Freiburg University Hospital, Medical Informatics Group, Freiburg, Germany

[3]Freiburg University Hospital, Department of Dermatology, Freiburg, Germany

[4]German National Library of Medicine, Cologne, Germany

[5]Bertelsmann Stiftung, Gütersloh, Germany

[6]Paraná Catholic University, Health Informatics Laboratory, Curitiba, Brazil

[7]Universidade Tecnológica Federal do Paraná, CPGEI, Curitiba, Brazil

Summary
Objectives
The increasing amount of electronically available documents in bibliographic databases and the clinical documentation requires user-friendly techniques for content retrieval.

Methods
A domain-specific approach on semantic text indexing for document retrieval is presented. It is based on a subword thesaurus and maps the content of texts in different European languages to a common interlingual representation, which supports the search across multilingual document collections.

Results
Three use cases are presented where the semantic retrieval method has been implemented: a bibliographic database, a department EHR system, and a consumer-oriented Web portal.

Conclusions
It could be shown that a semantic indexing and retrieval approach, the performance of which had already been empirically assessed in prior studies, proved useful in different prototypical and routine scenarios and was well accepted by several user groups.

The full English version can be found in /Methods of Information in Medicine 2010 49 2: 141-147/ or http://www.schattauer.de/de/magazine/uebersicht/zeitschriften-a-z/methods/issue/special/manuscript/12759/show.html.