

From SeaDataNet II to SeaDataCloud

Michèle Fichaut, IFREMER (France), michele.fichaut@ifremer.fr

Dick Schaap, MARIS (Netherlands), dick@maris.nl

On behalf of SeaDataNet and SeaDataCloud consortium

Scientists, environmental policy makers, and industry need accurate, homogeneous, integrated marine data sets and reliable services to access these. Oceanographic and marine data include a very wide range of measurements and variables, covering a broad, multidisciplinary spectrum of projects and programmes. They are collected by over a thousand research institutes, governmental organisations and private companies in the countries bordering the European seas. Various heterogeneous observing sensors are installed on research vessels, submarines, aircraft, and moorings, drifting buoys, gliders, floats, fixed platforms and satellites. These sensors measure physical, chemical, biological, geological and geophysical parameters, with further data resulting from the analysis of water and sediment samples for a wide variety of parameters. These data are collected at a very considerable cost. Besides, they are of prime value because they are the reference for any study and, if lost, cannot be replaced. In order to make best use of these data for science and for society, a robust operational infrastructure, based on European and internationally agreed standards, is mandatory. It has to cover data quality and long term stewardship as well as technical and semantic aspects of interoperability.

The SeaDataNet pan-European infrastructure has been developed by NODCs and major research institutes from 34 countries, in the frame of two European projects (FP6/SeaDataNet – 2006-2011 and FP7/SeaDataNet II – 2011-2015). Over 100 marine data centres are connected and provide discovery and access to data resources for all European researchers. SeaDataNet is a key infrastructure driving several portals of the European Marine Observation and Data network (EMODnet); it complements the Copernicus Marine Environmental Monitoring Service (CMEMS). However, more effective and convenient access is needed to better support European researchers. The standards, tools and services developed must be reviewed and upgraded to keep pace with demand, such as developments of new sensors, and international and IT standards. Also EMODnet and Copernicus pose extra challenges to boost performance and foster INSPIRE compliance. More data from more data providers must be made available, from European and international research projects and observing programmes.

In this context the new SeadataCloud project, aimed at building upon and expanding the achievements of the SeaDataNet infrastructure, has the following main aims and objectives:

- 1. To enhance and innovate the SeaDataNet standards, products and services** offered to an expanded multi-disciplinary community:
 - To innovate and improve the functionality, performance and quality of the present data discovery and access services of the SeaDataNet infrastructure by adopting a European cloud environment (EUDAT);
 - To expand the range of services of the SeaDataNet infrastructure by specifying, developing and deploying advanced e-services to facilitate individual and collaborative research by using, handling, curating, quality controlling, transforming and processing marine and ocean data into value-added analyses, harmonised data collections, and data products;
 - To progress standardisation by adopting ISO, OGC, and W3C standards, in particular by refining existing and formulating new dedicated profiles for marine metadata, and data formats, expanding the SeaDataNet vocabulary services with new lists, developing and deploying enhanced services (Web Processing Services - WPS, Sensor Observation Services - SOS), and Semantic Web applications (e.g. Linked Data) for the SeaDataNet European directories;

- To facilitate the implementation of these standards by offering to data centres a preconfigured and pre-built system including all necessary data management tools, easily deployable and ready to use with minimal setup;
- To develop synergy and tuning between the SeaDataNet network of dedicated ocean and marine data centres, and the upcoming practice of academic data publishing with persistent identifiers (e.g. DOI) for citation purposes;

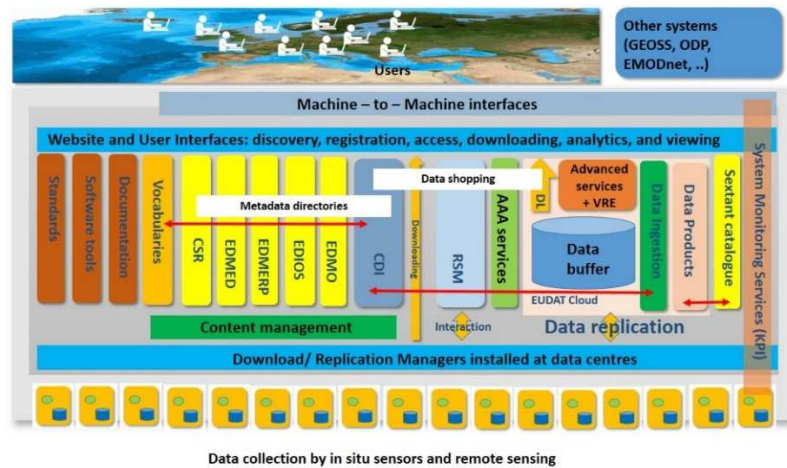


Figure 1 – Proposed SeaDataCloud architecture with data replication, advanced services and VRE in the cloud

2. To promote the adoption of the protocols and standards developed for interoperability to other key downstream initiatives in the field to expand the communities of data providers and users:

- To analyse and improve meeting the requirements of lead user communities, such as Copernicus Marine Environmental Monitoring Service (CMEMS), the European Marine Observation and Data network (EMODnet) portals, large marine observing networks (EuroGOOS, AtlantOS, Euro-ARGO, EMSO, JERICO-Next...);
- To contribute and tune standards with the INSPIRE community;
- To contribute to global portal initiatives, such as the IOC-IODE – Ocean Data Portal (ODP), GEOSS and the International Council for Science World Data System (ICSU WDS);
- To increase considerably the number of users and transactions by expanding and engaging potential user communities by promotion, dissemination, development and demonstration of use cases, and organising workshops;
- To secure more metadata, data input and data circulation in the SeaDataNet infrastructure by further development of the marine data centre network, organizing training workshops, thereby promoting and supporting adoption and implementation of SeaDataNet standards, tools and services;
- To deploy interoperability solutions and establish exchange with relevant international data management systems in the world, also making use of the Ocean Data Interoperability Platform (ODIP) results, for giving users overview and access to these data sets;

3. To present a long-term sustainable arrangement for the integrated SeaDataNet infrastructure and network of the key data centres in Europe for in situ and remote sensing data for marine research (including coastal research) and their resources:

- To achieve an overall capacity of all connected data centres on a pan-European scale to provide up-to-date and high quality metadata and data access services for the in situ and remote sensing data managed by their centres;
- To formalise the cooperation and to establish an official long-term arrangement with the CMEMS for synergy and mutual exchanges of data and data products;
- To develop an exploitation and business plan for long-term sustainability of the integrated infrastructure and take steps towards fostering long-term cooperation and integrated operation of the key data centres.