Semantic Methods and Tools for Information Portals The SemIPort Project

Jorge Gonzalez-Olalla, Gerd Stumme

Institute of Applied Informatics and Formal Description Methods AIFB,
University of Karlsruhe, D–76128 Karlsruhe, Germany
http://www.aifb.uni-karlsruhe.de/WBS
{jgo, stumme}@aifb.uni-karlsruhe.de

1 Objectives and Research topics.

The goal of the SemIPort project is to develop innovative methods for presenting, storing and accessing scientific information within a Semantic Information Portal. The project will enable users to access a semantically structured information inventory, while allowing them to integrate their own information.

Several scenarios as the AIFB portal and the DBLP online bibliography will be used to test the developed methods. Later, the results of the project will be implemented in the form of an Internet Portal for the German Informatics Society (GI) as a competency and service network .

The project is financed by the German bmb+f for a period of 3 years starting June 2002. The participants are the Institute for Applied Informatics and Formal Description Methods at the University of Karlsruhe (coordinator), the German Research Center for Artificial Intelligence, the Fraunhofer Institute for Integrated Publication and Information Systems, and the Data Bases and Information Systems Group at the University of Trier.

The **research tasks** within the project are:

(1) Ontology Modelling and Metadata Standards. (2) Scalable Storing, Processing and Querying of Integrated XML and RDF Inventories (Knowledge Warehouse). (3) Web Mining and Ontology-based Knowledge Integration. (4) Visualization and Browsing of Complex Data Inventories. (5) Personalization and Agent-based Interaction.

The most related to the workshop is task (3). This topic will concentrate on:

i) Semantic Web Content and Structure Mining. The explicit semantics of semantic web data will be used to improve the results of 'classical' web and Data Mining techniques. E.g. for 'hubs and authorities' one could distinguish between different kinds of authorities.

ii) Semantic Web Usage Mining. The behavior of the portal's visitors will be used to optimize the portal's structure and usability, by developing tools that will help to manage and update the ontology according to the actual needs of the users. A 'semantic log file' will be created and analyzed for this purpose.

iii) Flexible techniques for knowledge integration will be developed. Existing methods of schema integration for relational data bases and ontologies will be analyzed and expanded for dynamic integration.