The integration platform of Roshydromet for data exchange within Russian and international projects

Lazareva Alina, RIHMI-WDC (Russia), alazareva@meteo.ru
Belov Sergey, RIHMI-WDC (Russia), belov@meteo.ru

In modern world every day we receive and process great amount of hydrometeorological information, obtained from different parts of the world for ensuring activity. Ocean data is used in different spheres of human activities. Many organizations handle collection, storage, accumulation, processing, and dissemination of information of global scale. Projects that are interested in these processes have similar but at the same time various options of registration and formalization of data and metadata of so called “information resources”. It makes difficulties in the migration of information resources between systems. Information exchange can be done using local systems "adapters", providing transformation from one particular data format to another, but this is economically unprofitable, resource-intensive and simply inexpedient. For this reason mutual exchange of data between systems is almost not possible.

Creation of a platform for integration information resources can complement in solving interoperability problem between systems. There are several major problems of integration that such platform should solve: unification of structural presentation of information resources, actuality of information, monitoring of the data transmission as well access rights to information resources and services.

Such integration platform is an intermediate element in the process of data transferring between data source to destination system – end users, acting as some sort of “interpreter” between two interacting systems, originally speaking different “languages”. Integration of heterogeneous information resources take place by using definite data and metadata formalization techniques.

All interaction in the Roshydromet data integration platform is based on the concept of “information resource”. This concept describes metadata about the data as well as data source (database, file catalogues, web services, etc.). Comprehensive metadata descriptions helps to identify the location of the data source, access and retrieve data from the local data sources, transform to appropriate transport data format (point, profile or grid using netCDF). Roshydromet data integration platform has strict rule that metadata should be always accompanied by data. No incorporeal metadata descriptions allowed. Metadata is based on ISO-19115 standard, controlled code lists and Roshydromet unified parameter vocabulary and has transformation procedures to be exported as common ISO 19139 or WMO Core Profile metadata standards. Information resource can describe a complete data set (e.g. historical oceanographic cruises as a whole data store) or subsets (e.g. cruises in Arctic, Atlantic ocean, Pacific ocean, etc.) with multiple conditions, for instance, combined by vessel types, temporal extents, measurements, instruments, etc. Based on such rules information resource is producing single or multiple instance descriptions that contains metadata for individual subset. Instance descriptions are derived from “parents” metadata records. Using metadata, data originators assign permissions - no conditions for data access and use or restricted access based on agreement model.

Integration platform contains a number of software components that provides web interfaces to administrators and operators in the system. It utilizes intelligent algorithms and set of rules to simplify metadata creation, update in automatic or semi-automatic mode based on data update frequency. Therefore metadata in Roshydromet integration platform is always adequate to data in term of spatial and temporal coverage.

At present moment developed integration platform directly or indirectly is used in national and international projects. Developed under aegis of ESIMO project (www.esimo.ru) it has been transferred to IODE OceanDataProject as in-kind contribution. Further it has been used for creation of Russian segment of WMO Information System, Informational and telecommunication network of Roshydromet for collecting

geophysical data, state fund system of RIHMI-WDC. As a success story, ESIMO now has 36 institutions from 12 ministries and agencies connecting over 500 databases and data storages that producing 3300 information resources and around 30000 instances respectively. This integration platform is in new evolution phase, but even being a promising development in Roshydromet, it is not limited to hydrometeorology and can be used in other fields of human activity where there is a need to integrate various heterogeneous information.