Creation of unified parameter vocabulary services using Semantic Web technologies to ensure data integration and interoperability of information systems

Introduction

Unified parameter vocabulary allows to uniformly encode environmental parameters in the formation, transmission, and processing of observational data on the marine environment. It ensures the unification and interoperability of data during their further use. Roshydromet is using unified parameter vocabulary for more than 10 years and it is used in many existing projects and systems.

Problem

Existing approach is using XML to exchange the unified parameter vocabulary for further use in projects and systems. This approach has number of limitations and restrictions (specific metadata model is used, need for parsers on consumer side, delays in transmission processes, etc.). More promising is the use of the Semantic Web to improve the access and search capabilities on the top of existing vocabulary data model. Semantic Web will allow to take into account the semantics of the domain of the parameter vocabulary, organize a general structure for describing the data, create a model of relationships between individual parameters, parameter groups, phenomena and processes to assess their completeness and quality.

Results

As a result, the ontological model of unified parameter vocabulary, model update services, SPARQL-endpoint and REST-services for access to unified parameter vocabulary there were developed . These tools allowed to solve the problems of limited access to a parameter vocabulary (see Figure 1).
Approaches to interoperability at the level of the contents of the vocabulary with BODC NVS vocabulary are realized.

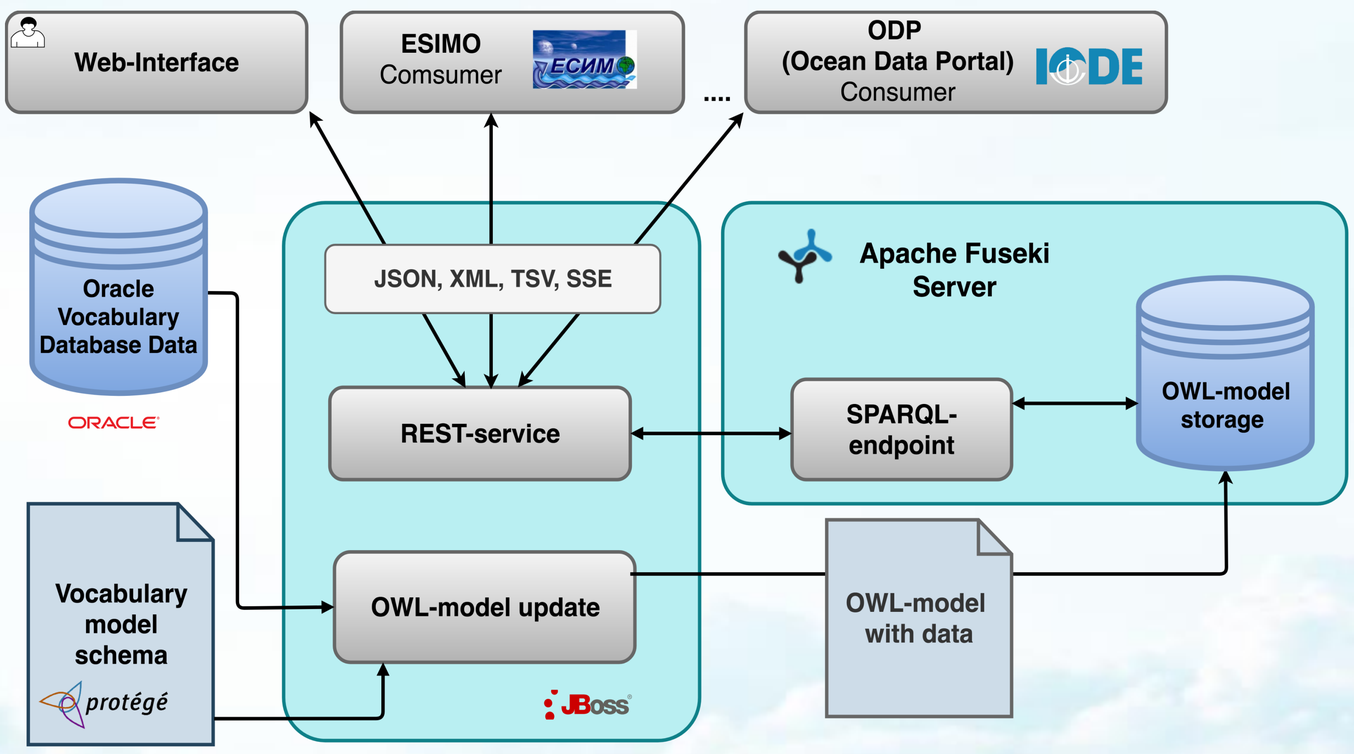

Figure 1 – Architecture of vocabulary parameter services using Semantic Web technologies

4. Storage of the ontological model and the content access service.

The ontological model is uploaded to the storage on the Apache Fuseki server, which also hosts the SPARQL-endpoint, a service that transmits requests to the storage and extracts the necessary information.

5. REST-service for accessing information from the ontological model.

Used to transfer information from the SPARQL service to the consumers. The output formats are: json, tsv, xml, sse. Similar approaches are implemented in international vocabularies, such as BODC NVS and SYSSVoc.

6. Web-interface to demonstrate access to the REST-service.

It helps to demonstrate various types of search, perform sorting/ filtering of results, export results to EXCEL, provide the ability to conveniently view the necessary information.

Implementation

1. Analysis of the structure and composition of the unified parameter vocabulary.

The structure of the vocabulary allows to identify entities, their attributes and relationships between them.

2. Creating an ontological model of the unified parameter vocabulary.

In this step OWL technology was used. All entities of the vocabulary are described as “classes” of the ontological model, and the relationships between them are described as “properties” of classes. The basic Oceanographic parameters of Roshydromet were mapped to the BODC parameter vocabulary using URN identifiers created on the BODC NVS side.

3. Software implementation of the complete ontological model of the unified parameter vocabulary.

The Apache Jena Java framework was used in the implementation of the OWL model. Each vocabulary entity in the model is described by a unique URI-identifier. The model is updated daily.

Perspectives

Further use of Semantic Web technologies will be introduced by RIHMI-WDC in national systems and projects, such as ODP, and will provide unified formalized platform for integrating the unified parameter vocabulary with world vocabularies, such as SISVoc and BODC NVS.

CONTACT US:
Anna Maslennikova (amaslennikova@meteo.ru)
Sergey Belov (belov@meteo.ru)

Russian National Oceanographic Data Center (NODC)
All-Russian Research Institute of Hydrometeorological Information – World Data Center (RIHMI-WDC)
6, Korolev St., Obninsk, Kaluga Region, 249035 Russian Federation
Phone: +7 (48439)74679 / Fax: +7 (48439)68611

Partnership Centre for the IODE Ocean Data Portal