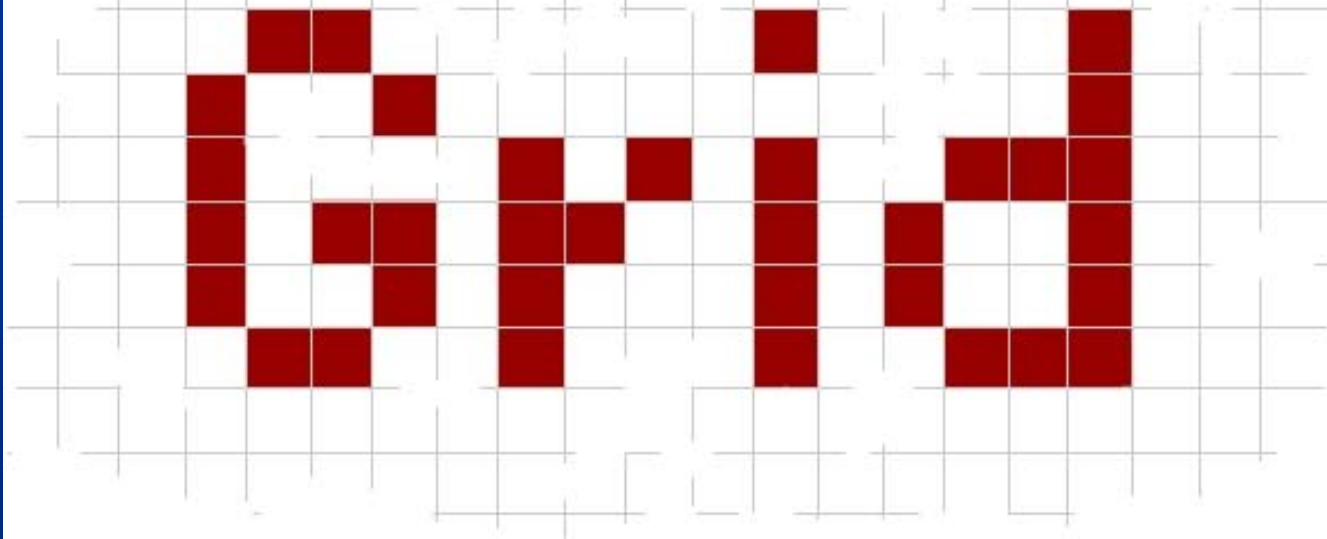


DataMining

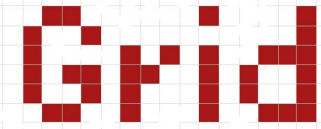


DataMiningGrid

Data Mining Tools and Services for Grid Computing Environments

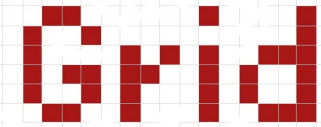
Context, Gap and Aim

- Context: Future and emerging complex problem-solving environments are characterized by increasing amounts of digital data and rising demands for co-ordinated resource sharing across geographically widely dispersed sites.
- Gap: Currently there exists no coherent framework for developing and deploying data-mining applications on the Grid.
- Aim: The DataMiningGrid project will address this gap by developing generic and sector-independent data mining interfaces and tools for the Grid.



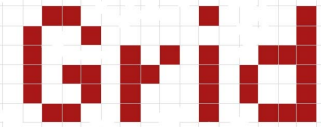
Main Objectives

- Grid interfaces/tools facilitating
 - Grid interoperation of data mining tools and data sources
 - Knowledge discovery in text and ontology-learning in Grid environments
 - Analysis of network and Grid failures
- User-friendly workflow editor for defining and configuration data analysis tasks in Grid environments
- Test-bed comprising several demonstrator applications from bioinformatics, health care, and automotive industry
- Alignment and integration of these technologies with emerging Grid standards and infrastructures



Expected Results and Benefits

- **Interfaces/tools** facilitating data mining applications in Grid computing environments to be used by **developer/system integrators**
 - (1) Interoperation data mining/sources; (2) Text mining/ontology learning; (3) Network/Grid failures; and (4) Workflow editor
- Test-bed: **demonstrators/use cases** for use by **developer/system integrators and end users**
 - (1) Text mining for quality management in automotive industry; (2) Mining of distributed medical/biological databases; (3) Modelling of genetic networks; (4) Mining of historic text documents and digital libraries; and (5) Analysis of network and Grid failures
- Main users:
 - (1) Developers of data mining and Grid applications; (2) System integrators; and (3) End users
- Contribution to related Grid and data mining standards such as web services, Grid standards, data/text mining, data warehousing

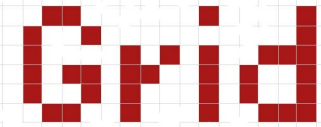


Exploitation Potential

- Strong industrial exploitation focus (DC, FhG, Technion)
- Identified market opportunities at three levels
 1. **Enhancements of Grid software** – available as open source
 2. **Data mining interfaces/tools** addressing common data mining tasks – developers, integrators, software companies)
 3. **Data mining demonstrators** addressing end user applications – end users, solution providers, developers, integrators, software companies, etc.
- Positioning of products and services through cooperation, industrial projects and possibly joint ventures
- Inimitable competitive advantage: combination of innovation and project-related expert know-how (Grid and data mining)

Partners

DataMining



DAIMLERCHRYSLER



Isreal Institute of Technology



University of Ljubljana