Pan–European infrastructure for marine and ocean data management

Dick M.A. Schaap – MARIS (SeaDataNet Technical Coordinator)

Optimising FAIRness of federated Blue Data Infrastructures webinar
6 Dec 2023

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

What is SeaDataNet?

A pan–European infrastructure, initiated and set up by the NODCs and marine data focal points of 34 countries bordering the European seas

90s	Metadata directories Medar/MedAtlas
1998–2001	EuroNODIM (FP3)
2002–2005	Sea–Search (FP5)
2006–2011	SeaDataNet (FP6)
2011–2015	SeaDataNet II (FP7)
2016–2021	SeaDataCloud (H2020)

SeaDataNet AISBL since 2019

- Developing and maintaining of standards and associated tools, services, and guidance for metadata and data formats, and controlled vocabularies for handling many data types and disciplines, deploying FAIR and INSPIRE principles

- Providing training and support to data centres for uptake of standards, tools, and services in their operation

- Developing skills for uptake of emerging technologies and principles

- Developing and publishing integrated data products such as T&S climatologies

- Being a major player in the European ocean and marine data management landscape supporting EU initiatives like EMODnet, CMEMS, Blue-Cloud, and DTO and working together with several Research Infrastructures (RIs)

European Directory services

EDMERP
Projects

EDMO
Organisations

CSR
Cruises

EDIOS
Observatories

CDI
Data

EDMED
Data sets
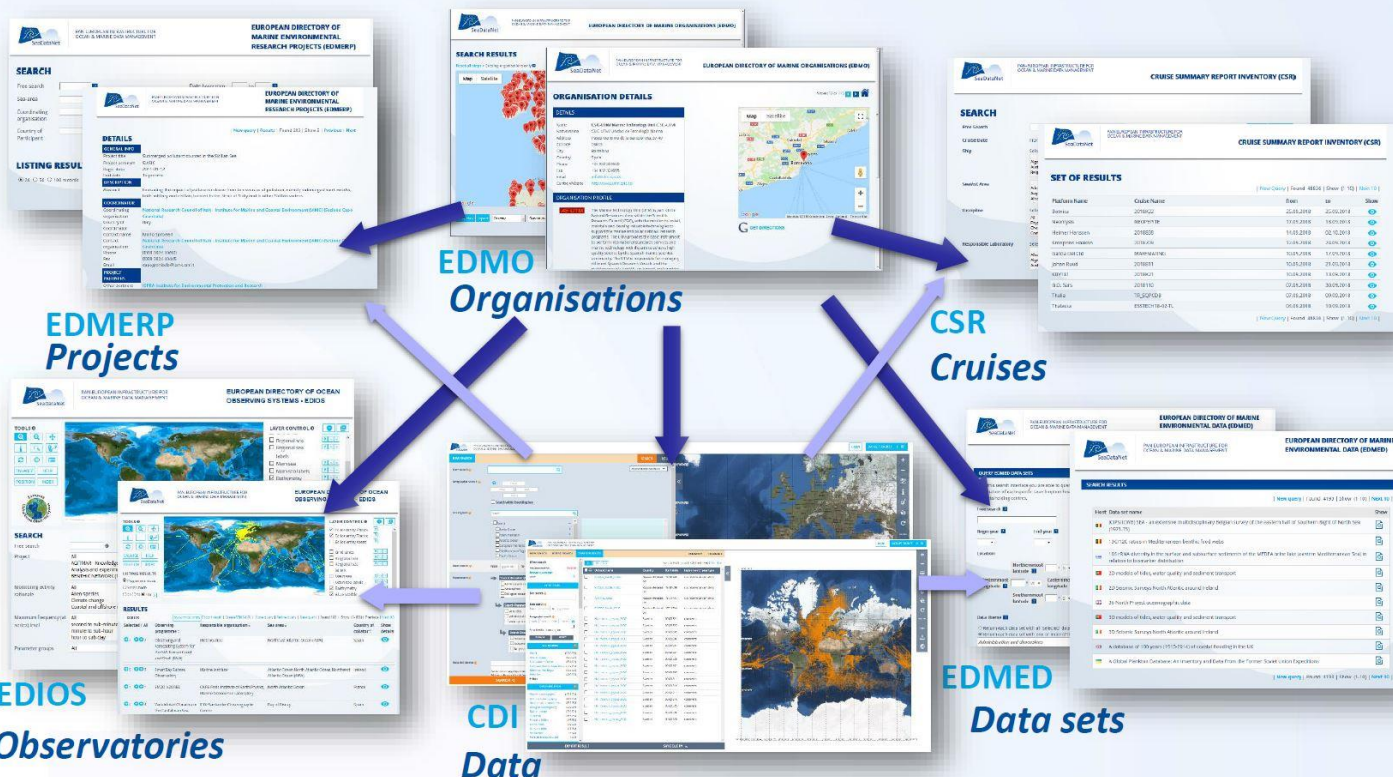
- *User Interfaces*
- *Machine-to-Machine services:*
  - *SparQL*
  - *SOAP web services*
  - *API's*
  - *Linked Data Principle*

Maintaining and publishing a series of Pan–European directories

SeaDataNet standards

"Making Data and

Services:

- **F**indable

- **A**ccessible

- **I**nteroperable

- **R**e-usable

for machines and

people."

