

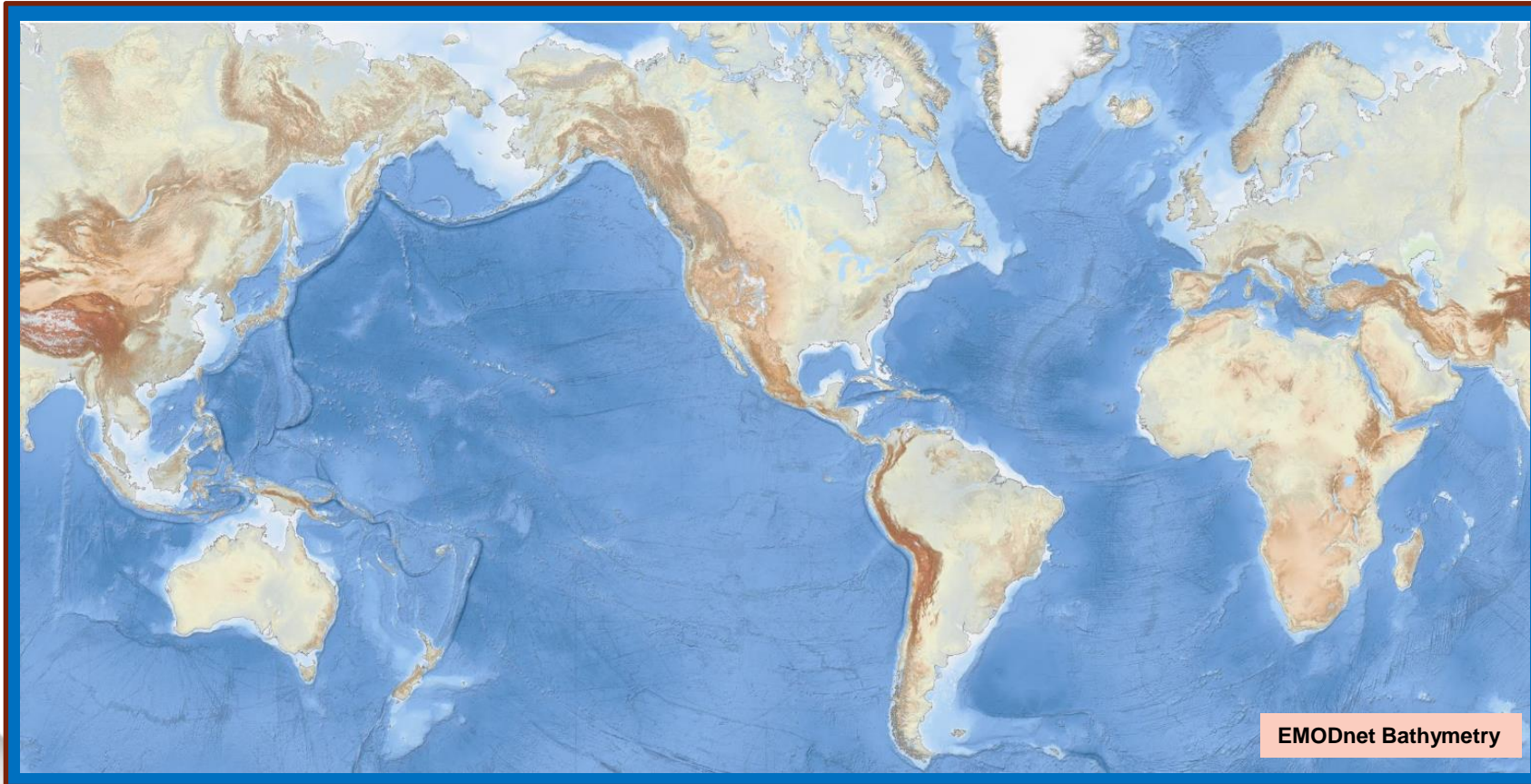
European marine and ocean data management and how to handle data from scientific cruises in Eurofleets+

