SeaDataCloud – further developing the pan-European SeaDataNet infrastructure for marine and ocean data management

Dick M.A. Schaap (MARIS – Netherlands)
Technical Coordinator SeaDataCloud Project

Chris Ariyo (CSC – Finland)
Project Manager SeaDataCloud Project (EUDAT)

DIGITAL INFRASTRUCTURES for RESEARCH 2017
“Connecting the building blocks for Open Science”
Brussels, Belgium, 30th November – 1st December 2017
Acquisition of ocean and marine data in many ways
Ocean and Marine Data acquisition

- Data are collected by governments, research institutes, and private industry (in Europe more than 1,000 organisations)

- Data are gathered for physics, geophysics, chemistry, biology, geology, and bathymetry

- Acquisition of oceanographic and marine data is expensive; annual costs in Europe estimated at 1.4 Billion Euro (1.0 = in-situ; 0.4 = satellites)

Professional data management is required with agreements on standardization, quality control protocols, archiving, catalogues, and access.
Collect once; use many times!
What is SeaDataNet?

A pan-European infrastructure set up and operated for managing marine and ocean data in cooperation with the NODCs and data focal points of 35 countries bordering the European seas.

<table>
<thead>
<tr>
<th>Period</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006-2011</td>
<td>Metadata directories</td>
</tr>
<tr>
<td></td>
<td>MEDAR/MedAtlas</td>
</tr>
<tr>
<td>2002-2005</td>
<td>Sea-Search (FP5)</td>
</tr>
<tr>
<td>2006-2011</td>
<td>SeaDataNet (FP6)</td>
</tr>
<tr>
<td>2011-2015</td>
<td>SeaDataNet II (FP7)</td>
</tr>
<tr>
<td>2016-2020</td>
<td>SeaDataCloud (H2020)</td>
</tr>
</tbody>
</table>
Portal with standards, tools, and services, both for users and data centres

www.seadatanet.org
SeaDataNet metadata directories

- EDMER: Projects
- EDIOS: Observing programmes
- EDMO: Organisations
- EDMED: Data sets
- CSR: Research cruises
- CDI: Data index
SeaDataNet portal

CDI service for discovery and unified data access

SeaDataNet

Already 102 data centres connected and more underway

European data sources
data centres > 600 originators

Data centres

Search and Shop

Metadata + transaction data
1.95 million CDI entries from 34 countries, 102 data centres and 612 originators for physics, chemistry, geology, geophysics, bathymetry and biology; from 1805 to 2017; 87.6% unrestricted or under SDN License
SeaDataNet products

Analysis of data anomalies
QC analysis
File and parameter aggregation
Data harvesting
CENTRAL CDI
SeaDataNet Quality Checks Strategy (QCS)

Aggregated datasets and climatologies
Improvement of the data quality

Access to the SeaDataNet Products Catalogue
Regional products
SeaDataNet cooperation and involvement

- **Copernicus Marine Environmental Monitoring Services (CMEMS):** providing long-term archives and standards
- **Marine Strategy Framework Directive (MSFD):** providing infrastructure, standards and data collections for several indicators
- **Large ocean monitoring systems (EuroGOOS, AtlantOS, Euro-ARGO, JERICO-Next, ..):** providing standards and validation + long-term archiving services
- **Ocean Data Interoperability Platform (ODIP):** exploring and demonstrating common standards and interoperability with leading data management infrastructures in USA and Australia
- **GEOSS - EuroGEOSS:** Maintaining the GEOSS portal with SeaDataNet in-situ data collections from large community of European data holders (> 100 data centres; >600 data originators)
- **European Open Science Cloud (EOSC):** shaping the pilot Blue Cloud
SeaDataNet and EMODNet

- EU initiative for an overarching European Marine Observation and Data Network (EMODNet) driven by Marine Knowledge 2020 and Blue Growth

- SeaDataNet qualified as a leading infrastructure for the EMODnet data management component and is driving several thematic portals from the start in 2008

- This synergy has resulted in many more data centres adopting SeaDataNet standards and connecting to the CDI Data Discovery and Access service while it gave a flying start to EMODnet
PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

