

Semantic Web Mining for Building Information Portals

Jens Hartmann, Andreas Hotho, Gerd Stumme
Institute AIFB
University of Karlsruhe

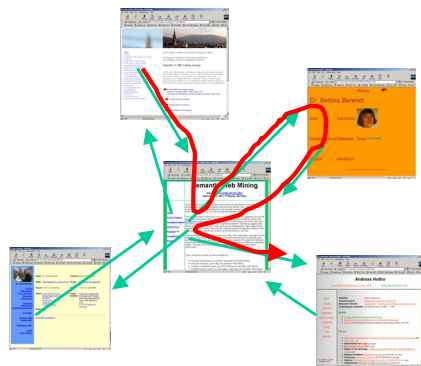


Semantic Web Mining for Building Information Portals

1. The Semantic Web and Web Mining
 - Web Mining
 - Semantic Web
2. Extracting Semantics from the Web
3. Mining the Semantic Web
4. Project SemiPort

Slide 2

What is Web Mining?



Goals:

- the **improvement of site design** and **site structure**,
- the generation of **dynamic recommendations**,
- and **improving marketing**.

Web Mining Areas

- Web content mining
- Web structure mining
- Web usage mining

Slide 3

Limitations of Web Mining ...

Structure/Content Mining

- Search and navigation by structure
- Content included in structure
- Content *only* understandable for humans
- Huge amount of data available

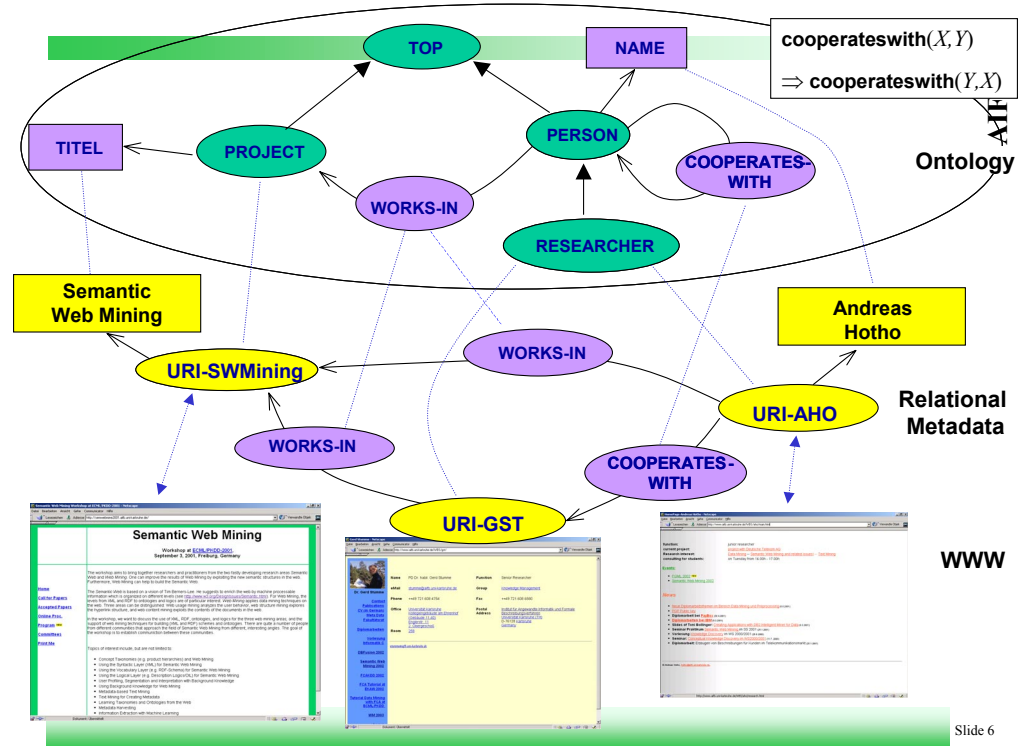
Usage Mining

- Usage tracking by technical access logs (url's)
- mapping of url's to general concepts is difficult and expensive
- events and url's are n to m

Slide 4

What is the Semantic Web?

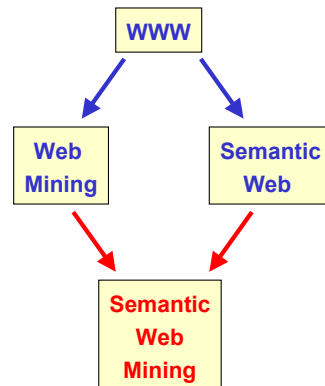
- As said before, The current WWW is a great success with respect to
 - the amount of available information
 - the number of users
- However, one problem of the current WWW is that the information may only be interpreted by humans.
- The **Semantic Web** tries to overcome this problem by using machine-processable **metadata** and **ontologies**.
- Goals** include:
 - more effective retrieval of Web information
 - knowledge-based inference



The Semantic Web and Web Mining

What are the benefits of bringing the Semantic Web and Web Mining together?