Dick M.A. Schaap





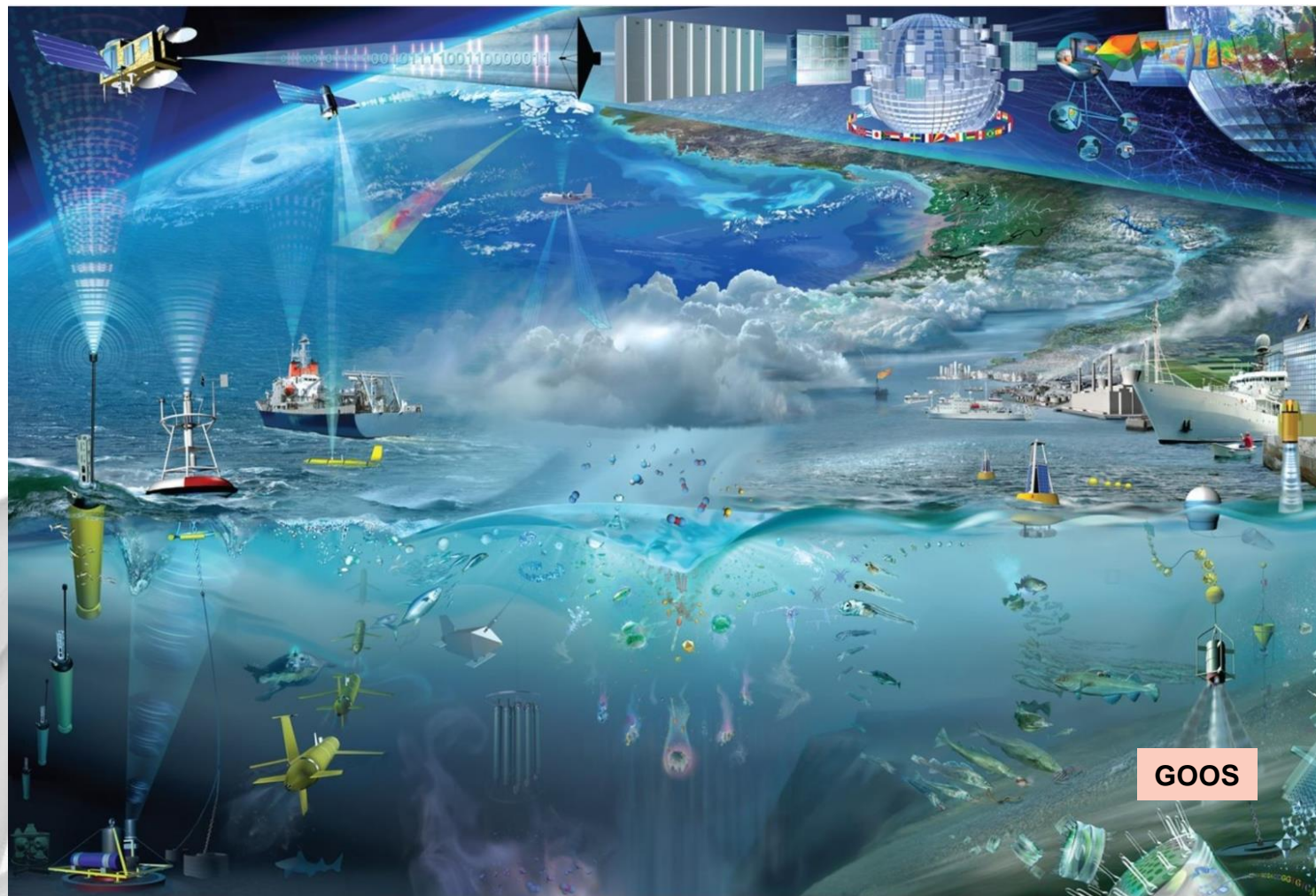
Oceans and seas are important



Climate, Energy, Food, Tourism, Trade, Health,



Acquisition of marine and ocean data

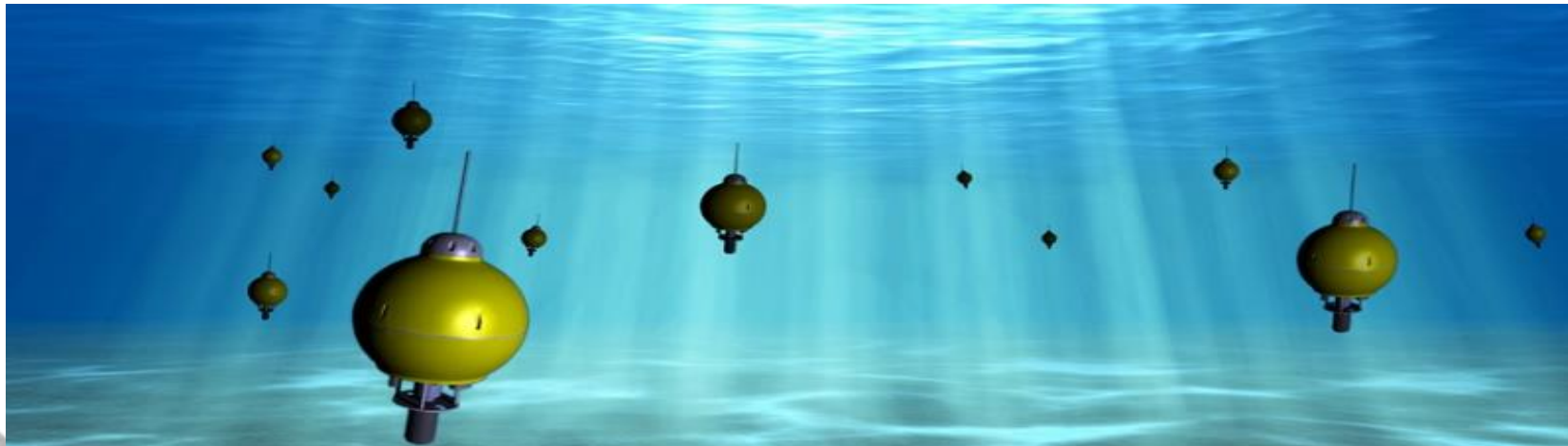


- Scientific Research to gain knowledge and insight
- Modelling (including hindcast, nowcast, forecast)
- Economic activities: shipping, offshore industry, dredging industry, fisheries, tourism, engineering ..
- Environmental Management: monitoring and assessment (water quality, climate status, stock assessment)
- Marine Conventions and Directives, in Europe: Water Framework Directive (WFD), Marine Strategy (MSFD), Marine Spatial Planning (MSP), Coastal Zone Management
- EU Strategies, such as Green Deal, Blue Environment, Blue Economy



