Creating Visual Summaries of Clinical Practice Guidelines with Natural Language Processing

Contact: Florian Borchert

Evidence synthesis in the form of Clinical Practice Guidelines (CPGs) serves as a basis for evidence-based decision making in clinical practice. However, CPGs today are disseminated mostly as large collections of free-text documents, complicating accessibility to humans and machines alike. Our goal is to use Natural Language Processing (NLP) to automatically extract information from the unstructured texts and transform them into a structured format.

While transforming CPGs into fully machine-readible formats, such as process models, is a long-term vision, as an initial step we want to create structured visual summaries of the landscape of CPGs targeted at medical practitioners. The British National Institute for Health and Care Excellence (NICE) has done pioneering work in this regard (see for example Figure 1), and we want to create similar graphics using CPGs from the German Guideline Program in Oncology. These guidelines are available to us as a pre-processed text corpus (GGPONC), simplifying automated processing with NLP significantly.

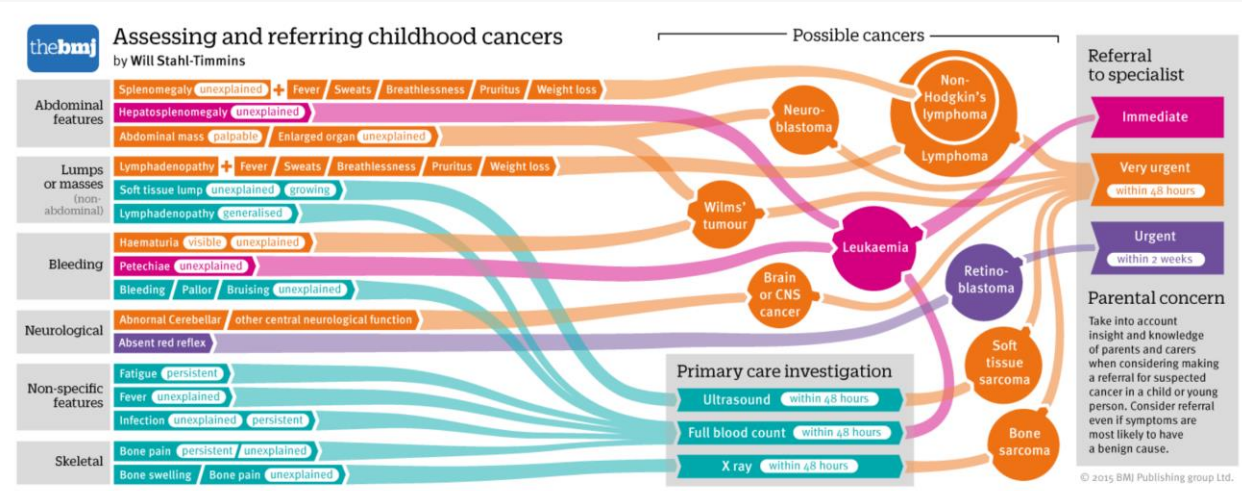

Figure 1 Visual Summary of NICE guidance, https://www.bmj.com/content/350/bmj.h3036/infographic

The goals of this thesis are to (1) identify information elements (like symptoms, diagnoses and procedures) relevant for visual summarization of guidelines, (2) enhance existing and develop new NLP solutions for German medical text to extract these elements from the GGPONC corpus and (3) use the information to automatically create visual summaries that facilitate navigating the landscape of evidence in oncology.