

The Centre for the Study of Manuscript Cultures (CSMC)
announces a Lecture
by
Prof. Dr. rer. nat. Ralf Möller
University of Luebeck

eHumanities 4.0: Pitfalls and Potentials

Manuscripts have provided the fundamental medium for information sharing for ages, and in the context of now-available commodity information technology (e.g., web, cloud computing) the tremendous importance of manuscript analysis for our society is well accepted such that scholarly interest in manuscripts has increased significantly over the last two decades. In the area of eHumanities, the insight that digital data processing can do more than just store manuscripts as electronic documents in order to make them available on request for printing or reading is ubiquitous. Given new techniques for character recognition (OCR) being applied in a data driven way, Google technologies such as string search for finding manuscripts are now standard, and are indeed successfully complemented with new computer vision technologies.

In the talk we will analyze experiences from previous EU- and DFG-funded research projects on deriving document-local symbolic content descriptions (annotations) as well as repository-integrating (holistic) latent descriptions of documents (in particular web pages) in order to demonstrate the potential of contemporary information system technology for research and applications in the area of eHumanities. Emphasizing machine learning techniques used in latest commercial tools, we will also shed light on overcoming pitfalls of relying too much on (tedious and expensive) manual work to derive annotations.

While Google derives symbolic data descriptions in order to foster the construction of sensible follow-up queries to retrieval results (see Google's Knowledge Vault project), their focus is on making business, but for niche topics, which can very well be socially important, we cannot rely on private initiatives from US companies. Slightly provocative, but in accordance with the notion "Industry 4.0", which emphasizes the impact of modern computerized information technology for industry, we sketch the idea of "eHumanities 4.0" to best support also researchers and other "knowledge workers" in their highly specialized daily work.

Thursday, 28 April 2016 at 6 pm
Room 0001, CSMC