Data Infrastructures: European and Global

Mark Hebden

British Oceanographic Data Centre
Introduction

European

• SeaDataNet infrastructure

(SEA-SEARCH > SeaDataNet > SeaDataCloud)

• Metadata to data to cloud

• Linkage with other European data infrastructures
e.g. EMODnet, Copernicus Marine Environment
Monitoring System (CMEMS), EuroGOOS

Global

• IOC/IODE – ICES – IHO – WMO – ISC - RDA
Acquisition of marine data

- Fixed buoys
- Profilers
- Oceanographic vessels
- Commercial ships
- Fishing vessels
- Planes
- HF Radar
- Gliders
- Sea elephants
- Sea floor observatories
- Satellites
- Profilers
- Submarines
- Surface
- Water column
- Sea bottom
European marine landscape
A pan-European infrastructure set up and operated for managing marine and ocean data in co-operation with National Oceanographic Data Centres and oceanographic data focal points from 35 countries bordering European seas.

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990s</td>
<td>Metadata directories, MEDAR/ MedAtlas</td>
</tr>
<tr>
<td>2002-05</td>
<td>SEA-SEARCH</td>
</tr>
<tr>
<td>2006-11</td>
<td>SeaDataNet</td>
</tr>
<tr>
<td>2011-15</td>
<td>SeaDataNet II</td>
</tr>
<tr>
<td>2016-20</td>
<td>SeaDataCloud</td>
</tr>
</tbody>
</table>
SeaDataNet partners and involvement

• SeaDataNet core partners are the National Oceanographic Data Centres (NODCs), major marine research institutes (RIs) in Europe and EU Joint Research Centre (JRC)

• Members have institute and national tasks for marine data management, dealing with data as collected through scientific cruises with research vessels and collected by various other platforms. Emphasis on in situ data (delayed mode) and providing validation and long term stewardship

• NODC members also participating in International Oceanographic Data and Information Exchange (IODE) programme of the Intergovernmental Oceanographic Commission of UNESCO (IOC-IODE) and in the Data Management Workgroup of the International Council for the Exploration of the Sea (ICES), discussing marine data management topics on an international basis.
SeaDataNet standards

- Set of common standards for the marine domain, adapting ISO and OGC standards and achieving INSPIRE compliance
- **Adoption of ISO 19115 – 19139 standard for describing metadata** on data, data sets, research cruises, monitoring networks, and research projects => marine metadata profiles, schemas, schematron rules
- **Controlled vocabularies** for the marine domain (>84,000 terms in 111 lists), with international governance and web services
- **Standard data exchange formats** : ODV ASCII and NetCDF (CF) and MedAtlas ASCII fully supported by controlled vocabularies
- Maintenance and dissemination of standard **QA-QC procedures**, together with IOC/IODE and ICES (QC-loop)
SeaDataNet principles

• Traditionally, SeaDataNet is based on a semi-distributed system (centralised metadata catalogues and distributed data), relying on the existing data centre network.

• The SeaDataNet RI enables a user to see the data from all the data centres connected as a virtual unique data centre, delivering integrated data, metadata and products.

• SeaDataCloud – aligning with the European Open Science Cloud (EOSC) vision…
Centralised SeaDataNet metadata directories

- EDIos
- CDI
- EDMED

- Projects (3,105)
- Observing programmes (363)
- Data index (2,276,264)
- Data sets (4,187)

- EDMERp
- Organisations (4,068)
- Research cruises (56,694)
CDI (Common Data Index) Service
Data collection by in situ sensors and remote sensing
SeaDataNet partners in many projects

SeaDataNet members are participating in many EU projects in which the SeaDataNet infrastructure and standards are adopted and further developed.
SeaDataNet cooperation

- **Copernicus Marine Environmental Monitoring Services (CMEMS):** providing long-term archives and data management standards. MoU describing cooperation and synergy options, in particular with CMEMS-INSTAC

- **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 as integrator

- **Ocean Data Interoperability Platform (ODIP):** exploring and demonstrating common standards and interoperability with leading data management infrastructures in USA and Australia

- **GEOSS - EuroGEOSS:** serving 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
• Initiative of EU DG MARE, driven by Marine Knowledge 2020 Strategy and Blue Growth

• A network of organisations supported by the EU’s integrated maritime policy. Work together to observe the sea, process data to international standards and make freely available as **interoperable data layers and data products**

• “Collect once and use many times”