- The results of Web Mining can be improved by exploiting the new semantical structures in the Web.
- Web Mining can help to build the Semantic Web.



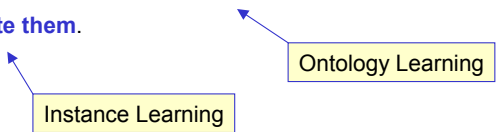
Semantic Web Mining for Building Information Portals



- The Semantic Web and Web Mining
- Extracting Semantics from the Web
 - Ontology Learning
 - Instance Learning
- Mining the Semantic Web
- Project SemlPort

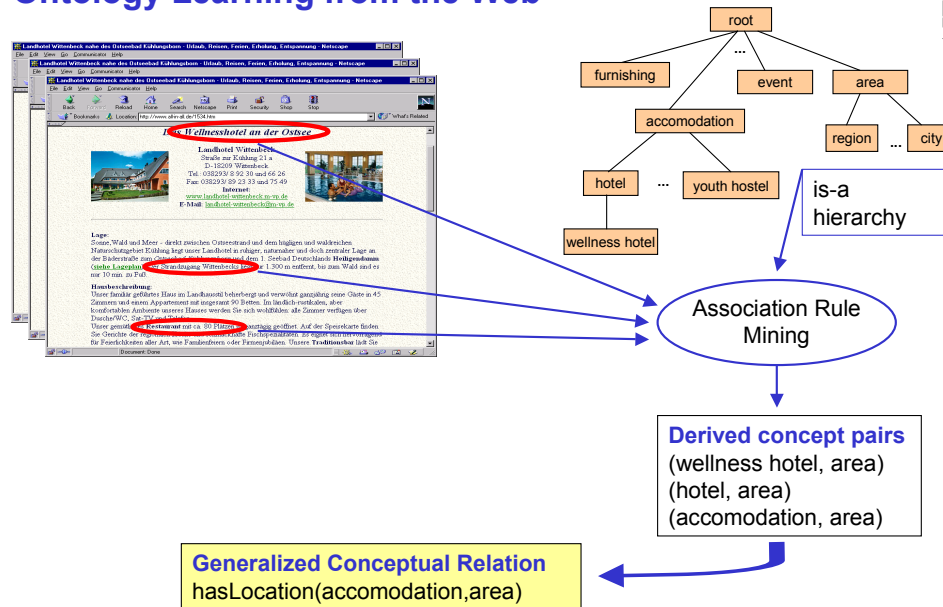
Extracting Semantics from the Web

- Web Mining can help
 - to learn structures for knowledge organization (e.g. ontologies)
 - and to populate them.

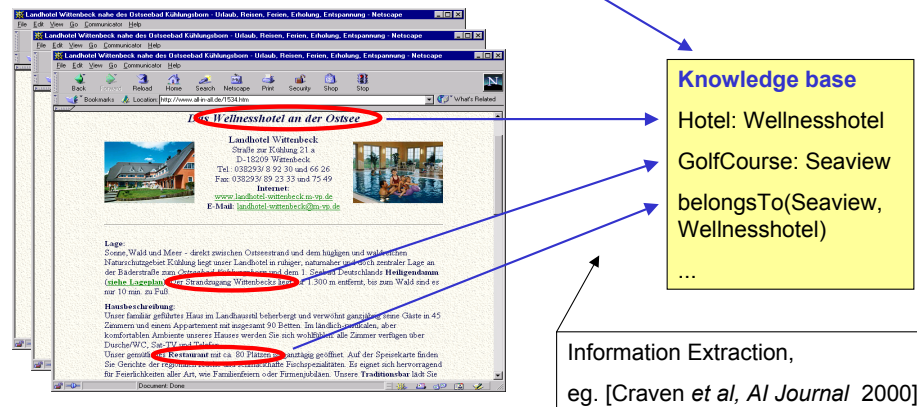
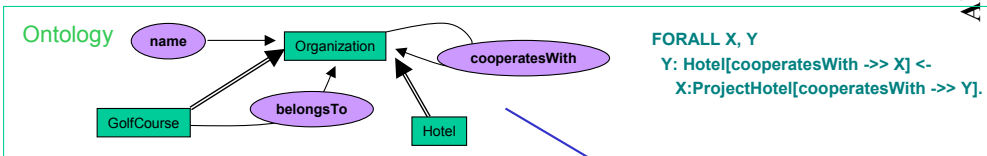


- The approaches are semi-automatic:
 - A lot of tacit background knowledge, experiences, social conventions, etc is involved in the modelling process.
 - In order to obtain high quality results, a human has to be in the loop.
 - If this were not the case, the Semantic Web would be superfluous!

