

SeaDataNet – panEuropean infrastructure for marine and ocean data management – an update

Dick Schaap¹ (dick@maris.nl), Michele Fichaut² (m.fichaut@free.fr),
Steven Piel² (steven.piel@ifremer.fr), Marine Vernet² (marine.vernet@ifremer.fr),
Peter Thijsse¹ (peter@maris.nl), Serge Scory³ (serge.scory@naturalsciences.be),
Alessandra Giorgetti⁴ (agiorgetti@inogs.it), Anastasia Iona⁵ (sissy@hnodc.hcmr.gr),
Hong Minh Le³ (hongminh.le@naturalsciences.be), Lesley Rickards⁶ (ljr@noc.ac.uk)

¹MARIS (The Netherlands)

²Ifremer (France)

³RBINS (Belgium)

⁴Istituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS (Italy)

⁵HCMR (Greece)

⁶National Oceanography Centre -NOC, British Oceanographic Data Centre - BODC (UK)

Introduction

SeaDataNet is a major pan-European infrastructure for managing, indexing and providing access to marine data sets and data products, acquired by European organisations from research cruises and other observational activities in European coastal marine waters, regional seas and the global ocean. Founding partners are National Oceanographic Data Centres (NODCs), major marine research institutes, UNESCO-IOC, and ICES. The SeaDataNet network was initiated in the nineties and over time its network of data centres and infrastructure has expanded, during a successive series of dedicated EU RTD EU projects, such as SeaSearch, SeaDataNet, SeaDataNet 2, and SeaDataClod, and by engaging as core data management infrastructure and network in leading European initiatives such as the European Marine Observation and Data network (EMODnet), Copernicus Marine Service (CMS), and the European Open Science Cloud (EOSC) and closely cooperating with ocean observing communities such as EuroGOOS and Euro-Argo. These facilitated an ongoing development and evolution of the SeaDataNet technical infrastructure, standards, tools, and services, while managing and further expanding a large network of connected data centres and data providers.

SeaDataNet Data Management services

SeaDataNet develops, governs and promotes common standards, vocabularies, software tools, and services for marine data management, which are freely available from its portal and widely adopted and used by many communities and projects. The SeaDataNet network of data centres maintains and publishes a series of European directory services. These give a wealth of data and information, such as overviews of marine organisations in Europe (EDMO), their engagement in marine research projects (EDMERP), managing large datasets (EDMED), data acquisition by research vessels (CSR) and monitoring programmes (EDIOS) for the European seas and global oceans. A core SeaDataNet service is the Common Data Index (CDI) data discovery and access service which provides online unified discovery and access to vast resources of data sets, managed by 115+ connected SeaDataNet data centres from 34 countries around European seas, both from research and monitoring organisations. Currently, it gives access to more than 2.9 Million data sets, originating from 950+ organisations in Europe, covering physical, geological, chemical, biological and geophysical data, and acquired in European waters and global oceans.

SeaDataNet CDI Data Discovery and Access Service

The CDI service infrastructure consists of several components. Offline software packages

(MIKADO, NEMO, OCTOPUS) facilitate manual and machine generation of metadata entries in SeaDataNet standard profiles, based upon ISO19115-19139, and converting data sets into SeaDataNet standard data formats (ODV ASCII and NetCDF), supported by an ever increasing set of controlled vocabularies. Moreover, the separate online directories (EDMO, EDMED, EDMERP, CSR, and EDIOS) are maintained with input from NODCs for activities in their countries. Together, these contribute to marking up the metadata profiles in a semantically controlled way. Overall, this results in rich and highly FAIR metadata and data sets, useful for various applications. Other online CDI service components aim at populating the CDI metadata and harmonised related data sets into a central online CDI catalogue and a cloud based data cache. Once imported, validated, and accepted, the CDI metadata and data sets can be discovered and retrieved by users through an online shopping mechanism with dashboard for users and data centres for overseeing requests and download transactions.

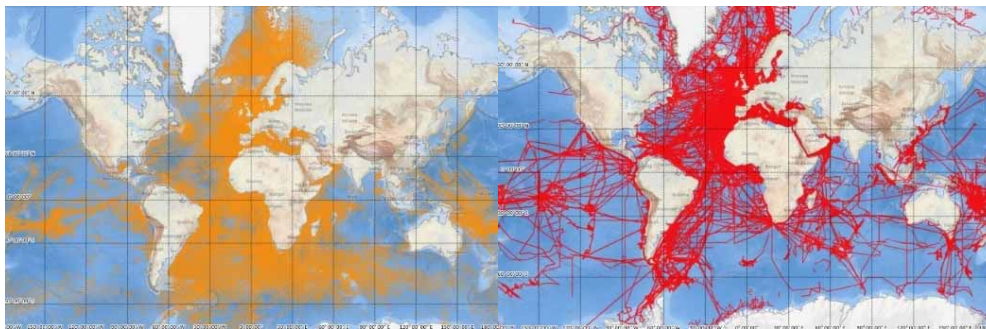


Figure 1 Coverage of data sets for point and track observations in the SeaDataNet CDI data discovery and access service.

The FAIRness of the SeaDataNet services is further amplified by running SPARQL endpoints for machine-to-machine services, using Linked Data principles, and a Swagger API for the CDI service. Use of common vocabularies in all metadatabases and data formats is an important prerequisite towards consistency, interoperability, and FAIRness.

Major pillar under EMODnet

SeaDataNet is providing core services in EMODnet Chemistry, Bathymetry, Physics, and Ingestion for bringing together and harmonizing large amounts of marine data sets from hundreds of data originators (currently >950), which are then used by EMODnet groups for generating thematic data products. Such products are for example: a Digital Terrain Model for all European seas and NW Atlantic Ocean (Bathymetry), and European aggregated and validated data collections for eutrophication, contamination, and marine litter (Chemistry). These products are very popular and find their way to many users from government, research, industry, and public. Users also pose further requirements, which leads to further optimising of the data flows from originators to data centres to SeaDataNet and to EMODnet thematics. Moreover, it implicates regular maintenance, upgrading and expanding of standards, tools, and services, also because of adding new data types, such as e.g. marine litter or vessel-mounted ADCP. Further uptake of standards and services is promoted and established, resulting in more data population, and expansion of the networks of data centres and data originators. Currently, SeaDataNet not only counts NODCs, but also many other data centres from marine research institutes, geological surveys, national hydrographic services, environmental management agencies, and companies.

Further innovation of services

Innovative developments are taking place for adopting new technologies and responding to new challenges. The European Open Science Cloud (EOSC) gives a suitable framework, promoting FAIR data services and interoperability. SeaDataNet is well engaged in several EOSC projects, such as ENVRI-FAIR, EOSC-FUTURE, Blue-Cloud pilot, and currently, in the Blue-Cloud 2026 project. The Blue-Cloud pilot deployed a versatile cyber platform with smart federation of multidisciplinary data repositories, analytical tools, and computing facilities in support of exploring and demonstrating the potential of cloud based open science for ocean sustainability. Blue-Cloud 2026 project aims for a further evolution into a Federated European Ecosystem to deliver FAIR & Open data and analytical services, instrumental for deepening research of oceans, coastal & inland waters. SeaDataNet benefits from further innovating and expanding its services. Using its large knowledge base, network, and experience, it is also a major driver towards other blue data infrastructures stimulating uptake and refinement of SeaDataNet standards, and engagement in open science and virtual research. This contributes to optimizing the overall FAIRness in the European marine data landscape. This is highly important, also in the perspective of the Digital Twins of the Oceans (DTO) initiative as part of the EU Green Deal and DestineE.