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

- Bathymetry
  - Minimum cell water depth
  - Maximum cell water depth
  - Average cell water depth
  - Standard deviation of cell water depth
  - Number of values used for interpolation of cell water depth
  - Horizontal coordinate reference system
  - Depth reference system
  - Lowest Astronomical Tide

- Geology
  - Seabed substrate
  - Sediment accumulation rate
  - Sea-floor geology
  - Seabed lithology
  - Stratigraphy
  - Coastline migration
  - Aggregate resources
  - Geological events

- Seabed Habitats
  - Data on modelled seabed habitats (depth, seabed substrate, broad scale biological zone, T, S, light, oxygen, energy due to waves & current)
  - Broad-scale physical habitat map
  - Detailed habitat maps from surveys
  - Individual habitat modelling outputs
  - Habitat point data

- Chemistry
  - Pesticides & Biocides
  - Antifoulants
  - Pharmaceuticals
  - Heavy Metals
  - Hydrocarbons
  - Radionuclides
  - Fertilizers
  - Acidity
  - Dissolved Gasses
  - Plastics
  - Marine Litter Beach litter
  - Seafloor litter
  - Micro litter
  - Chlorophyll
  - Silicates
  - Organic Matter

- Biology
  - Biomass
  - Abundance
  - Gridded Abundance maps
    - species groups:
      - phytoplankton
      - zooplankton
      - angiosperms
      - macro-algae
      - invertebrate bottom fauna
        - birds
        - mammals
        - reptiles
        - Fish

- Physics
  - Waves
  - Water temperature
  - Water salinity/conductivity/density
  - Currents
  - Light attenuation/fluorescence
  - Sea level
  - Atmospheric parameters
    - Wind
    - Underwater noise
    - Rivers
    - Ice

- Human activities
  - Aggregate Extraction
  - Dredging
  - Fisheries
  - Hydrocarbon Extraction
  - Main Ports
  - Mariculture
  - Ocean Energy
  - Facilities
  - Pipelines and Cables
  - Protected Areas
  - Waste Disposal
  - Wind Farms
  - Other Forms of Area Management / Designation
CDI service as driver for portals

Total collection

Data discovery and access

> 110 data centres

NODCs; HOs; GEOs; BIoS; ICES;

> 650 European data originators

CDI Data Discovery and Access service

GEOSS portal

IODE ODP portal

Aggregated collection

Regional subsets

Black Sea portal

Caspian portal

Geo-Seas portal

Thematic subsets

Biology

Chemistry

Geology

Physics

EMODnet

Bathymetry
EMODnet Physics
Pillars under EMODnet Physics

**European Global Ocean Observing System**

- Association of agencies to further develop operational oceanography in European Sea areas and adjacent oceans
- Component Regional Ocean Observing Systems (ROOSs) responsible for collection of in situ data.

**Copernicus Marine Environment Monitoring System (CMEMS)**

- Operational service responding to emerging issues in environmental, business and scientific sectors.
- Analyses and forecasting using satellite/in situ observations
- Operates the In Situ Thematic Assembly Centre (CMEMS-INSTAC) packaging observations from EuroGOOS ROOSs into QC’d data collections - forecasting services, validation processes and downstream NRT applications.

**SeaDataNet**

- Pan-European marine data management infrastructure and network of NODCs
Welcome to the Data Submission service of the EMODnet Data Ingestion portal. This service facilitates you to submit marine datasets by completing a form and uploading your data as a package. Press "Continue" to start.
Principle data flow

- Mobilise ‘sleeping’ data to serve ‘Blue Society’
Workflow in practice…

Example:
Marine Scotland Science cruises:
Data now in SeaDataNet, via EMODnet Ingestion and feeding other EMODnet portals
International/Intergovernmental Organisations

- UNESCO
- IOC
- Intergovernmental Oceanographic Commission
- IHO
- International Hydrographic Organization
- ICES
- CIEM
- International Council for the Exploration of the Sea
- WMO
- World Meteorological Organization
- RDA
- Research Data Alliance
International/Intergovernmental Organisations

Intergovernmental Oceanographic Commission (IOC)

- International Oceanographic Data and Information Exchange (IODE)
- Global Ocean Observing System (GOOS)

IOC-WMO

- Joint Technical Committee on Oceanography and Marine Meteorology (JCOMM)

IOC-IHO – GEBCO

- Guiding Committee
- Seabed2030

International Council for the Exploration of the Sea (ICES)