- **Common standards for the marine domain**, adapting ISO and OGC standards and achieving INSPIRE compliance:

  - **Metadata formats for data sets, research cruises, monitoring networks, organisations, and research projects**

  - **Standard data exchange formats : ODV ASCII and NetCDF (CF),** fully supported by controlled vocabularies

  - **Controlled Vocabularies** for the marine domain (>90.000 terms in 110+ lists), with international governance and web services

Maintenance and dissemination of standard QA-QC procedures, together with IOC/IODE and ICES

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

CDI Data Discovery & Access Service

SeaDataNet

Cooperation with EUDAT, European e-infrastructure of academic computing centres

EUDAT

As part of:

SeaDataCloud

SeaDataNet portal

Cloud Data Cache

Already 110 data centres connected and more underway
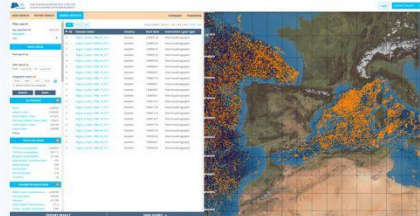
Search and Shop

Data download

Metadata + transaction data

European data sources
data centres holding data of ≈ 850 data originators

https://cdi.seadatanet.org/search

Providing harmonized discovery and access to marine and ocean data sets for physics, chemistry, geology, bathymetry, biology, and geophysics

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

CDI Data Discovery & Access Service

SeaDataNet
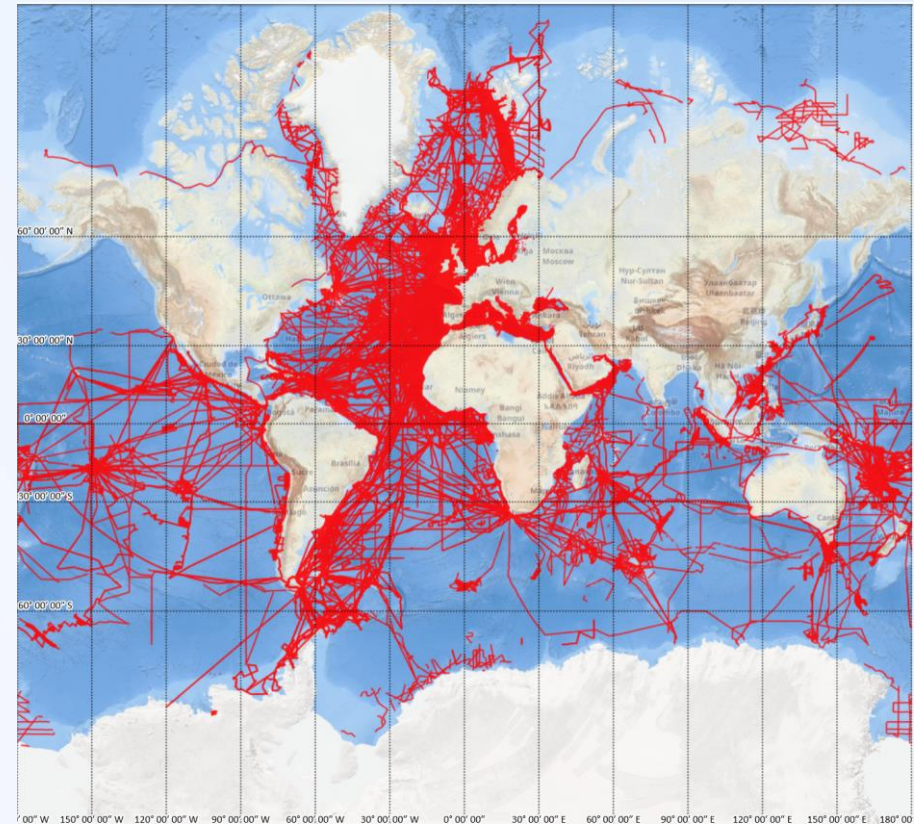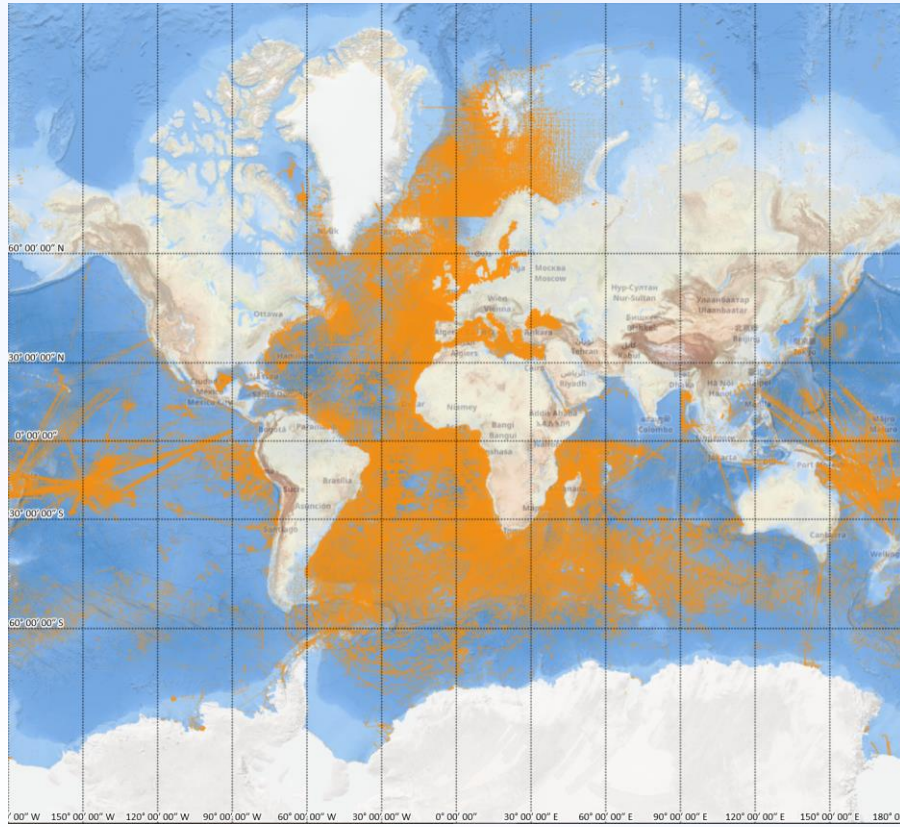
December 2023: more than 2.9 million CDI entries for physics, chemistry, biology, geology, bathymetry, and geophysics, from 117 data centres, located around the European seas, and 970 data originators.

