Abstract der Dissertationsarbeit

Making Inter- and Intranet resources available in a structured way is one of the most important and challenging problems today. An underlying structure allows users to search for information, documents or relationships without a clearly defined information need. While search and filtering technology is becoming more and more powerful, the development of such explorative access methods lacks behind. This work is concerned with the development of large-scale data mining methods that allow to structure information spaces based on loosely coupled user annotations and navigation patterns.

An essential challenge, that was not yet fully realized in this context, is heterogeneity. Different users and user groups often have different preferences and needs on how to access an information collection. While current Business Intelligence, Information Retrieval or Content Management solutions allow for a certain degree of personalization, these approaches are still very static. This considerably limits their applicability in heterogeneous environments.

This work is based on a novel paradigm, called collaborative structuring. This term is chosen as a generalization to the term collaborative filtering. Instead of only filtering items, collaborative structuring allows users to organize information spaces in a loosely coupled way, based on patterns emerging through data mining.
A first contribution of the work is to define the conceptual notion of collaborative structuring as combinatorial optimization problem and to put it into relation with existing research in the areas of data and web mining. As second contribution, highly scalable, distributed optimization strategies are proposed and analyzed.
Finally, the proposed approaches are quantitatively evaluated against existing methods using several real-world data sets. Also, practical experience from two application areas is given, namely information access for heterogeneous expert communities and collaborative media organization.