Economy of data acquisition

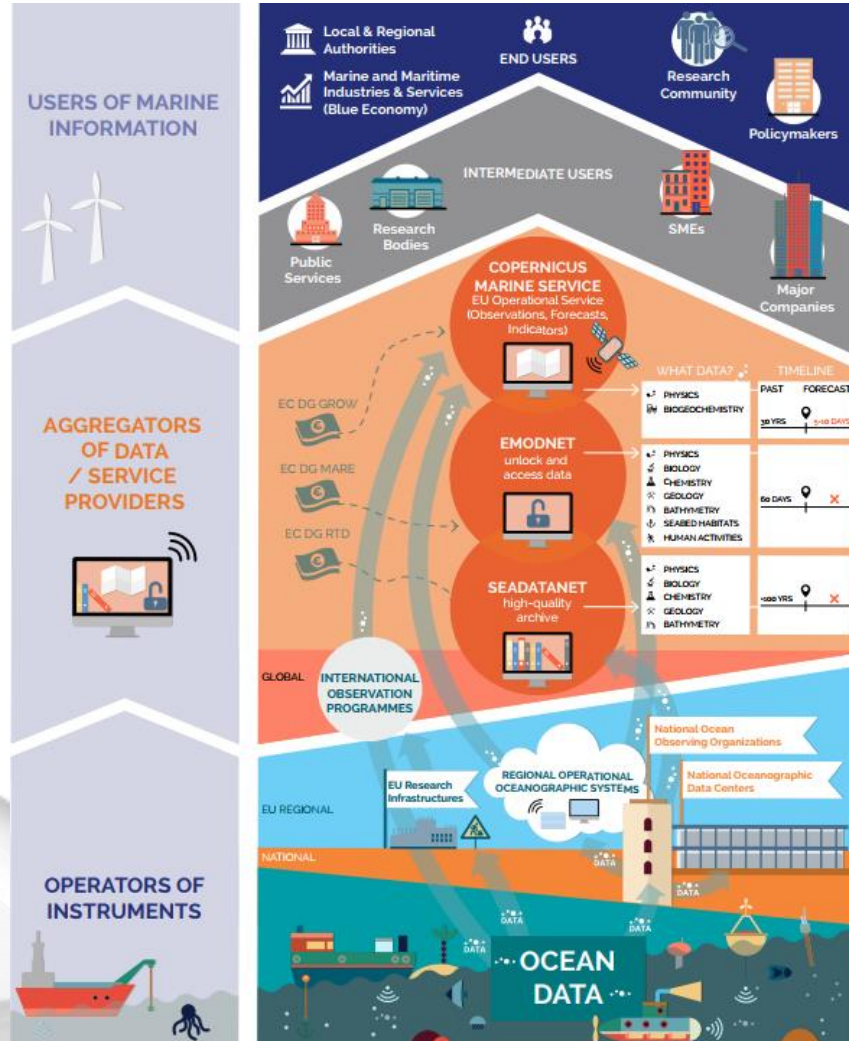
- Data are collected by governments, research institutes, and private industry: in Europe already more than several thousands of organisations)
- Data for physics, geophysics, meteorology, chemistry, biology, geology, bathymetry
- Acquisition of oceanographic and marine data is expensive; annual costs in Europe estimated at **1.4 Billion Euro** (1.0 = in-situ; 0.4 = satellites)



Professional data management is required with agreements on standardisation, quality control protocols, long term archiving, catalogues, and access



European landscape of marine data management



Data aggregators and providers of data products and services



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 technological 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, and EOSC and working together with several Research Infrastructures (RIs), also in the Blue-Cloud



8



SeaDataNet standards

“Making Data and Services:

- ***Findable***
- ***Accessible***
- ***Interoperable***
- ***Re-usable***

for machines
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



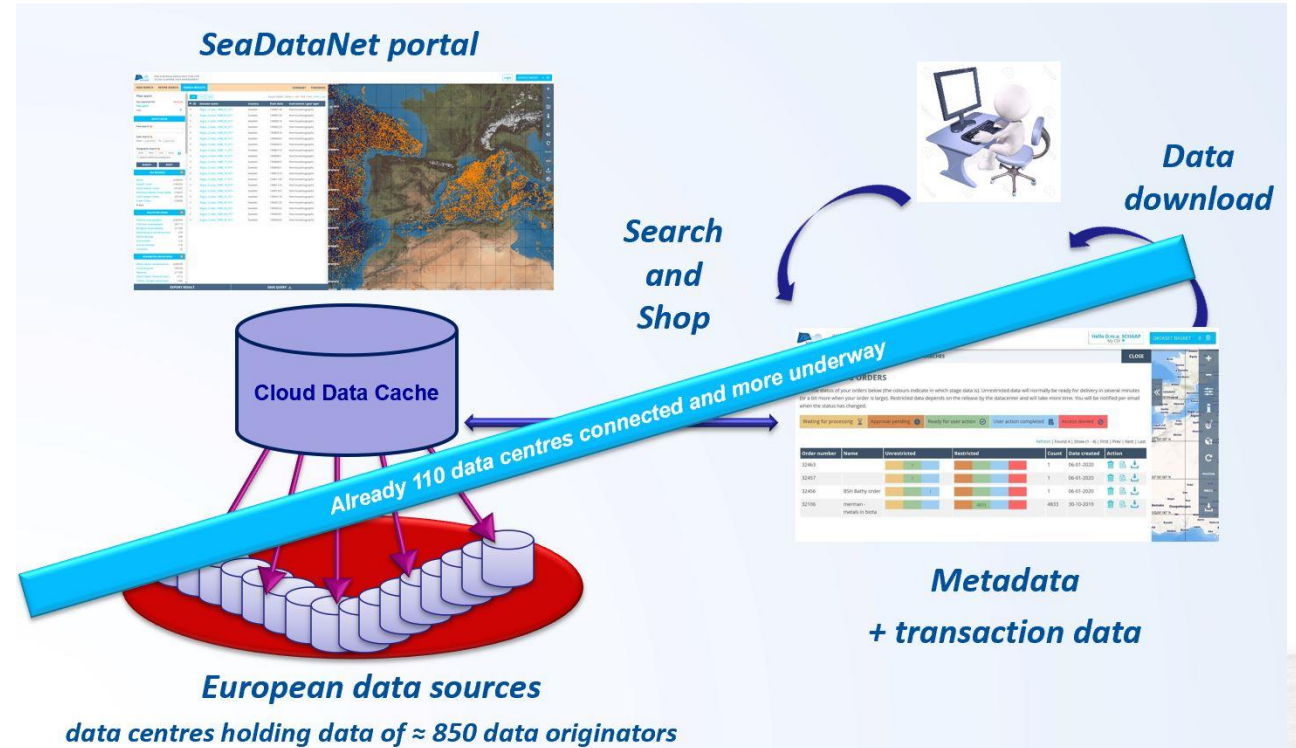


CDI Data Discovery & Access Service

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



As part of:

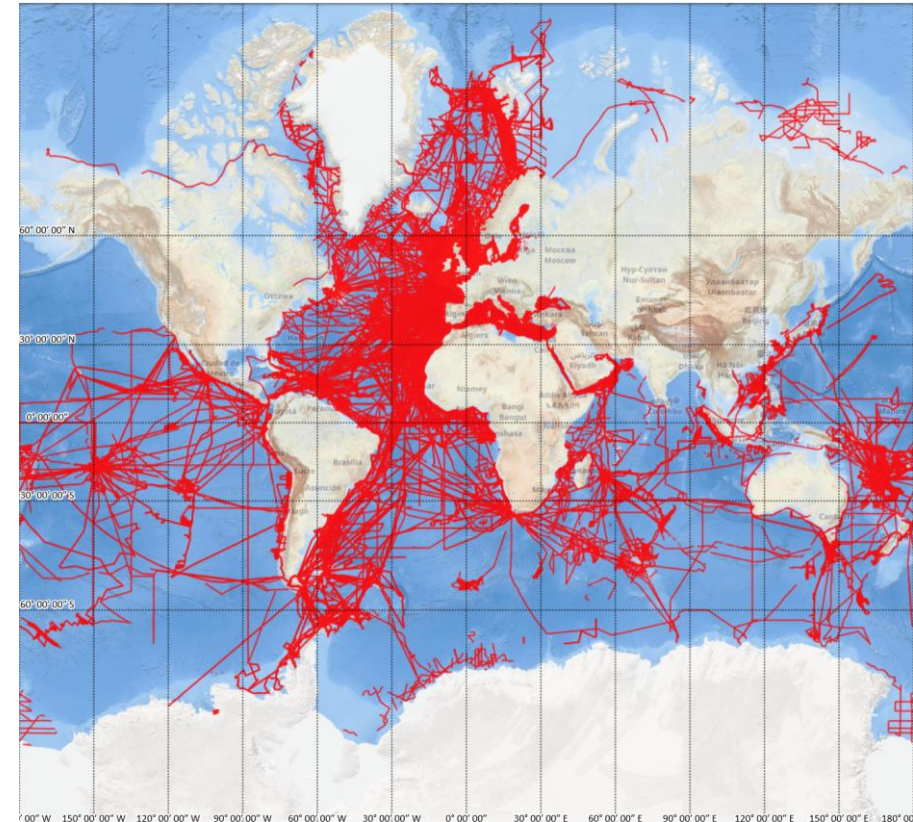
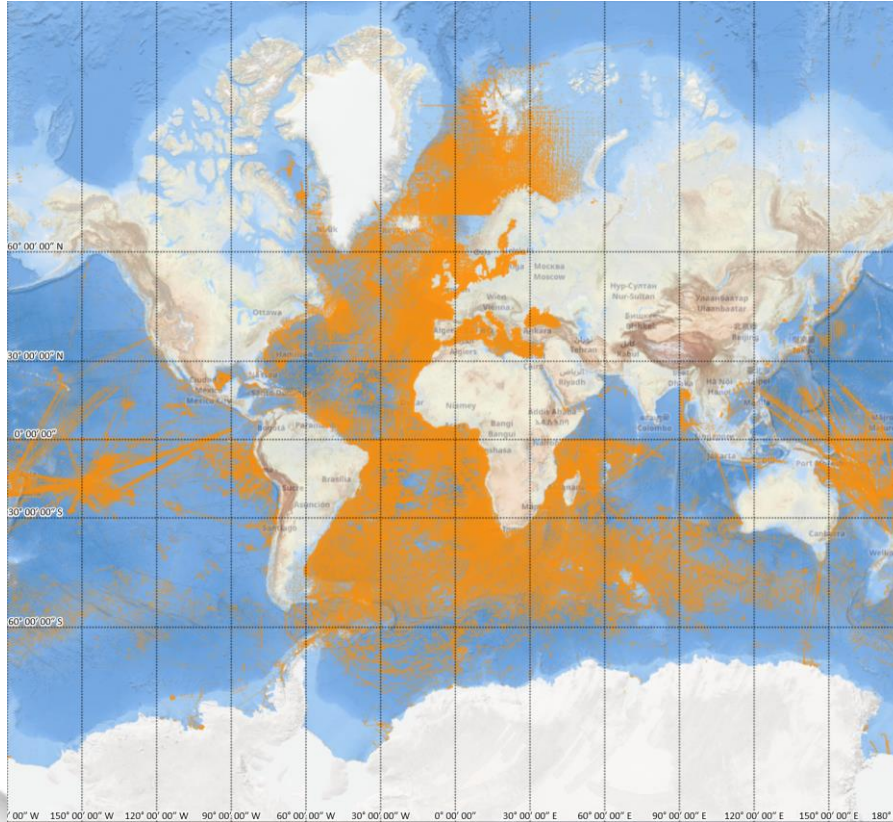


