ENRICHMENT OF CLINICAL PRACTICE GUIDELINES BY SEMANTICS

Katharina Kaiser and Silvia Miksch [1])

Medical science is an area, which offers a great potential for the application of computer-support. Amongst others a huge amount of free text documents exist that could be used in a far better and more efficient way, if they are better structured, categorized, and computer-interpretable.

In our group we deal with so called *clinical practice guidelines* (CPGs), "systematically developed statements to assist practitioners and patient decisions about appropriate healthcare for specific circumstances" [1]. To optimally support the medical personnel approaches exist to transform these documents in a computer-interpretable (and thus -executable) format – a cumbersome and time-consuming task: it demands for detailed medical knowledge and knowledge of the complex formal methods.

We try to develop methods that automate parts of this task. Therefore, we also use semantic technologies, such as ontologies, in order to automatically process the documents, for instance, by using Information Extraction techniques. Due to the complexity of the medical area the structuring and description of the medical knowledge has already started very early. In the meantime, endless medical ontologies exist used by both physicians and computer scientists. We utilize these ontologies to semantically enrich CPGs in order to apply rules to structure and formalize them.

Thereby, ontologies provide us with the domain knowledge so we (as non-physicians) can understand the texts. Without this domain knowledge we would not be able to process the documents. A number of methods for formalizing CPGs exist that still fall back on the assistance of physicians for that part. But they are expensive and often manual semantic annotation is less accurate and it relies on the experts' skills.

Therefore, using semantic systems implicates a great advantage and facilitation. Text data evolves to information. Due to the semantic enrichment of documents they not only become computer-interpretable, but also comprehensible for non-experts. Furthermore, errors may be easier detectable.

Today, we are able to automate the modeling of CPGs in a computer-interpretable format using semantic technologies. In the near future, we will probably be able to unambiguously annotate very context-specific text data by semantics. This is especially important in such highly complex

[1] Danube University Krems, Department of Information and Knowledge Engineering in co-operation with Vienna University of Technology, Institute of Software Technology & Interactive Systems

domains as the medical science. The next great milestone will be an almost automatic processing of text data to "understand" textual information. For our work that may mean the integration of computer-interpretable CPGs in the daily routine and more precisely in patient data management systems (PDMSs). Furthermore, an almost automatic generation of computer-interpretable CPGs in terms of "living guidelines" may be possible. A "living guideline" is one that remains under scientific review on an ongoing basis, with updates published at set intervals (e.g., annually), to present up-to-date and state-of-the-art knowledge. To always apply the latest version of a CPG it is important to have it implemented immediately after publication.

References

[1]   Field, M.J. and Lohr, K.N. (eds.) Clinical Practice Guidelines: Directions for a New Program. National Academies Press, Institute of Medicine, Washington DC, 1990.