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

https://cdi.seadatanet.org/search

CDI service user interface

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

Cooperation with

- The overarching **European Marine Observation and Data Network (EMODnet) was initiated in 2008 by EU DG MARE**
- SeaDataNet qualified as a leading infrastructure for the EMODnet data management component and is driving several thematic portals from the start
- 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
- EMODnet has a focus on **European data products and services** in support of Blue Economy, Blue Environment and Marine Knowledge 2020 agendas
- The data sets as gathered, harmonised and delivered by SeaDataNet provide essential input for generating and regularly updating EMODnet data products

Cooperation with

One central map viewer to visualise all EMODnet data

1 OCEAN 1 EMODnet

One single portal

120 partners

+100 use cases

One central metadata catalogue to enhance data search and discovery

Discover, visualise and download marine data and products across 7 thematics and hundreds of parameters

BATHYMETRY

HUMAN ACTIVITIES

PHYSICS

GEOLOGY

SEABED HABITATS

CHEMISTRY

BIOLOGY

EMODNET.EC.EUROPA.EU

YOUR GATEWAY TO *IN SITU* MARINE DATA IN EUROPE AND BEYOND

https://emodnet.ec.europa.eu/en

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

Feeding data to several EMODnet Thematics

Thematic subsets → Bathymetry

Thematic subsets → Physics

Thematic subsets → Chemistry

Thematic subsets → Geology

Thematic subsets → Biology

Data discovery
And access

> 110 data centres

NODCs; HOs; GEOs; BIOs; ICES;

