

1st Workshop on Software Services: Frameworks and Platforms  
Timisoara, September 23-25, 2010



## ADVANCED DATA MINING AND INTEGRATION RESEARCH FOR EUROPE (ADMIRE)

PETER KRAMMER  
SLOVAK ACADEMY OF SCIENCE

**ABSTRACT.** The project ADMIRE aims to develop an easy-to-use, high-level, abstract data mining and data integration process management. To this end, we have designed an architecture, which comprises a high-level, abstract data integration and data mining (DMI) description language called DISPEL, a set of tools to support the language, and an underlying infrastructure enabling integration and mining of large, distributed data sets. A DMI process is described in DISPEL and executed through a Gateway, which parses the DISPEL document and initiates data streaming using OGSA-DAI. The data is then processed in a set of filters, corrected, cleaned, and finally fed into a data mining algorithm which provides the user with the required results. The whole process is managed through a visual interface, where the user can compose his/her process, create a DISPEL document, submit the document for processing, and visualize results. The interface enables cooperation of domain experts, data mining experts, and regular users. The architecture, language and tools of ADMIRE are verified using two pilot applications. One of the applications, being developed by the Institute of Informatics of the Slovak Academy of Sciences tries to predict various hydro-meteorological phenomena by using distributed data integration and data mining. The application consists of a set of scenarios, specified by domain experts. The scenarios use data from several different vendors, and use data mining algorithms to provide results which are currently not easily achievable with standard methods used in the Earth Sciences community.