Geophysical Research Abstracts,
Vol. 11, EGU2009-1487, 2009
EGU General Assembly 2009

SeaDataNet - Pan-European infrastructure for marine and ocean data management: Unified access to distributed data sets

D.M.A. Schaap (1) and G. Maudire (2)

(1) Mariene Informatie Service 'MARIS', Voorburg, The Netherlands, (dick@maris.nl), (2) IFREMER, Brest, FRance

SeaDataNet is an Integrated research Infrastructure Initiative (I3) in EU FP6 (2006 – 2011) to provide the data management system adapted both to the fragmented observation system and the users need for an integrated access to data, meta-data, products and services. Therefore SeaDataNet insures the long term archiving of the large number of multidisciplinary data (i.e. temperature, salinity current, sea level, chemical, physical and biological properties) collected by many different sensors installed on board of research vessels, satellite and the various platforms of the marine observing system. The SeaDataNet project started in 2006, but builds upon earlier data management infrastructure projects, undertaken over a period of 20 years by an expanding network of oceanographic data centres from the countries around all European seas. Its predecessor project Sea-Search had a strict focus on metadata. SeaDataNet maintains significant interest in the further development of the metadata infrastructure, but its primary objective is the provision of easy data access and generic data products.

SeaDataNet is a distributed infrastructure that provides transnational access to marine data, meta-data, products and services through 40 interconnected Trans National Data Access Platforms (TAP) from 35 countries around the Black Sea, Mediterranean, North East Atlantic, North Sea, Baltic and Arctic regions. These include:

- National Oceanographic Data Centres (NODC's)
- Satellite Data Centres.

Furthermore the SeaDataNet consortium comprises a number of expert modelling centres, SME's experts in IT, and 3 international bodies (ICES, IOC and JRC).

Planning:

The SeaDataNet project is delivering and operating the infrastructure in 3 versions:

- Version 0: maintenance and further development of the metadata systems developed by the Sea-Search project plus the development of a new metadata system for indexing and accessing to individual data objects managed by the SeaDataNet data centres. This is known as the Common Data Index (CDI) V0 system
- Version 1: harmonisation and upgrading of the metadatabases through adoption of the ISO 19115 metadata standard and provision of transparent data access and download services from all partner data centres through upgrading the Common Data Index and deployment of a data object delivery service.
- Version 2: adding data product services and OGC compliant viewing services and further virtualisation of data access.

SeaDataNet Version 0:

The SeaDataNet portal has been set up at http://www.seadatanet.org and it provides a platform for all SeaDataNet services and standards as well as background information about the project and its partners. It includes discovery services via the following catalogues:

- CSR - Cruise Summary Reports of research vessels;
- EDIOS – Locations and details of monitoring stations and networks / programmes;
- EDMED – High level inventory of Marine Environmental Data sets collected and managed by research institutes and organisations;
- EDMERP - Marine Environmental Research Projects ;
- EDMO - Marine Organisations.

These catalogues are interrelated, where possible, to facilitate cross searching and context searching. These catalogues connect to the Common Data Index (CDI).

- Common Data Index (CDI)

The CDI gives detailed insight in available datasets at partners databases and paves the way to direct online data access or direct online requests for data access / data delivery. The CDI V0 metadatabase contains more than 340.000 individual data entries from 36 CDI partners from 29 countries across Europe, covering a broad scope and range of data, held by these organisations. For purposes of standardisation and international exchange the ISO19115 metadata standard has been adopted. The CDI format is defined as a dedicated subset of this standard. A CDI XML format supports the exchange between CDI-partners and the central CDI manager, and ensures interoperability with other systems and networks. CDI XML entries are generated by participating data centres, directly from their databases. CDI-partners can make use of dedicated SeaDataNet Tools to generate CDI XML files automatically.

