MyOcean Internal Information System (Dial-P)

Frederique Blanc (1), Tony Jolibois (1), Thomas Loubrieu (2), Giuseppe Manzella (3), Paolo Mazzetti (4), and Stefano Nativi (4)

(1) CLS, Toulose, France, fblanc@cls.fr, (2) Ifremer, Brest, France, Thomas.Loubrieu@ifremer.fr, (3) ENEA, La Spezia, Italy, giuseppe.manzella@enea.it, (4) IMAA CNR, Prato, Italy, mazzetti@imaa.cnr.it

MyOcean is a three-year project (2008-2011) which goal is the development and pre-operational validation of the GMES Marine Core Service for ocean monitoring and forecasting. It's a transition project that will conduct the European "operational oceanography" community towards the operational phase of a GMES European service, which demands more European integration, more operationality, and more service. Observations, model-based data, and added-value products will be generated – and enhanced thanks to dedicated expertise – by the following production units:
• Five Thematic Assembly Centers, each of them dealing with a specific set of observation data: Sea Level, Ocean colour, Sea Surface Temperature, Sea Ice & Wind, and In Situ data,
• Seven Monitoring and Forecasting Centers to serve the Global Ocean, the Arctic area, the Baltic Sea, the Atlantic North-West shelves area, the Atlantic Iberian-Biscay-Ireland area, the Mediterranean Sea and the Black sea.

Intermediate and final users will discover, view and get the products by means of a central web desk, a central re-active manned service desk and thematic experts distributed across Europe. The MyOcean Information System (MIS) is considering the various aspects of an interoperable – federated information system. Data models support data and computer systems by providing the definition and format of data. The possibility of including the information in the data file is depending on data model adopted. In general there is little effort in the actual project to develop a 'generic' data model. A strong push to develop a common model is provided by the EU Directive INSPIRE. At present, there is no single de-facto data format for storing observational data. Data formats are still evolving, with their underlying data models moving towards the concept of Feature Types based on ISO/TC211 standards. For example, Unidata are developing the Common Data Model that can represent scientific data types such as point, trajectory, station, grid, etc., which will be implemented in netCDF format. SeaDataNet is recommending ODV and NetCDF formats. Another problem related to data curation and interoperability is the possibility to use common vocabularies. Common vocabularies are developed in many international initiatives, such as GEMET (promoted by INSPIRE as a multilingual thesaurus), UNIDATA, SeaDataNet, Marine Metadata Initiative (MMI). MIS is considering the SeaDataNet vocabulary as a base for interoperability. Four layers of different abstraction levels of interoperability an be defined:
[U+F02D] Technical/basic: this layer is implemented at each TAC or MFC through internet connection and basic services for data transfer and browsing (e.g FTP, HTTP, etc).
[U+F02D] Syntactic: allowing the interchange of metadata and protocol elements. This layer corresponds to a definition Core Metadata Set, the format of exchange/delivery for the data and associated metadata and possible software. This layer is implemented by the DIAL-P logical interface (e.g. adoption of INSPIRE compliant metadata set and common data formats).
[U+F02D] Functional/pragmatic: based on a common set of functional primitives or on a common set of service definitions. This layer refers to the definition of services based on Web services standards. This layer is implemented by the DIAL-P logical interface (e.g. adoption of INSPIRE compliant network services).
[U+F02D] Semantic: allowing to access similar classes of objects and services across multiple sites, with multilinguality of content as one specific aspect. This layer corresponds to MIS interface, terminology and thesaurus.
Given the above requirements, the proposed solution is a federation of systems, where the individual participants are self-contained autonomous systems, but together form a consistent wider picture. A mid-tier integration layer

mediates between existing systems, adapting their data and service model schema to the MIS. The developed MIS is a read-only system, i.e. does not allow updating (or inserting) data into the participant resource systems. The main advantages of the proposed approach are:

• to enable information sources to join the MIS and publish their data and metadata in a secure way, without any modification to their existing resources and procedures and without any restriction to their autonomy;

• to enable users to browse and query the MIS, receiving an aggregated result incorporating relevant data and metadata from across different sources;

• to accommodate the growth of such a MIS, either in terms of its clients or of its information resources, as well as the evolution of the underlying data model.