

SeaDataNet

by Robert Keeley¹

Introduction

SeaDataNet (www.seadatanet.org) is the pan-European infrastructure for ocean and marine data management as developed and operated by National Oceanographic Data Centres and other marine data centres in countries around the European seas, with co-funding from the European Commission (EC). The first steps were undertaken in the Sea-Search project (2002 – 2005), which was succeeded in 2006 with the five-year SeaDataNet project and continued in 2011 for another four years with the SeaDataNet II project. The EU (European Union) had adopted or was poised to adopt a number of directives that require the accessibility of a wide variety of marine data and information that was being collected. At the same time, the EC recognized the wide diversity of capabilities for managing data in the EU. The SeaDataNet infrastructure and its standards and services are meeting these requirements. There are 55 partner institutions [including the Intergovernmental Oceanographic Commission (IOC), the International Council for the Exploration of the Seas (ICES), and the EU Joint Research Centre (EU-JRC)] in the SeaDataNet II project coming from 35 countries. Partner institutions are responsible for assembling the variety of marine data and information collected within their nation (shared responsibilities are arranged when two or more institutions cooperate in the data collection). The project began with targeting physical oceanographic data, but that has expanded now to include biological, geological, bathymetric, and chemical data and information from moored platforms, autonomous vehicles, and ships. As part of building the pan-European infrastructure, SeaDataNet partners are active in disseminating their approach. That is resulting in the adoption of SeaDataNet protocols, infrastructure, and tools by other EU projects and projects outside of Europe. In addition to assembling, processing, and distributing data and information, there is also a component within SeaDataNet to generate products.

Components and Technologies

SeaDataNet has a great variety of component developments. Most of them focus on standardizing the assembly, treatment, archiving, and dissemination of marine data and information. Other components are concerned with product development, communications, and exploring options for sustainability including seeking continued EU co-funding after the end of SeaDataNet II. The SeaDataNet activities have greatly enlarged and harmonized the European capacity for marine data management. The SeaDataNet infrastructure has expanded beyond the

SeaDataNet consortium partners to include other marine data centres. This is reflected in a pan-European infrastructure which at present is connecting more than 100 data centres from 35 countries and which is used by many institutes and projects at national, regional, and EU scales. Altogether SeaDataNet handles data sets from more than 500 organizations.

SeaDataNet does not contribute to data acquisition. However, one of the first responsibilities of SeaDataNet partners is to be sure that what is collected is reported in a standard form called a Cruise Summary Report (CSR). This report has been around for many years, but SeaDataNet has built software standardized on ISO-19115 - 19139, to streamline the reporting requirements and allow for cruise track maps to be included. The German Data Centre is responsible within the project for acquiring CSRs from national SeaDataNet partners and has constructed a database containing them. This is supported by a number of other standardized catalogues that are maintained and managed by SeaDataNet. These include lists of projects underway or completed, European organizations, descriptions of data set collections, descriptions of monitoring activities and networks, and controlled vocabularies. The various catalogues are interrelated and each is managed at the European level by one of the SeaDataNet partners while individual project partners are responsible for ensuring their national entries are correct. These catalogues and vocabularies are used to ensure a common naming convention for all data collected. CSRs enter the SeaDataNet realm even before the data do.

Once data and information arrive at a SeaDataNet partner institution, a metadata record is created in a common structure called the Common Data Index (CDI). This is a centralized facility operated within the SeaDataNet project by the Marine Information Service in the Netherlands (MARIS). The format of the CDI is an ISO 19115 - 19139 profile and each partner is responsible for creating records using software tools developed by IFREMER (French Research Institute for Exploitation of the Sea) in France, another SeaDataNet partner. Through software and network connectivity to the various data centres contributing to the project, MARIS harvests CDI entries and updates an on-line catalogue of holdings. This catalogue is the point of entry for users wanting to identify data of interest and for requesting access to the data sets that are managed at each of the connected data centres.

¹ Retired oceanographer; formerly with the Marine Environmental Data Service of the Department of Fisheries and Oceans, Ottawa, Ontario, Canada