- Data Portals
- Data and Information Group

International Science Council

- World Data System (WDS) and WDS International Technology Office (ITO)
- Scientific Committee on Oceanic Research (SCOR)
- Scientific Committee on Antarctic Research (SCAR)

Research Data Alliance
Established in 1960
Promotes international cooperation and coordinates programmes in:
- marine research
- observation systems
- services (including data and information)
- hazard mitigation
- capacity development
To understand and effectively manage the resources of the ocean and coastal areas
Contributing to the UN Decade of Ocean Science for Sustainable Development (2021-2030)
Over 60 National Oceanographic Data Centres (NODCs)
30 Associate Data Units (ADUs)
Ocean Biogeographic Information System (OBIS)
Standards and best practices
Global data projects
Ocean Data Portal
Regional Ocean Data and Information Networks (ODINs)
Capacity Building
Ocean Biogeographic Information System (OBIS)

- A global open-access data and information clearing-house on marine biodiversity for science, conservation and sustainable development
- Emanates from the Census of Marine Life (2000-2010)
- Adopted as a project under IOC-UNESCO’s IODE programme in 2009
- More than 20 OBIS nodes around the world connect 500 institutions from 56 countries
- Provided over 58 million observations of nearly 130,000 marine species, from Bacteria to Whales, from the surface to 10,900 meters depth, and from the Tropics to the Poles.
- Datasets are integrated so you can search and map them all seamlessly by species name, higher taxonomic level, geographic area, depth, time and environmental parameters
Main in situ Elements of the Global Ocean Observing System

January 2020

Profiling Floats (Argo)
- Core (3983)
- Deep (136)
- BioGeoChemical (387)

Data Buoys (DBCPC)
- Surface Drifters (1460)
- Offshore Platforms (94)
- Ice Buoys (31)
- Moored Buoys (325)
- Tsunameters (31)

Timeseries (OceanSITES)
- Interdisciplinary Moorings (309)
- Research Vessel Lines (63)

Repeated Hydrography (GO-SHIP)
- Automated Weather Stations (250)
- Manned Weather Stations (1269)
- Radiosondes (12)
- eXpendable BathyThermographs (32)

Sea Level (GLOSS)
- HF Radars (270)
- Animal Borne Sensors (53)

Generated by www.jcommps.org, 06/02/2020
General Bathymetric Chart of the Ocean (GEBCO)

- Aims to provide the most authoritative, publicly available bathymetry data sets for the world’s oceans
- Initiated in 1903 - First Edition published in 1905
- Now under the joint auspices of IHO and IOC
- Seabed2030 – funding from Nippon Foundation
- Aims to bring together all available bathymetric data to produce the definitive map of the world ocean floor by 2030 and make it available to all
International Council for the Exploration of the Sea (ICES)

• Intergovernmental marine science organization, meeting societal needs for impartial evidence on the state and sustainable use of our seas and oceans set up in 1902

• Well-established data centre

• Provides marine data services to ICES member countries, expert groups, HELCOM and OSPAR, EEA) and various other European projects and biodiversity portals

• Large dataset collections organised around specific thematic data portals:
  • Biological community
  • Contaminants and biological effects
  • Eggs and larvae
  • Fish predation (stomach contents)
  • Fish trawl survey
  • Historical plankton
  • Ocean physics and chemistry

• Data and Information Group (DIG)
• Interdisciplinary body of the International Science Council (ISC; formerly ICSU) created by its 29th General Assembly in Maputo, Mozambique in 2008

• Enabling universal and equitable access to quality-assured scientific data, data services, products and information

• Supported by:
  • International Project Office in Tokyo
  • International Technology Office in Canada

• 123 member organisations, including IODE
The Research Data Alliance (RDA) builds the **social and technical bridges that enable open sharing of data**

The RDA vision is researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society – membership ~10,000

Major funding for RDA is provided by the following organisations:

- The Australian Commonwealth Government through the Australian National Data Service
- European Commission’s 7th Framework Programme through the RDA Europe project
- The National Science Foundation

Currently working through 66 Interest Groups and 36 Working Groups, e.g.:

- Metadata – FAIR – PIDs – Data Publication
- Active Data Management Plans
- Vocabulary services (VSSIG)
- InteroperAble Descriptions of Observable Property Terminology WG
- Persistent Identification of Instruments
- Certification of Digital Repositories
- …