“Semantic Web Mining”
Workshop Proposal for ECML/PKDD-2001

1 Workshop Topic and Goals

In the *Semantic Web Mining* workshop we intend to bring together researchers and practitioners from the two fastly developing research areas *Semantic Web* and *Web Mining*. One can improve the results of Web Mining by exploiting the new semantic structures in the web. Furthermore, Web Mining can help to build the Semantic Web.

The *Semantic Web* is based on a vision of Tim Berners–Lee. He suggests to enrich the web by machine processable information which is organized on different levels (see www.w3.org/DesignIssues/Semantic.html). For Web Mining, the levels from XML and RDF to ontologies and logics are of particular interest.

Web Mining applies data mining techniques on the web. Three areas can be distinguished: Web usage mining analyzes the user behavior, web structure mining explores the hyperlink structure, and web content mining exploits the contents of the documents in the web.

In the workshop, we want to discuss the use of XML, RDF, ontologies, and logics for the three web mining areas; and the support of web mining techniques for building (XML and RDF) schemes and ontologies. There are quite a number of people from different communities that approach the field of Semantic Web Mining from different, interesting angles. The goal of the workshop is to establish communiction between these communities.

Topics of interest include, but are not limited to:

- Concept Taxonomies (e. g. product hierarchies) and Web Mining
- Using the Syntactic Layer (XML) for Semantic Web Mining
- Using the Vocabulary Layer (e. g. RDF-Schema) for Semantic Web Mining
- Using the Logical Layer (e. g. Description Logics/OIL) for Semantic Web Mining
- User Profiling, Segmentation and Interpretation with Background Knowledge
- Using Background Knowledge for Web Mining
- Metadata-based Text Mining
- Text Mining for Creating Metadata
- Learning Taxonomies and Ontologies from the Web
- Metadata Harvesting
- Information Extraction with Machine Learning

2 Intended Audience

The workshop aims at bringing together researchers and practitioners from both the semantic web and the web mining community. We have sent out last week a request for statements of interest in the workshop. Up to now, we have received 86 responses, including the following:

- “I support your proposal. This is a promising research direction.” Guus Schreiber, University of Amsterdam, Area Editor of the ETAI Journal on the Semantic Web

- “I would certainly support the inclusion of a Workshop on this topic. Most timely.” Brian Garner, Deakin University

- “We want to give you a statement of interest of DaimlerChrysler Research Department: Data Mining Solution FT3/AD in a workshop ”Semantic Web Mining”. Jürgen Franke, DaimlerChrysler AG, Research Center Ulm

- “Indeed, if the success of the Semantic Web would be dependent on explicitly written semantic markup, the whole enterprise would be doomed. Therefore, techniques like web-mining and machine-learning will be crucial facilitators for the Semantic Web, and I warmly encourage the involvement of the ECML/PKDD community in this enterprise!” Frank van Harmelen, Vrije Universiteit Amsterdam, project co-manager of the EU On-To-Knowledge project

- “Our group has an interest in automating ontology population and generally in moving knowledge from informal to formal.” John Domingue, Open University, U.K.

- “Hereby CognIT A.S states its interest in the topic of the ”Semantic Web Mining” proposal for a workshop at the ECML PKDD. We feel that the topic is relevant and interesting for a larger audience stemming from industry as well as the academic world.” Rob Engels, CognIT a.s, Oslo

- “I am quite interested in the workshop considered, and it is quite likely that I or someone from my group would participate on it. We are working in the intersection of ontological engineering, web analysis, and data mining.” Vojtech Svatek, University of Economics, Prague

- “We are involved in the UK AKT project so this is very much of interest to us.” Derek Sleeman, University of Aberdeen

- “I shall be interested in this workshop because I am applying Machine Learning in 2 relevant areas: Web usage mining and learning image content for image retrieval.” Susan Craw, The Robert Gordon University, Aberdeen

- “At TNO we are actually working on the implementation of a semantic network based on the ideas I developed with Notion System [...] So you see we can’t hardly be more interested in the topics of this workshop.” Ronald Poell, Netherlands Organization for Applied Scientific Research (TNO)

We will publicize the workshop via mailing lists (e. g., mlnet, kdd-nuggets, KR, DL) and address lists of previously held ontology and KDD workshops. Especially we will invite all who sent us a statements of interest. Additionally, we will provide links from our institute home pages.

3 Organization of the Workshop

The workshop will consist of presentations followed by discussions. In order to obtain an intensive exchange of ideas between the different communities, there will be left extensive time for discussion following each presentation. A final panel will discuss the research agenda for the coming years.

If there are related workshops scheduled at ECML/PKDD, a joint session might be of interest.

Schedule

First call for papers	5 March 2001
Deadline for paper submission	8 June 2001
Notification of acceptance	29 June 2001
Deadline for final contributions	13 July 2001

4 Organizing Committee

Gerd Stumme
AIFB, Universität Karlsruhe
D–76128 Karlsruhe, Germany
email: `stumme@aifb.uni-karlsruhe.de`
`www.aifb.uni-karlsruhe.de/WBS/gst`
phone: +49-721-608 4754
fax: +49-721-693 717

Andreas Hotho
AIFB, Universität Karlsruhe
D–76128 Karlsruhe, Germany
email: `hotho@aifb.uni-karlsruhe.de`
`www.aifb.uni-karlsruhe.de/WBS/aho`
phone: +49-721-608 6558
fax: +49-721-693 717

Bettina Berendt
Abteilung Pädagogik und Informatik
Humboldt-Universität zu Berlin
Geschwister-Scholl-Straße 7
D-10117 Berlin
berendt@educat.hu-berlin.de
http://www.educat.hu-berlin.de/ berendt/index.html

5 Program Committee

- Rosine Cicchetti (Université de la Méditerranée, Marseille)
- Stefan Decker (Stanford University, Palo Alto)
- Klaus-Peter Huber (SAS, Heidelberg)
- Alexander Mädche (Universität Karlsruhe)
- Bamshad Mobasher (DePaul University, Chicago)
- Katharina Morik (Universität Dortmund)
- Claire Nedellec (Université Paris Sud)
- Myra Spiliopoulou (Otto–von–Guericke–Universität Magdeburg)
- Rudi Studer (Universität Karlsruhe)