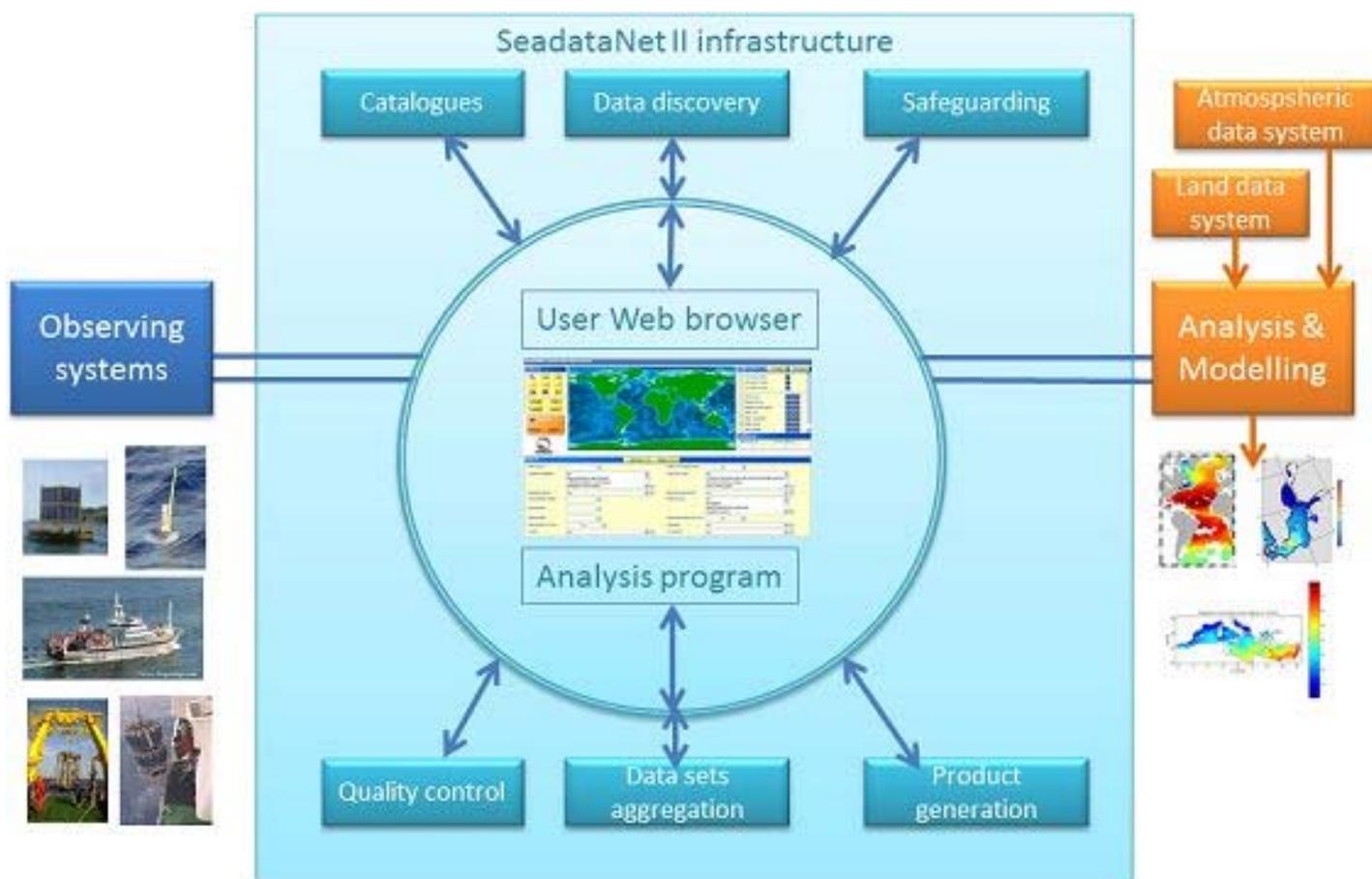


Figure 1: A schematic of the SeaDataNet infrastructure

A single CDI entry is quite detailed. For example, a single CTD cast, a single drifting buoy location, a single time series, a single biological observation all generate single CDIs. This makes for many, many entries in the CDI, but allows for a very precise search capability. All data, whether available with no restrictions or with some distribution restrictions, are recorded in the CDI.

When data arrive at SeaDataNet partners, each uses a common software tool to check the quality of the data. This tool has been built by greatly extending the software application called Ocean Data View (ODV) developed by Reiner Schlitzer of the Alfred Wegener Institute in Germany. He wrote the original software to provide a means of data visualization for himself. As part of SeaDataNet, he has extended this tool so that it now handles a variety of data types, allows for quality control operations, checks for data duplications, permits the inclusion of detailed metadata such as the SeaDataNet CDI and records data processing history within ODV.

All of these tools already mentioned and others are tied together by a common vocabulary. This is maintained by the British Oceanographic Data Centre (BODC) in the UK. It began as an effort to standardize naming of ocean variables

and units of measurement. It has evolved into a full fledged vocabulary server using modern technology that allows machine-to-machine interactions. This means that software that is constructing a CDI in one country can consult the vocabulary server to check that the name used for a particular variable is known and well described. BODC has built a community of experts to help extend and improve a combined SeaDataNet and MarineXML Vocabulary Content Governance Group (SeaVoX). More information about SeaVoX can be found at https://www.bodc.ac.uk/data/codes_and_formats/seavox/.

Data Access

The user interface to the CDI catalogue was built by MARIS and uses standard Web Map Service technologies to display information. The user can search in a variety of ways such as by type of data, area, time frame, etc. Once data of interest are identified, a request to receive the data can be launched.