> 950 European data originators

EMODnet
European Marine
Observation and
Data Network
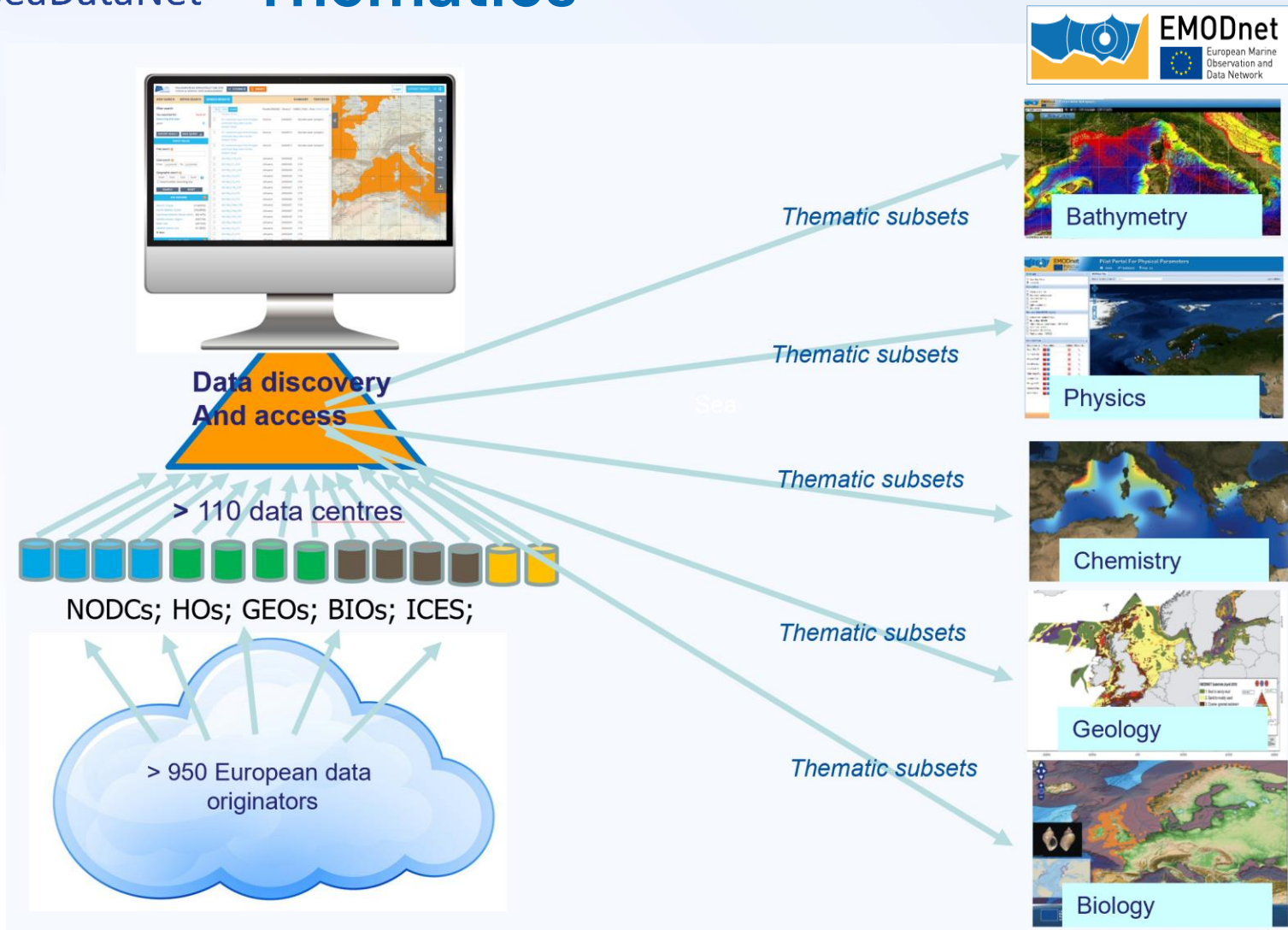
Example: EMODnet Bathymetry

Source reference layer with direct links to CDI service for metadata about used data

The best Digital Terrain Model for European seas:
* Resolution 115 * 115 meters
* Based upon > 17.000 survey and SDB data sets
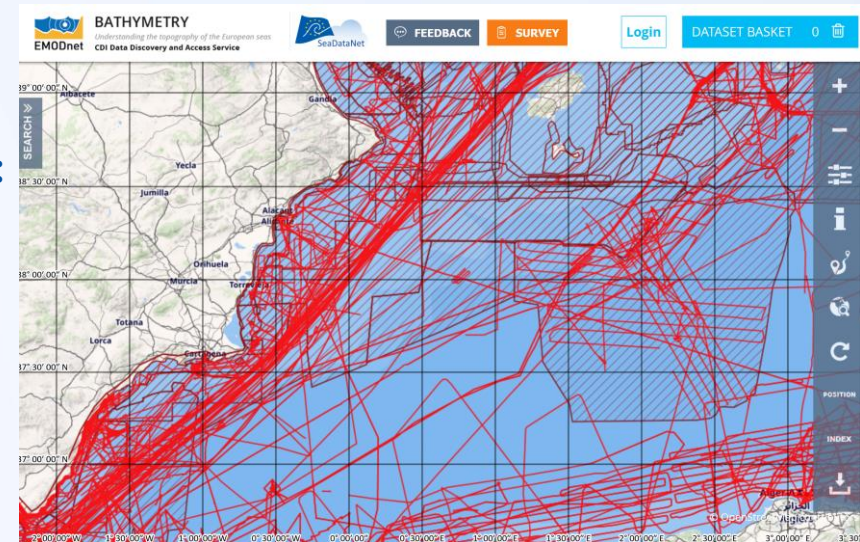
SeaDataNet CDI metadata format

