Dear Reader,

the World Wide Web has become a significant target for data mining, due to several reasons: The web is a huge resource of any kind of information, the increase of commercial applications on the Web requests the extraction of knowledge from the Web, and the immense amount of data available calls for automatic means for the extraction.

The Web differs in many regards from other mining applications. Web pages consist of (sometimes structured) natural language text, calling for text mining techniques; hyperlinks provide additional structure, that can be handled with graph mining approaches; Web servers log user activities, which also can be analyzed; and the Web is very dynamic in terms of growth, content changes, and structural changes. The combination of all of these aspects makes the Web a unique setting for data mining. During the last decade, researchers have attacked many of these challenges for Web Mining.

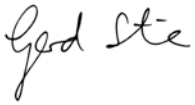
Recently, with the emergence of Semantic Web and Web 2.0, the attention of the research community has shifted to a new focus. With the Semantic Web, a more conceptual view on (Web and other) data arose, leading to the desire to discover topics and trends (which then can be captured in an ontology); and Web 2.0 platforms facilitate simplified participation of untrained users, who started to build social and topic-oriented networks, leading to the desire to discover significant substructures and communities. Even though not specifically requested in the Call for Papers, most submissions to this Special Issue addressed one of these two topics.

The discovery of trends and topics is the subject of four contributions. B. Berendt and M. Draheim present an analysis of a very large Email corpus to detect new topics. Hoser et al report on a project with Siemens AG in which they try to detect topic trends in newsgroups related to mobile phones. B. Stein and S. Meyer zu Eißen classify and evaluate different labeling strategies for document clusters. Recommending scientific literature based on semantic profiles is addressed in G. Semeraro et al. The detection and evolution of communities in an online student social networking system is the topic in T. Falkowski and M. Spiliopoulou.

A survey over these recent developments, including an overview over Web Mining foundations, is given in our introductory paper. The collection of papers is complemented by a dictionary entry on recommender systems by B. Mobasher and by a dissertation report about new developments in text categorization by S. Meyer zu Eißen. One of the next issues will include a closely related description of the TAGora project.

We want to thank all the authors of this special issue for giving exciting insights into recent developments, and the reviewers for their valuable comments. We wish all readers a pleasant lecture.

Andreas Hotho



Gerd Stumme

Schwerpunkt
Web Mining

Auszug aus: Künstliche Intelligenz, Heft 3/2007, ISSN 0933-1875, BöttcherIT Verlag, Bremen, **www.kuenstliche-intelligenz.de/order**