<https://cdi.seadatanet.org/search>

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



CDI Data Discovery & Access Service



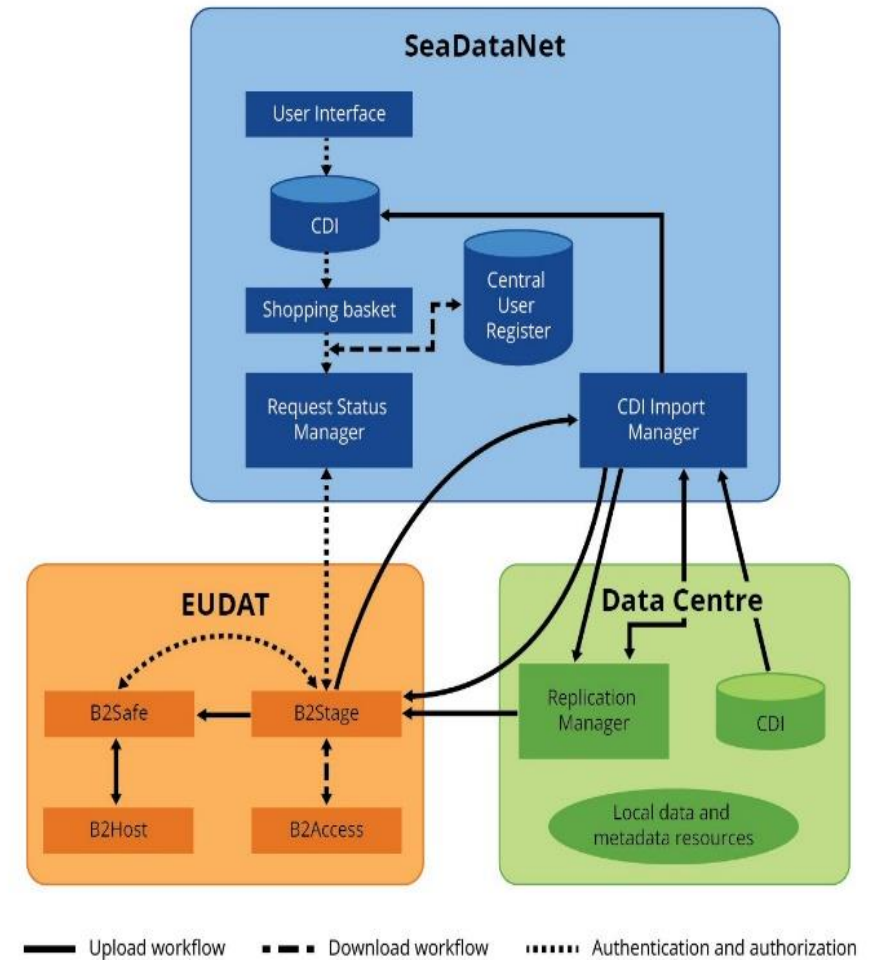
Nov 2022: more than 2.8 million CDI entries for physics, chemistry, biology, geology, bathymetry, and geophysics, from 117 data centres, located around the European seas, and 907 data originators.

<https://cdi.seadatanet.org/search>




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







CDI service user interface



PAN-EUROPEAN INFRASTRUCTURE FOR
OCEAN & MARINE DATA MANAGEMENT

 [FEEDBACK](#)  [SURVEY](#)

[Login](#) [DATASET BASKET 0](#)

NEW SEARCH

REFINE SEARCH

SEARCH RESULTS

SUMMARY

TIMESERIES

Filter search

You searched for:

Instrument Type:

CTD

EXPORT RESULT

SAVE QUERY

INPUT FIELDS

Free search

Date search

From

yyyymmdd

To

yyyymmdd

Geographic search

North

West

East

South

Search within bounding box

SEARCH

RESET

SEA REGIONS

Atlantic Ocean (295552)

North Atlantic Ocean (284862)

Northeast Atlantic Ocean (40... (271726)

Baltic Sea (210188)

Mediterranean Region (156882)

Mediterranean Sea (121244)

+ More

DISCIPLINE (P08)

Physical oceanography (751057)


Administration and dimensi... (320780)

Chemical oceanography (315854)

100 1000 10000

Found 763627 | Show (1 - 100) | First | Prev | Next | Last

	Dataset name	Country originator	Start date	Instrument / gear type
<input type="checkbox"/>	08V1_LT1B_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT1_CTD	Lithuania	20080211	CTD
<input type="checkbox"/>	08V1_LT2C_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT2_CTD	Lithuania	20080211	CTD
<input type="checkbox"/>	08V1_LT3_CTD	Lithuania	20080211	CTD
<input type="checkbox"/>	08V1_LT46_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT4C_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT4_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT5_CTD	Lithuania	20080213	CTD
<input type="checkbox"/>	08V1_LT64A_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT64_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT65_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT66_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LT6B_CTD	Lithuania	20080213	CTD
<input type="checkbox"/>	08V1_LT6_CTD	Lithuania	20080213	CTD
<input type="checkbox"/>	08V1_LT7_CTD	Lithuania	20080213	CTD
<input type="checkbox"/>	08V1_LTB-1_CTD	Lithuania	20080211	CTD
<input type="checkbox"/>	08V1_LTB-4_CTD	Lithuania	20080212	CTD
<input type="checkbox"/>	08V1_LTN-6_CTD	Lithuania	20080213	CTD
<input type="checkbox"/>	08V2_LT1_CTD	Lithuania	20080507	CTD



+

-

Full Screen

Layers

Info

Position


Index

Download


CDI search example for CTD measurements




CDI Data Discovery & Access Service




PAN-EUROPEAN INFRASTRUCTURE FOR
OCEAN & MARINE DATA MANAGEMENT

 **FEEDBACK**

 **SURVEY**


Login


DATASET BASKET 0 

NEW SEARCH **REFINE SEARCH** **SEARCH RESULTS** **SUMMARY** **TIMESERIES**


Filter search


You searched for: **Reset all**

Free search:
tara oceans 


EXPORT RESULT **SAVE QUERY** 


INPUT FIELDS

Free search 

Date search 


From: To:

Geographic search 




☐ Search within bounding box

SEARCH **RESET**

SEA REGIONS 

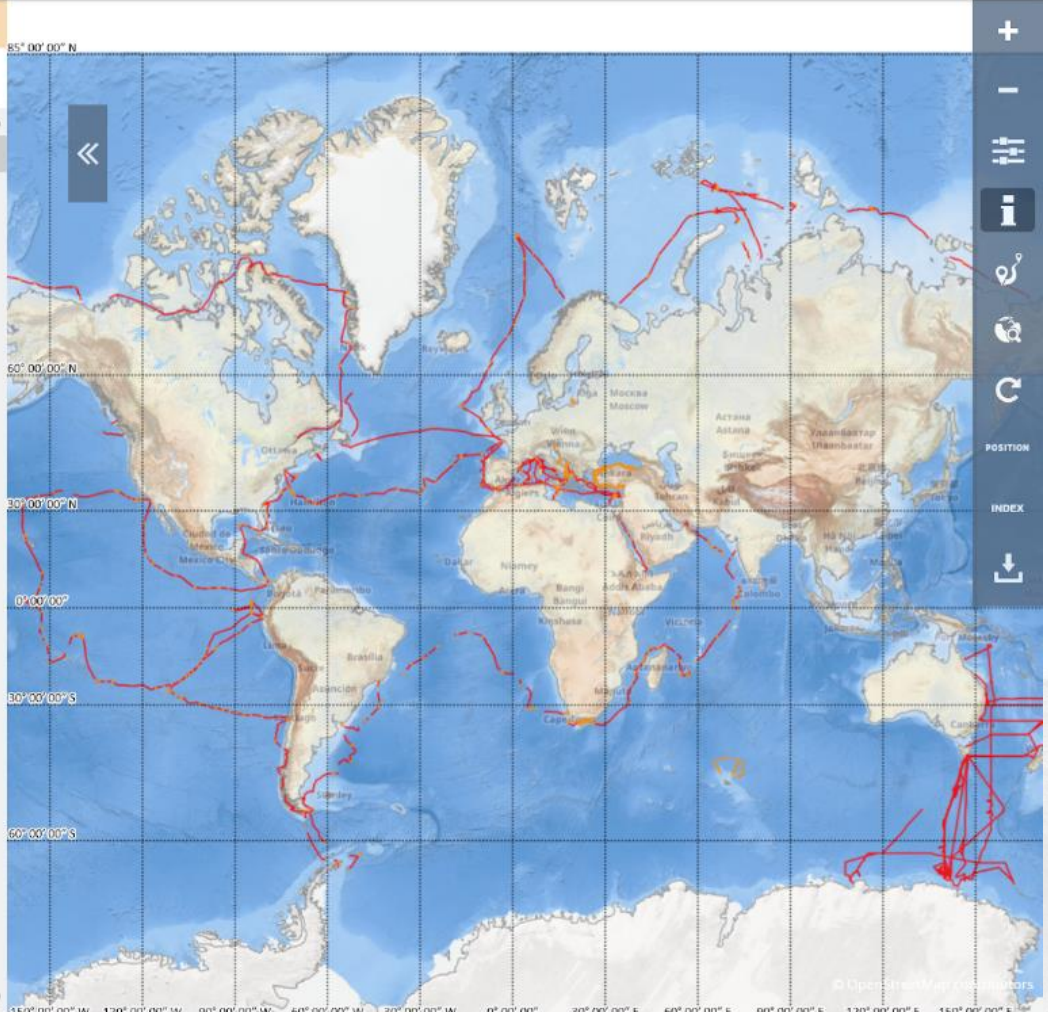
Mediterranean Region (2491)
Mediterranean Sea (2218)
Mediterranean Sea, Eastern ... (2173)
Atlantic Ocean (677)
Pacific Ocean (558)
North Atlantic Ocean (508)
+ More

DISCIPLINE (P08) 

Physical oceanography (4296)
Chemical oceanography (2613)

Found 4386 | Show (2701 - 2800) | [First](#) | [Prev](#) | [Next](#) | [Last](#)

SELECT	<input type="checkbox"/>
DATASET NAME	Environmental parameters measured on discrete water samples during the Tara Oceans expedition 2009-2013.
COUNTRY ORIGINATOR	France
START DATE	20101009
INSTRUMENT / GEAR TYPE	discrete water samplers
SELECT	<input type="checkbox"/>
DATASET NAME	Environmental parameters measured from physical and optical sensors during Tara Oceans expedition 2009-2013.
COUNTRY ORIGINATOR	France
START DATE	20101009
INSTRUMENT / GEAR TYPE	CTD
SELECT	<input type="checkbox"/>
DATASET NAME	Environmental parameters measured on discrete water samples during the Tara Oceans expedition 2009-2013.
COUNTRY ORIGINATOR	France
START DATE	20101009
INSTRUMENT / GEAR TYPE	discrete water samplers
SELECT	<input type="checkbox"/>
DATASET NAME	Environmental parameters measured from physical and optical sensors during Tara Oceans expedition 2009-2013.
COUNTRY ORIGINATOR	France
START DATE	20101012
INSTRUMENT / GEAR TYPE	CTD
SELECT	<input type="checkbox"/>
DATASET NAME	Environmental parameters measured on discrete water samples during the Tara Oceans expedition 2009-2013.
COUNTRY ORIGINATOR	France



CDI search example for Tara-Oceans cruises

<https://cdi.seadatanet.org/search>



Fit for handling many data types

- **SeaDataNet metadata and 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 > 91000 terms
 - Available as Web Services, SparQL endpoint, User Interfaces, and with P01 Vocabulary builder and decomposer



Tools & services provided

- Software tools for generating XML entries (MIKADO), format conversions and format checking (NEMO and OCTOPUS)
- Software tools for analyses (ODV) and interpolations (DIVA)
- Online versions of ODV (WebODV) and DIVA (DIVAnd) as part of SeaDataNet Virtual Research Environment (VRE)
- Sensor Web Enablement (SWE) toolkit for operational data streams
- SEANOE data publishing and DOI minting service
- Vocabulary and Directories web services
- Brokerage service for discovery and access of several international repositories (NCEI-USA, WOD-USA, AODN-Australia)