Abstract metadata model specification
- ✓ SeaDataNet metadata profile of ISO 19115 documentation

XML encoding implementation:
- ✓ Schema definition
- ✓ Schematron rules
- ✓ Sample metadata
- ✓ XML implementation documentation
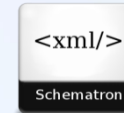
- **SeaDataNet Controlled Vocabularies for marking up metadata**:
  - Currently, 115 lists with > 90000 terms
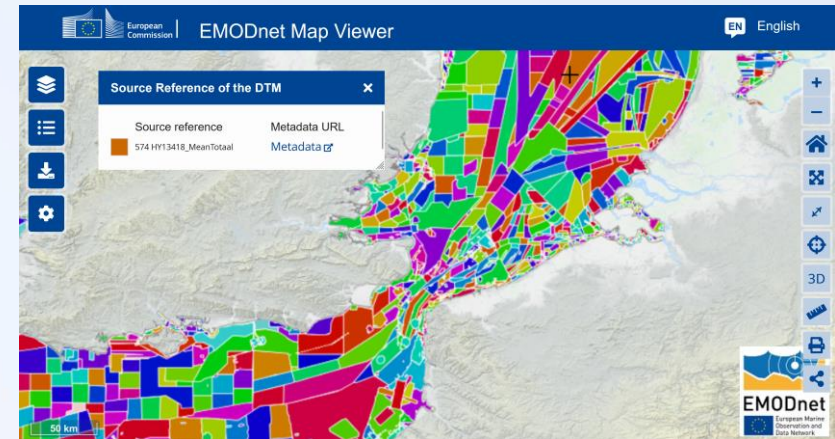 - Available as Web Services, SparQL endpoint, User Interfaces, and with PO1 Vocabulary builder and decomposer
- **SeaDataNet European Directories** for marking up metadata and following 'linked data' principles
  - EDMO, EDMED, CSR, EDMERP

SeaDataNet fit for handling different data types
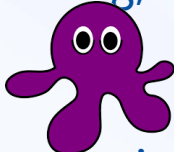
- **SeaDataNet data formats for:**
  - **Physical** data sets, developed with NODCs
  - **Chemistry** datasets, developed with EMODnet Chemistry
  - **Biological** data sets, developed with EurOBIS
  - **Geological and geophysical** data sets, developed with EuroGeoSurveys
 - **Bathymetry** data sets, developed with EMODnet Bathymetry
  - **HF-Radar** data sets, developed with EuroGOOS and EMODnet Physics
  - **Glider** data sets, developed with Ocean Glider network
  - **Flow Cytometry** data together with CNRS and JERICO
  - **Marine Litter** data (beach, seafloor, and micro litter), developed with EMODnet Chemistry and TG ML

- **SeaDataNet Controlled Vocabularies** expanded with new lists and new terms:
  - Currently, 115 lists with > 90000 terms
  - Available as Web Services, SparQL endpoint, User Interfaces, and with P01 Vocabulary builder and decomposer

CDI service architecture

- **Local software tools** at data centres to prepare CDI metadata and data ingestions