Access to the CDI is freely available (see above SeaDataNet link). However, a request for data download requires user registration. The registration is so that SeaDataNet can know who is in their user community and facilitates tracking requests. Also, although a large fraction

of the data made available through the CDI is open to all, there are some that require consultations with data originators before they may be released. Obtaining release may involve communication with the user, hence the need for registration. The SeaDataNet Data Policy is easily found under the "Data Access" tab on the SeaDataNet web pages.

Data within SeaDataNet are maintained by the connected data centres and in their own institutions. Once a query for access to data has been formulated through the on-line interface, a request manager tool distributes the request to all of the institutions holding the data. Software tools built through SeaDataNet and installed at partner institutions receive the request and assemble the data that they hold. A user can monitor the progress of data assembly. When the request is completed, the user is notified. Data are delivered in the SeaDataNet ODV ASCII format and soon also in a netCDF-CF (Common Data Format-Climate and Forecast) compliant format.

SeaDataNet has developed tools to monitor the state of network connections among partners. The capabilities of this software are quite sophisticated and it allows for the careful scrutiny of accessibility of partner operations. The monitoring of all services and data centre connections supports the objective to provide a robust and operational infrastructure.

Products

The main product currently under development is a set of climatologies for waters surrounding Europe. It is being created using the DIVA (Data-Interpolating Variational Analysis) software developed by the University of Liège. The software is free and available through the SeaDataNet link provided on the previous page.

Collaborations

SeaDataNet has collaborations with institutions such as ICES (International Council for the Exploration of the Sea) to ensure common naming of ships used in ocean research. IOC/IODE (Intergovernmental Oceanographic Commission/International Oceanographic Data and Information Exchange) provides training facilities at its offices in Belgium. The EU-JRC is leading an activity to analyze legal models for sustaining SeaDataNet as a legal entity. SeaDataNet has contributed ideas to the European Marine Observation and Data Network (EMODnet) initiative which was started in 2008 by the EU. SeaDataNet infrastructure and standards have been adopted by EMODnet projects for physics, chemistry, and bathymetry, and they are also contributing to biology and geology projects. This engagement has extended the types of data that SeaDataNet can manage and deliver. Moreover it has resulted in more data centres connecting to the SeaDataNet infrastructure and making their data available through the SeaDataNet CDI service. More recently, a collaboration with USA and Australia data centres within the Ocean Data Interoperability Platform (ODIP) has seen the adoption of

many of the vocabularies and a recognition of the usability of software tools.

The Future

The SeaDataNet II project has one year to complete its agreed deliverables. There are still many things to do. The objective of formulating Sensor Web Enablement (SWE) standards and implementing these for selected instruments and platforms is well underway. The suite of climatologies building upon the SeaDataNet data collections is making good progress but the final and public version remains to be constructed. Some data centres have yet to contribute additional descriptions of their data holdings and must install and configure upgraded versions of the software tools built by the project.

In September 2015, the SeaDataNet II project will officially end. A final report will be delivered to the EC some months later. It will include comprehensive documentation of the project's accomplishments. They are extensive, and have gone beyond the original objectives in a number of areas. There is little doubt that the experience gained in unifying the handling of the many types of data collected in the marine environment, the software tools that draw on the standardized vocabularies, the adoption of ISO standards for these tools and the quality control, duplicate detection and visualization tools will impact ocean data management throughout the world. This will be promoted further by the SeaDataNet consortium that will continue the operation of the SeaDataNet infrastructure for at least another three years as part of a SeaDataNet exploitation agreement. The National Oceanographic Data Centres in Europe and most other connected marine data centres are also actively engaged in the EMODnet projects and other EU projects which are making use of the SeaDataNet infrastructure, standards, tools, and services. On top of that, SeaDataNet is starting an active lobby towards their national authorities and the EU for another phase of the SeaDataNet research and technical developments. There are many new challenges for ocean and marine data management, considering new instruments and platforms, new data types, big data, cloud processes, visualization techniques, and societal questions which demand further development and maintenance of standards, services, and integrated capacities. It is hoped that the EU will recognize the importance to support the continued collaboration, use of the expertise, mass and momentum gained by SeaDataNet as a pan-European approach rather than by fragmented national developments.

Note from the Co-Editor: Dick Schaap (dick@maris.nl) and Michele Fichaut (Michele.Fichaut@ifremer.fr), co-chairs of SeaDataNet, provided comments and changes to the original text. Stephen Miller (Emeritus scientist at Scripps Institution of Oceanography), Margarita Gregg (Director of the US National Oceanographic Data Center), Hans Dahlin (formerly at the Swedish Meteorological and Hydrological Institute), and Robert Keeley constitute the Advisory Board to the SeaDataNet project.