Approach for SeaDataNet V1 and V2:

The approach for SeaDataNet V1 and V2, which is in line with the INSPIRE Directive, comprises the following services:

- Discovery services = Metadata directories
- Security services = Authentication, Authorization & Accounting (AAA)
- Delivery services = Data access & downloading of datasets
- Viewing services = Visualisation of metadata, data and data products
- Product services = Generic and standard products
- Monitoring services = Statistics on usage and performance of the system
- Maintenance services = Updating of metadata by SeaDataNet partners

The services will be operated over a distributed network of interconnected Data Centres accessed through a central Portal. In addition to service access the portal will provide information on data management standards, tools and protocols.

The architecture has been designed to provide a coherent system based on V1 services, whilst leaving the pathway open for later extension with V2 services. For the implementation, a range of technical components have been defined. Some are already operational with the remainder in the final stages of development and testing. These make use of recent web technologies, and also comprise Java components, to provide multi-platform support and syntactic interoperability. To facilitate sharing of resources and interoperability, SeaDataNet has adopted SOAP Web Service technology. The SeaDataNet architecture and components have been designed to handle all kinds of oceanographic and marine environmental data including both in-situ measurements and remote sensing observations.

The V1 technical development is ready and the V1 system is now being implemented and adopted by all participating data centres in SeaDataNet.

Interoperability:

Interoperability is the key to distributed data management system success and it is achieved in SeaDataNet V1 by:

- Using common quality control protocols and flag scale
- Using controlled vocabularies from a single source that have been developed using international content governance
- Adopting the ISO 19115 metadata standard for all metadata directories
- Providing XML Validation Services to quality control the metadata maintenance, including field content verification based on Schematron.
- Providing standard metadata entry tools
- Using harmonised Data Transport Formats (NetCDF, ODV ASCII and MedAtlas ASCII) for data sets delivery
- Adopting of OGC standards for mapping and viewing services
- Using SOAP Web Services in the SeaDataNet architecture

SeaDataNet V1 Delivery Services:

An important objective of the V1 system is to provide transparent access to the distributed data sets via a unique user interface at the SeaDataNet portal and download service. In the SeaDataNet V1 architecture the Common Data Index (CDI) V1 provides the link between discovery and delivery. The CDI user interface enables users to have a detailed insight of the availability and geographical distribution of marine data, archived at the connected data centres, and it provides the means for downloading data sets in common formats via a transaction mechanism.

The SeaDataNet portal provides registered users access to these distributed data sets via the CDI V1 Directory and a shopping basket mechanism. This allows registered users to locate data of interest and submit their data requests. The requests are forwarded automatically from the portal to the relevant SeaDataNet data centres. This process is controlled via the Request Status Manager (RSM) Web Service at the portal and a Download Manager (DM) java software module, implemented at each of the data centres. The RSM also enables registered users to check regularly the status of their requests and download data sets, after access has been granted. Data centres can follow all transactions for their data sets online and can handle requests which require their consent. The actual delivery of data sets is done between the user and the selected data centre.

The CDI V1 system is now being populated by all participating data centres in SeaDataNet, thereby phasing out CDI V0.

0.1 SeaDataNet Partners:

IFREMER (France), MARIS (Netherlands), HCMR/HNODC (Greece), ULg (Belgium), OGS (Italy), NERC/BODC (UK), BSH/DOD (Germany), SMHI (Sweden), IEO (Spain), RIHMI/WDC (Russia), IOC (International), ENEA (Italy), INGV (Italy), METU (Turkey), CLS (France), AWI (Germany), IMR (Norway), NERI (Denmark), ICES (International), EC-DG JRC (International), MI (Ireland), IHPT (Portugal), RIKZ (Netherlands), RBINS/MUMM (Belgium), VLIZ (Belgium), MRI (Iceland), FIMR (Finland), IMGW (Poland), MSI (Estonia), IAE/UL (Latvia), CMR (Lithuania), SIO/RAS (Russia), MHI/DMIST (Ukraine), IO/BAS (Bulgaria), NIMRD (Romania), TSU (Georgia), INRH (Morocco), IOF (Croatia), PUT (Albania), NIB (Slovenia), UoM (Malta), OC/UCY (Cyprus), IOLR (Israel), NCSR/NCMS (Lebanon), CNR-ISAC (Italy), ISMAL (Algeria), INSTM (Tunisia)