INFORMATION RESOURCES INTEGRATION OF WORLD DATA SYSTEM

Sergii Telenyk[1], Kostiantyn Yefremov[2]

[1] National Technical University of Ukraine “Kyiv Polytechnic Institute”, Faculty of Informatics and Computer Science, 37 Peremogy ave, Kyiv, email: telenik@acts.kiev.ua
[2] National Technical University of Ukraine “Kyiv Polytechnic Institute”, ICSU World Data Center for Geoinformatics and Sustainable Development, 37 Peremogy ave, Kyiv, Email: k.yefremov@wdc.org.ua

An existing Data Centers of World Data Systems provide users with access to accumulate over a long time period of non-uniform data and a variety of data analysis and processing facilities. However, users experience difficulties when trying to obtain the necessary data, if they are spread across multiple sources and require processing with multiple data analysis and processing facilities. The one effective solution to this problem would be creation of Semantic Data Directory (SDD), which will provide access to remote data sources on a single-window principle.

Multi-level architecture of user information service system, which provides on-demand access to mixed data, their transformation and processing, and presenting them in convenient form is proposed.

The user needs only to know necessary data in terms of SDD and be able to query in a pre-known format of SPARQL language. Further, user-friendly interface will help him/her to specify the query without knowing the syntax of SPARQL.

Proposed approaches were used for development and implementation of pilot version of SDD for Russian-Ukrainian WDS Segment.

Proposed architecture is based on three basic concepts. First, the semantic content of data sources and data analysis and processing facilities should be described in a formalized manner in global namespace according to Semantic Web concepts. The information unit is presented as directed graph, which opens up wide possibilities for data manipulating. For example, merging of two ontologies is reduced to the union of their graphs. Access to data and processing scheme are defined on the basis of this description.

Second, access to data is based on the integration of data sources concept. The specialized “Mediator” component determines data source to which request is addressed, if necessary - makes sub-queries to multiple data sources with the ability to combine the results. It creates the “Wrappers” to provide uniform access interface to mixed data sources customized for each source. “Wrappers” broadcast each request of “Mediator” so that understood for data source (this can be database or any other data), and transmit result back to “Mediator”.

The third component of the approach is a system of services integration. Whatever means of services implementation, it can be constructed series-parallel function call scheme with the appropriate input and output parameters based on the user's query. For the selection of required functions and services is used mentioned above description of services, part of which is a system of preconditions and postconditions imposed on parameters of services and their functions. The formal base of function call scheme is inference logic engine. During execution of scheme it is possible to connect services, their components and functions independently from the technology of implementation. To do this, meta-description of services and their components should be prepared at semantic level. For service interactions can be applied technologies such as JMS, FIPA ACL, or WCF.

The proposed architecture requires no changes to data sources, storage structures, and maintenance mechanisms. It can function away autonomously within the existing systems, supplemented and modified by the owner.

Keywords: **Semantic Data Directory,**