SeaDataNet

EMODnet Central Portal www.emodnet.eu

Data
- Bathymetry: Data on bathymetry (water depth), coastlines, and geographical location of underwater features such as wrecks
- Geology: Data on seabed substrate, seafloor geology, coastal behaviour, geological events and probabilities, and minerals
- Human activities: Data on the intensity and spatial extent of human activities at sea
- Seabed habitats: Data on modelled seabed habitats based on seabed substrate, energy, biological zone and salinity
- Physics: Data on salinity, temperature, waves, currents, sea level, light attenuation and FerryBox data
- Biology: Data on temporal and spatial distribution of species abundance and biomass from several taxa
- Chemistry: Data on concentrations of chemicals (pesticides, heavy metals, antifoulants) in water, sediments and biota

Metadata

Data Products

Data Services
Example of **EMODnet Bathymetry** – using > 7000 survey data sets to generate and provide a harmonized and higher resolution digital terrain model for all European seas – comparison with GEBCO
SeaDataCloud – cooperation with EUDAT

European Collaborative Data Infrastructure

A consortium of high performance computing (HPC) / data centres, libraries, scientific communities, data scientists.
EUDAT ENV Core Communities

IMDIS 2016, Gdańsk, 11-13 October 2016

EUDAT ENV Data Pilots

Institutions and partners

For more information visit - https://eudat.eu/use-cases
• Current emphasis on moving towards a Partnership Agreement between service providers (establishment in September 2016)

• Movement beyond Partnership (e.g. ERIC) will depend on circumstances and will be decided in close discussion with the other e-Infrastructures (e.g. PRACE, EGI, GEANT), in the context of the European Open Science Cloud
SeaDataCloud – general challenges

• It is about updating and further developing standards

• It is about improving and innovating services & products

• It is about adopting and elaborating new technologies

• It is about giving more attention to users and putting the user experience in a central position

• Moreover, it is about implementing a strategic and operational cooperation between the SeaDataNet consortium of marine and ocean data centres and the EUDAT consortium of e-infrastructure service providers
SeaDataCloud topics

• **Standards:**
  – Vocabularies
  – Data Formats
  – INSPIRE compliance
  – Sensor Web Enablement (SWE)
  – Data Management for new data types

• **Services:**
  – Implementing Linked Data for SeaDataNet directories for M-to-M services
  – Upgrading the CDI Data Discovery and Access service making use of the cloud
  – Integrating data offering from international programmes and organisations
  – Integrating INSPIRE Transformation services
  – Data publishing
Upgrading of CDI Data Discovery and Access service

- Introducing central data cache
  - Extra QA-QC
  - More efficient delivery
  - Higher performance
  - Transformations
- Replication for synchronisation
NEW FOCUS: ACHIEVING MORE USER ENGAGEMENT

WP10 Downstream Services

WP9 Upstream Services

Discovery and access to more datasets and information

WP8 Standards & Vocabularies

make it work!

Added-value services and applications
SeaDataCloud topics

- Developing and deploying a Virtual Research Environment (VRE):
  - Collaborative environment
  - Cloud computing
  - Big data
  - Advanced e-services to facilitate research for handling, curating, quality controlling, transforming and processing marine and ocean data into value-added analyses, harmonised data collections, and data products which can be integrated, visualised and published using OGC and high level visualisation services.

- Products:
  - Improved T&S Climatology – cooperation with Copernicus Marine Environmental Monitoring Service
Contacts information

Dick M.A. Schaap (MARIS – Netherlands)
Technical Coordinator SeaDataCloud Project

Chris Ariyo (CSC – Finland)
Manager SeaDataCloud Project (EUDAT)

SeaDataNet: https://www.seadatanet.org/
EUDAT: https://b2(service).eudat.eu/
EUDAT: http://www.eudat.eu/

EUDAT Partners in SeaDataCloud:
CINECA (Italy), CSC (Finland), DKRZ (Germany), GRNET (Greece), and STFC (United Kingdom)