From SeaDataNet II to SeaDataCloud

Michèle Fichaut (IFREMER, France), Dick Schaap (MARIS, The Netherlands)

On behalf of SeaDataNet and SeaDataCloud communities
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 34 countries bordering the European seas.

Metadata directories:

- 90s: MEDAR/MedATlases (MAST)
- 2002-2005: Sea-Search (FP5)
- 2006-2011: SeaDataNet (FP6)
- 2011-2015: SeaDataNet II (FP7)
- 2016-2020: SeaDataCloud (H2020)
Portal with standards, tools, and services, both for users and data centres

www.seadatatenet.org
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 sets, research cruises, monitoring networks, and research projects => marine metadata profiles, schemas, schematron rules
  - **Controlled vocabularies** for the marine domain (>65,000 terms in 82 lists), with international governance and web services
  - **Standards data exchange formats**: ODV ASCII and NetCDF (CF) fully supported by controlled vocabularies
- Maintenance and dissemination of standard Quality Assurance- Quality Check (QA-QC) procedures, together with IOC/IODE and ICES
SeaDataNet services and tools

- **Set of tools** to be used each data centre and freely available from the SeaDataNet portal: metadata editor (MIKADO), data conversion software (NEMO), download manager (DM), data analysis software (ODV), data interpolation software (DIVA)

- **Pan-European services** for harmonised discovery, access, visualisation of data and data products

- Common SeaDataNet **Data Policy** and SeaDataNet License

- **Capacity building** by training workshops for uptake of standards and tools by the data centres in order to achieve standardisation
SeaDataNet metadata directories (1)

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

IMDIS 2016, Gdańsk, 11-13 October 2016
CDI service for discovery and unified data access

SeaDataNet portal

European data sources
data centres $\approx$ 600 originators

Search and Shop

Metadata + transaction data

Already 102 data centres connected and more underway

IMDIS 2016, Gdańsk, 11-13 October 2016
SeaDataNet metadata directories (2)

- **Organisations (3747)**
- **Data sets (4 076)**
- **Projects (2 975)**
- **Observing programmes (362)**
- **Research cruises (47 770)**
- **Data index (1.87 M)**

*End of SDN2 project*
1.87 million CDI entries from 34 countries, 102 data centres and 597 originators for physics, chemistry, geology, geophysics, bathymetry and biology; from 1805 to 2016; 86% unrestricted or under SDN License
SeaDataNet products

- CENTRAL CDI
- Analysis of data anomalies
- Data harvesting
- QC analysis
- File and parameter aggregation

SeaDataNet Quality Checks Strategy (QCS)

Regional products

Aggregated datasets and climatologies

Improvement of the data quality

sdn-userdesk@seadatanet.org – www.seadatanet.org
SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

NODCs; HOs; GEOs; BIOs; ICES; PANGAEA

> 100 data centres

Black Sea portal

Caspian portal

Geo-Seas portal

GEOSS portal

IODE ODP portal

Aggregated collection

Regional subsets

Thematic portals

Data discovery and access

≈ 600 European data originators

CDI Data Discovery and Access service
**SeaDataCloud – a new opportunity**

- Standards and information technology are always evolving, and the SeaDataNet infrastructure must stay up-to-date to maintain and further expand its services.
- The EC has just accepted **SeaDataCloud** proposal as a new EU H2020 project for further developing SeaDataNet infrastructure and associated standards.
  - 10 M euros, consortium of 61 members, 32 countries
  - 4 year project. Start date November 1st 2016.
SeaDataCloud – some planned activities (1)

- Improve services to user and data providers
  - Utilise the benefits of a cloud environment with high performance computing to improve the performance of the CDI data access services
  - Develop online services to visualise and process data, in order to preview, subset, format, or analyse data of interest
  - Develop a Virtual Research Environment (VRE) to facilitate collaborative and individual research by users
  - Provide customized services (MySeaDataCloud services) to let users have his/her search profile, receive alerts on new available data, ingest and manage their own data sets
  - Scientific committee representing lead costumers (Copernicus CMEMS, ARGO, SOCAT, …) will contribute to the specification and definition of tools and services
SeaDataCloud – some planned activities (2)

- **Optimise connecting data centres and data streams to the infrastructure**
  - Ease connecting data centres to the SeaDataNet infrastructure by revising and upgrading the existing components ➔ Provide an integrated package for the new data centres by mean of a virtual application containing all necessary software and operating system
  - Facilitate connecting and ingesting data streams from operational observation networks using OGC Sensor Web Enablement (SWE) standards (in collaboration with ODIP Prototype 3 and other projects)

- **Improve interoperability with other European and International networks**
Present SeaDataNet architecture

Planned upgraded architecture with data replication, advance services and VRE in the cloud
Thank you for your attention