SeaDataNet cooperation

- **Many research projects:** ENVRI-FAIR, MARINET2, PHIDIAS, E-Shape, EOSC-EGI-ACE, EOSC-Future, AtlantOS,, adopting and adapting SeaDataNet standards and services, and use cases
- **Large ocean monitoring systems:** EuroGOOS, JERICO-S3, EuroFleets+, Euro-Argo, Gliders, ... adopting standards and services for validation + long-term archiving
- **Blue-Cloud project in G7 Future of the Oceans framework:** pillar under Blue-Cloud Data Discovery & Access service, federating SeaDataNet, EMODnet, ELIXIR-ENA, EurOBIS, EcoTaxa, ICOS, SOCAT, and Euro-Argo
- **GEOSS - EuroGEOSS:** populating the GEOSS portal with SeaDataNet in-situ data collections for global sharing
- **UNESCO IOC – IODE network and Ocean Data Portal:** global data exchange and interoperability solutions
- **Copernicus Marine Services (CMS):** providing standards, and cooperation in T&S climatologies

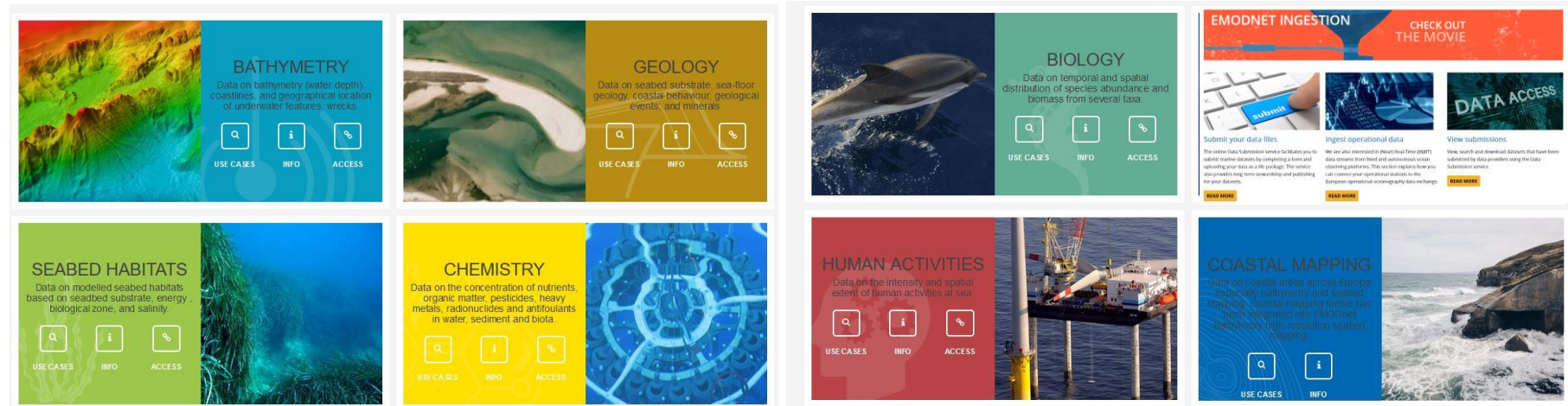
SeaDataNet is the essential data management link between marine and ocean data collectors (research cruises, operational monitoring, fixed network, autonomous floats, ..) and the overarching infrastructures such as EMODnet, CMS, Blue-Cloud, and EOSC



Cooperation with



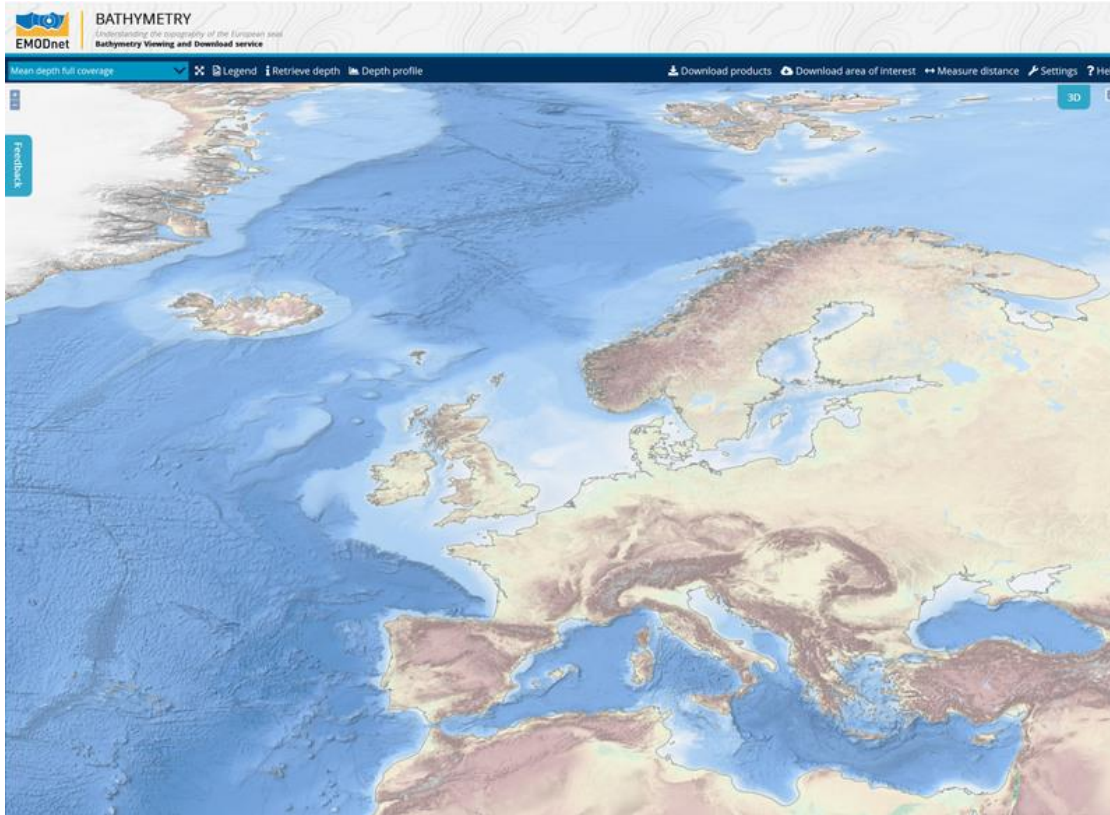
- The overarching **European Marine Observation and Data Network (EMODnet)** was initiated in 2008 by EU DG MARE



