



Information-computational system for storage, search and analytical processing of environmental datasets based on the Semantic Web technologies

A. Titov, E. Gordov, and I. Okladnikov

Institute of Monitoring of Climatic and Ecological Systems SB RAS, Siberian center for Environmental Research and Training, Tomsk, Russian Federation (office@scert.ru)

In this report the results of the work devoted to the development of working model of the software system for storage, semantically-enabled search and retrieval along with processing and visualization of environmental datasets containing results of meteorological and air pollution observations and mathematical climate modeling are presented.

Specially designed metadata standard for machine-readable description of datasets related to meteorology, climate and atmospheric pollution transport domains is introduced as one of the key system components. To provide semantic interoperability the Resource Description Framework (RDF, <http://www.w3.org/RDF/>) technology means have been chosen for metadata description model realization in the form of RDF Schema. The final version of the RDF Schema is implemented on the base of widely used standards, such as Dublin Core Metadata Element Set (<http://dublincore.org/>), Directory Interchange Format (DIF, <http://gcmd.gsfc.nasa.gov/User/difguide/difman.html>), ISO 19139, etc.

At present the system is available as a Web server (<http://climate.risks.scert.ru/metadatabase/>) based on the web-portal ATMOS engine [1] and is implementing dataset management functionality including SeRQL-based semantic search as well as statistical analysis and visualization of selected data archives [2,3]. The core of the system is Apache web server in conjunction with Tomcat Java Servlet Container (<http://jakarta.apache.org/tomcat/>) and Sesame Server (<http://www.openrdf.org/>) used as a database for RDF and RDF Schema.

At present statistical analysis of meteorological and climatic data with subsequent visualization of results is implemented for such datasets as NCEP/NCAR Reanalysis, Reanalysis NCEP/DOE AMIP II, JMA/CRIEPI JRA-25, ECMWF ERA-40 and local measurements obtained from meteorological stations on the territory of Russia. This functionality is aimed primarily at finding of main characteristics of regional climate dynamics. The proposed system represents a step in the process of development of a distributed collaborative information-computational environment to support multidisciplinary investigations of Earth regional environment [4].

Partial support of this work by SB RAS Integration Project 34, SB RAS Basic Program Project 4.5.2.2, APN Project CBA2007-08NSY and FP6 Enviro-RISKS project (INCO-CT-2004-013427) is acknowledged.

References

1. E.P. Gordov, V.N. Lykosov, and A.Z. Fazliev. Web portal on environmental sciences "ATMOS" // *Advances in Geosciences*. 2006. Vol. 8. p. 33 – 38.
2. Gordov E.P., Okladnikov I.G., Titov A.G. Development of elements of web based information-computational system supporting regional environment processes investigations // *Journal of Computational Technologies*, Vol. 12, Special Issue #3, 2007, pp. 20 – 28.
3. Okladnikov I.G., Titov A.G. Melnikova V.N., Shulgina T.M. Web-system for processing and visualization of meteorological and climatic data // *Journal of Computational Technologies*, Vol. 13, Special Issue #3, 2008, pp. 64 – 69.
4. Gordov E.P., Lykosov V.N. Development of information-computational infrastructure for integrated study of Siberia environment // *Journal of Computational Technologies*, Vol. 12, Special Issue #2, 2007, pp. 19 – 30.