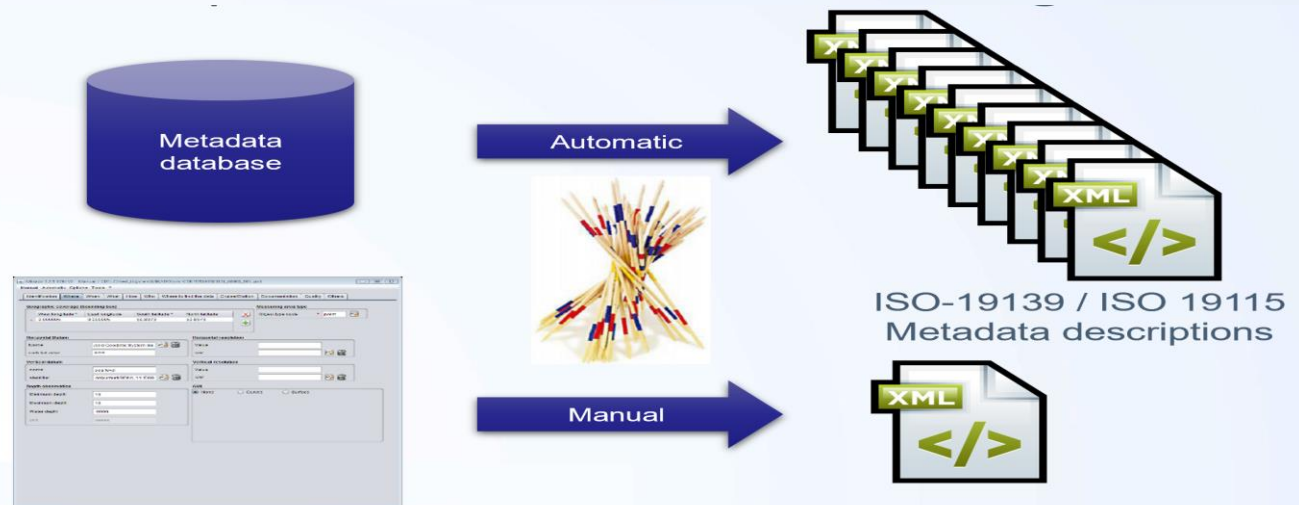- **Replication Manager** (RM) at data centres for transfer entries from data centres to central CDI catalogue and EUDAT data cloud

- **EUDAT cloud** with adapted EUDAT services to store unrestricted data sets

- **CDI User Interface** with central CDI metadata catalogue and facilities for ordering and downloading data sets
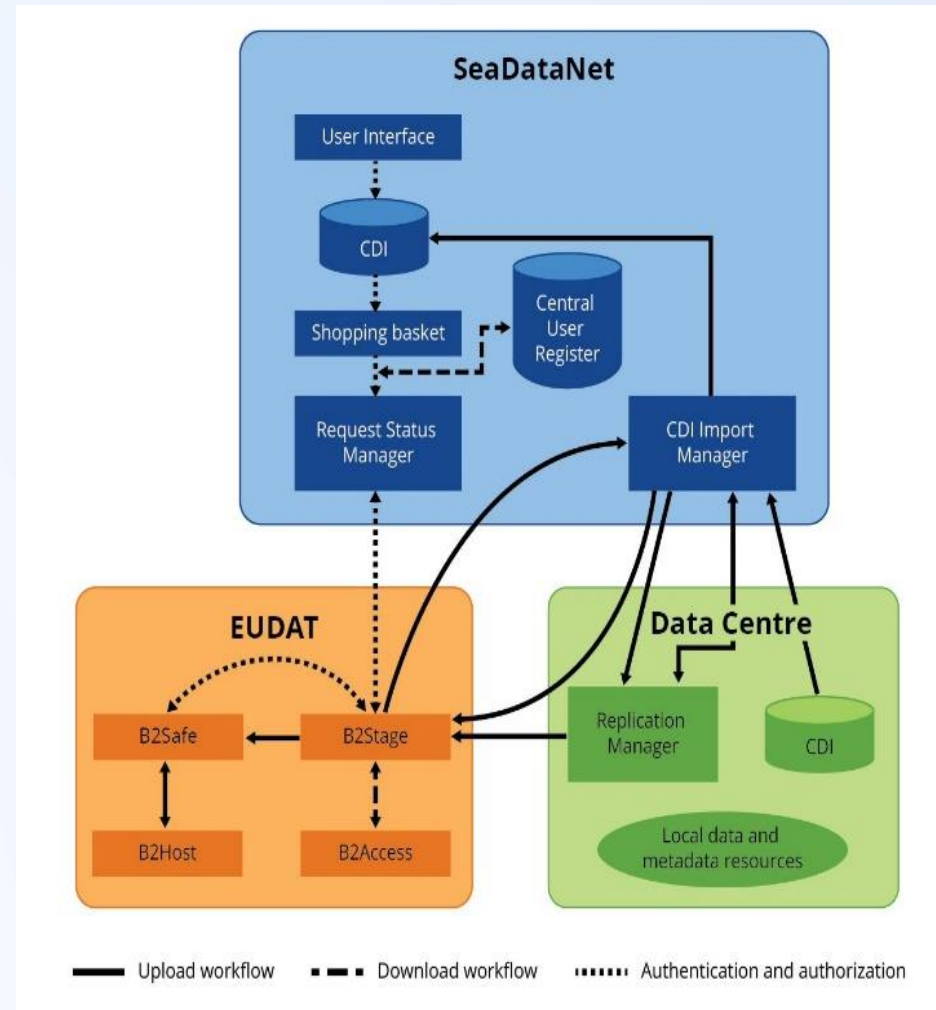
Local software tools at data centres

