

Datamining tools and services for Grid computing environments

DataMiningGrid

Future and emerging complex problem-solving environments are characterised by increasing amounts of digital data and rising demands for coordinated resource sharing across geographically dispersed sites. Next generation Grid technologies are promising to provide the necessary infrastructure facilitating seamless sharing of computing resources in complex problem-solving environments. Currently there exists no coherent framework for developing and deploying datamining applications on the Grid. The DataMiningGrid project will address this gap by developing generic and sector-independent data-mining tools and services for the Grid.

The **main objectives** of the DataMiningGrid project are:

- to develop Grid interfaces that allow datamining tools and data sources to interoperate within distributed Grid computing environments. A user-friendly workflow editor will be provided to facilitate the configuration of analysis tasks;
- to develop Grid-based text mining and ontology-learning services and interfaces for knowledge discovery in texts and ontology learning;
- to develop a testbed consisting of several demonstrator applications from a diverse set of sectors, including the bioinformatics, healthcare, and automotive industries;
- to align and integrate these technologies with emerging Grid standards and infrastructures.

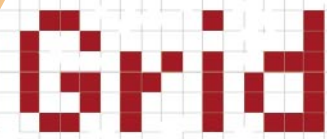
The **key technologies** developed in the project include:

- distributed datamining tools, facilitating novel approaches to mine data;
- Grid-aware datamining interfaces and services (e.g. data sources, algorithms);
- workflow-based management of mining in Grid environments;
- standardisation of datamining within Grid computing environments.

To demonstrate the technology developed, the project will implement a range of **demonstrator applications in e-science and e-business**. Other results include promotion of accessible data mining and the facilitation of industrial take-up.

Because of the increasing role of data in many sectors, the project's **impact** will be significant as it will be an important step towards more effective and efficient exploitation of available data and information resources. In the long run, the impact of the project will contribute to new business and R & D opportunities in the European market and an increase in quality of life. The project will also contribute to standardisation efforts of Grid and datamining technologies.

DataMining



Contract number
004475

Type of project
Specific targeted research project

Project coordinator
University of Ulster, School of
Biomedical Sciences

Contact person
Dr Werner Dubitzky
Cromore Road
BT52 ISA Coleraine
Northern Ireland
w.dubitzky@ulster.ac.uk

Project website
<http://www.datamininggrid.org>

**Maximum Community
contribution to project**
EUR1 883 000

Project start date
1 September 2004

Duration
36 months

continued overleaf ►



Information Society
Technologies



European Commission

Project partners*Organisation name and country*

UNIVERSITY OF ULSTER	UK
FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.	DE
DAIMLERCHRYSLER AG	DE
TECHNION – ISRAEL INSTITUTE OF TECHNOLOGY	IL
UNIVERZA V LJUBLJANI, FAKULTETA ZA GRADBENISTVO IN GEODEZIJO, D.O.O.	SI