

Thematic session Proposal: Tools for *Service-oriented Discovery of Knowledge*

Axel Hahn (University of Oldenburg, Germany), Nada Lavrac (Jozef Stefan Institute, Ljubljana, Slovenia), Joost Kok (Leiden University, The Netherlands)

A major challenge for third generation data mining and knowledge discovery systems is the integration of different data/knowledge resources (which are highly diverse in nature in terms of representation and data formats) and computer systems (tools for data integration, data mining and knowledge discovery) which are distributed across the network. The emerging third generation data mining and knowledge discovery systems should be able to mine distributed and highly heterogeneous data found on intranets/extranets/grid, and integrate efficiently with operational data/knowledge management and data mining systems.

Compared to contemporary search engines which provide a means of locating data on the net, third generation data mining and knowledge discovery systems will provide a means for discovering patterns, associations, changes and anomalies in networked data, where each data source comes with its own structure, semantics, data formats, names, concepts and access methods. Currently, the burden falls on the user to manually (via programs) convert between the data formats, resolve conflicts, integrate data and interpret results in order to make viable use of this information.

This thematic session addresses novel approaches and key *tools* which will make third generation data mining and discovery of knowledge possible:

- Providing meta-data (semantic annotations) of different information resources (data, human-coded knowledge, and machine-induced patterns and predictive models) and data mining and knowledge discovery systems (pattern mining and model discovery tools);
- Implementation of data mining and knowledge discovery tools as *services* available on the web. Such service-oriented data mining and knowledge discovery systems will enable meta-level search of data/knowledge resources and systems, enabling the construction of knowledge discovery scenarios (workflows of potentially repeatable sequences of data mining and data integration steps), resulting in improved pattern and model discovery.

Topics:

- Frameworks for third generation data mining and knowledge discovery
- Inductive databases, Constraint-Based Data Mining and Inductive Queries
- Learning from data and knowledge (texts, ontologies, ...)
- Service-oriented approaches to data mining
- Meta-level annotations and search for data mining services
- Data mining workflows/scenarios
- Data mining on the grid
- Applications of service-oriented data mining approaches in business, ecological modeling, medicine, health care, e-science, bioinformatics, ...

Foreseen contributions: we will use as a basis papers from partners in a (future) European project and solicit additional papers from various researchers in the field.