- Tool for generating XML entries (MIKADO)

- Tool for converting any ASCII data set to SDN ODV + NetCDF formats (NEMO)

- Tool for format checking, convert between SDN formats and handle multi-files (OCTOPUS)

- Vocabulary and Directories web services for marking up metadata

Preparing CDI XML entries

Metadata database

Automatic

Manual

ISO-19139 / ISO 19115
Metadata descriptions

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

Examples of data flow in a data centre

SeaDataNet

Example 1

Octopus

Files format X → NEMO ifremer → Files SDN format No QC → **QC** → Files SDN format QC → Local DB → **Metadata**

Coupling table

Mikado

Vocabularies

SDN Dir

RM

XML

Import Manager

DATA CLOUD

CDI central
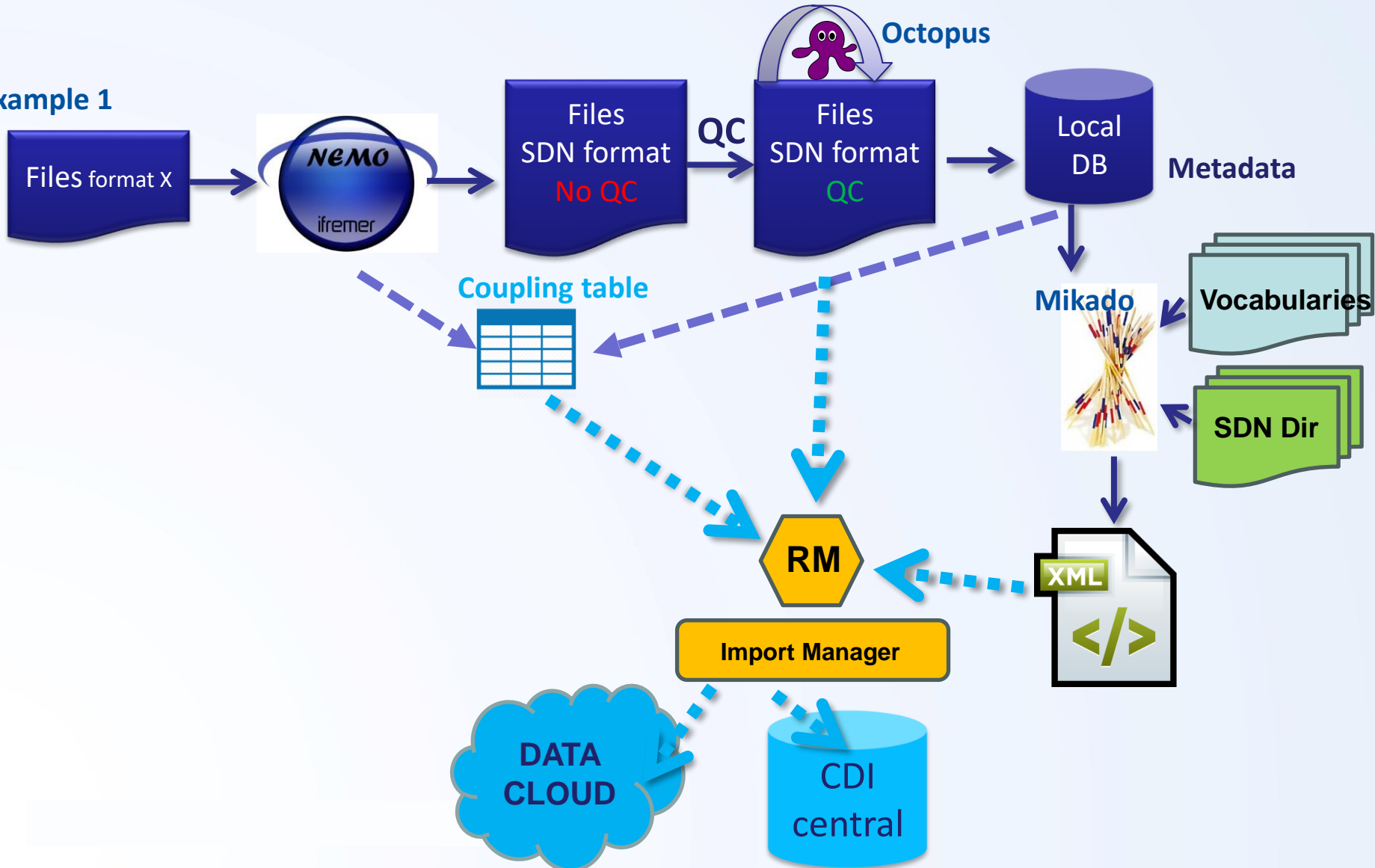
Examples of data flow in a data centre