Ontology Learning from the Web



Instance Learning from the Web

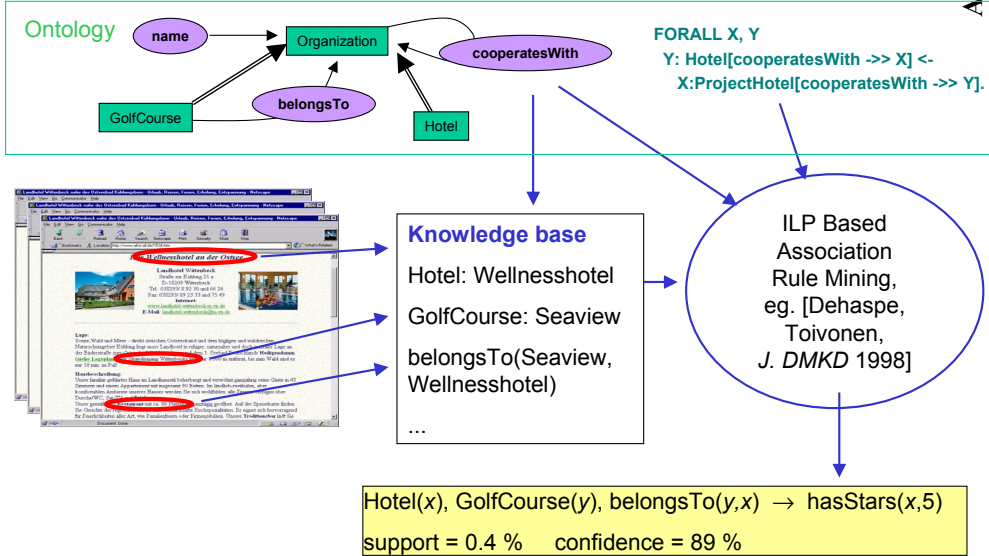


Semantic Web Mining for Building Information Portals



1. The Semantic Web and Web Mining
2. Extracting Semantics from the Web
3. Mining the Semantic Web
 - Structure/Content Mining
 - Usage Mining
4. Project SemlPort

Semantic Web Structure/Content Mining



Semantic Web Usage Mining

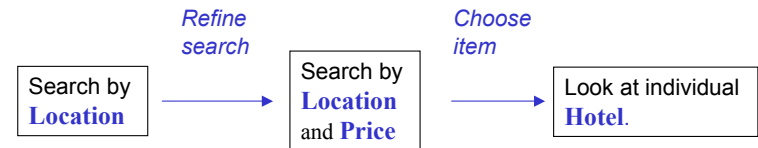
Basic idea: associate each requested page with one or more ontological entities, to better understand the process of navigation

From logfile analysis ...

```
p3ee24304.dip.t-dialin.net - - [19/Mar/2004:17:59:00] "GET /search.html?l=ostsee%20strand" 1759
p3ee24304.dip.t-dialin.net - - [19/Mar/2004:17:59:00] "GET /search.html?l=ostsee%20strand" 0" 200 8450
p3ee24304.dip.t-dialin.net - - [19/Mar/2004:17:59:00] "GET /mlesen.html?Item=3456&syn=02" 0
```

- Use the gained knowledge to
- understand search strategies
 - improve navigation design
 - personalization

... to semantic logfile analysis:



[Berendt & Spiliopoulou 2000; Berendt 2002; Oberle 2001]

Semantic Web Mining for Building Information Portals



- The Semantic Web and Web Mining
- Extracting Semantics from the Web
- Mining the Semantic Web
- Project SemiPort

AIFB Tasks in the SemiPort Project

Funded by bmb+f

- Ontology-Modelling and Metadata Standards
- Semantic Web Content & Structure Mining
- Semantic Web Usage Mining
- Ontology-Based Knowledge Integration

Objective

SemiPort objective is to develop innovative methods and tools for creating and maintaining semantic information portals for scientific communities.

Members

- AIFB Karlsruhe
- DFKI Saarbrücken
- IPSI Darmstadt
- DBIS Trier



Conclusion

- The results of Web Mining can be improved by exploiting the new semantical structures in the Web.
- Web Mining can help to build the Semantic Web.
- Both aspects can be combined by closing the loop.

- Semantic-based indexing and retrieval of *related* resources
- Providing *intelligent* recommendations
- Ontology-based access to large databases
- Knowledge Management for large distributed semantic databases