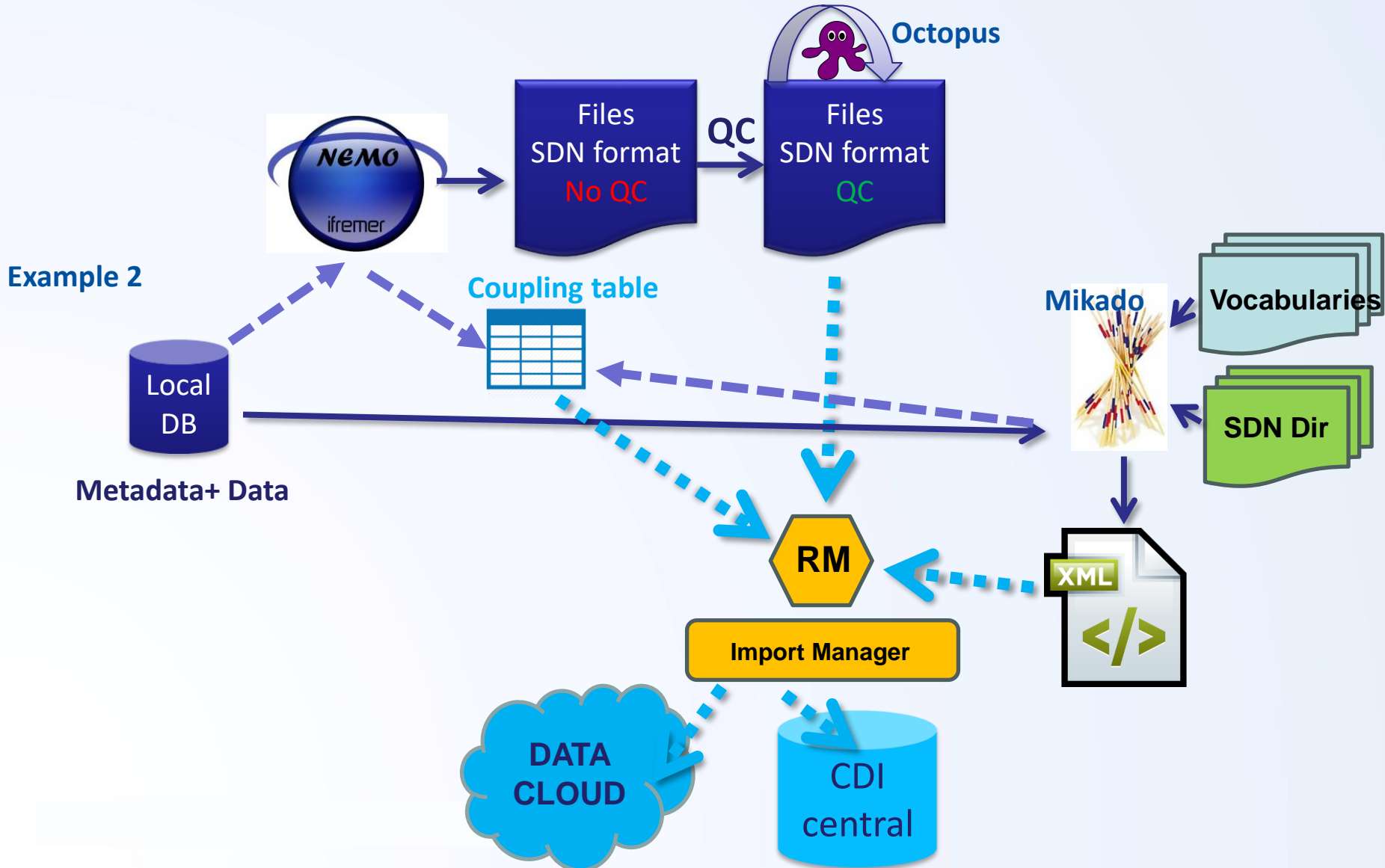
SeaDataNet

Conclusions

- SeaDataNet develops and maintains standards and services for marine data management which are widely adopted, and which have a major impact on the functioning of the marine data landscape in Europe and beyond

- SeaDataNet has a large network of data centres which have direct contacts with hundreds of in-situ data collectors, providing them services for validation, long term stewardship, and wider dissemination of data sets

- SeaDataNet data collections are input for many derived products, for instance for several European data products as generated and published by EMODnet and CMEMS

- SeaDataNet continues to innovate its standards and services to stay state-of-the-art for tomorrows needs. This is done by participating in many EU funded projects.