- 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

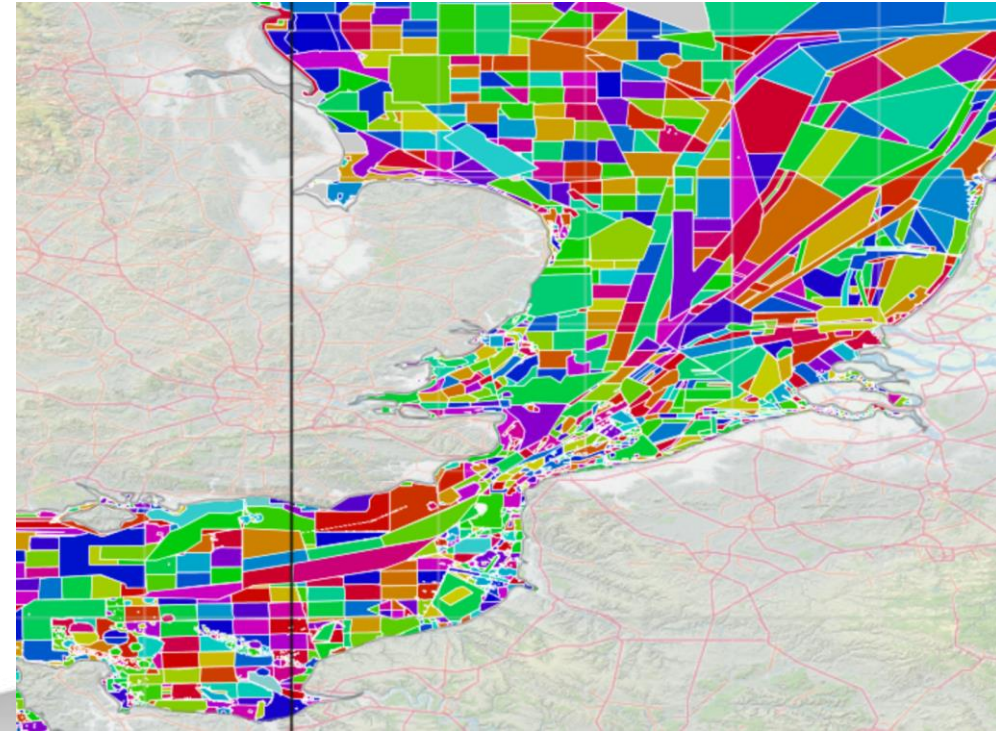


Example: EMODnet Bathymetry



The best Digital Terrain Model for European seas:

- * Resolution 115 * 115 meters
- * Based upon > 16.000 survey and SDB data sets



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

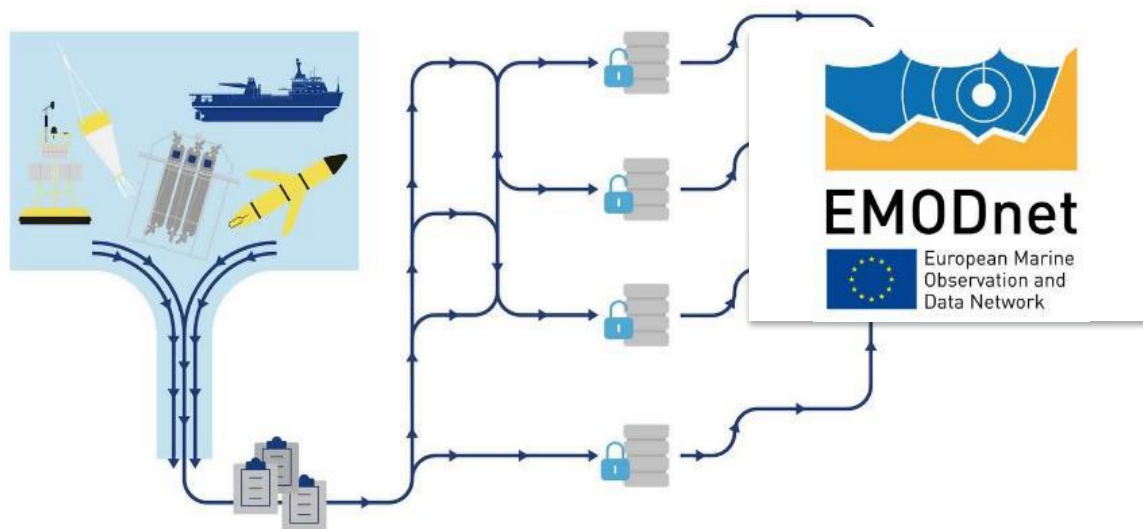


Example: EMODnet Bathymetry





Example: EMODnet Chemistry



- **C**ollection, **A**ggregation, **S**tandardization, **Q**uality check of EU marine water quality data relevant to the EU Marine Directives and to global climate change

Group of Parameters

Parameters

Marine Litter	Beach macrolitter, Seafloor macrolitter, Floating microlitter	Composition, Abundance, etc.
Ocean acidification	Acidity	pH, PCO_2 , etc.
Contaminants	Antifoulants, Hydrocarbons, Heavy metals, Pesticides, Polychlorobiphenyls (PCBs), Radionuclides	Anthracene, Fluoranthene, Me, Cd, Pb, TBT, DDTs, etc.
Eutrophication	Nutrients, dissolved gases, etc.	N, P, Si, Chl-a, O_2 , C, etc.



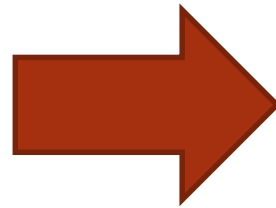


Example: EMODnet Chemistry

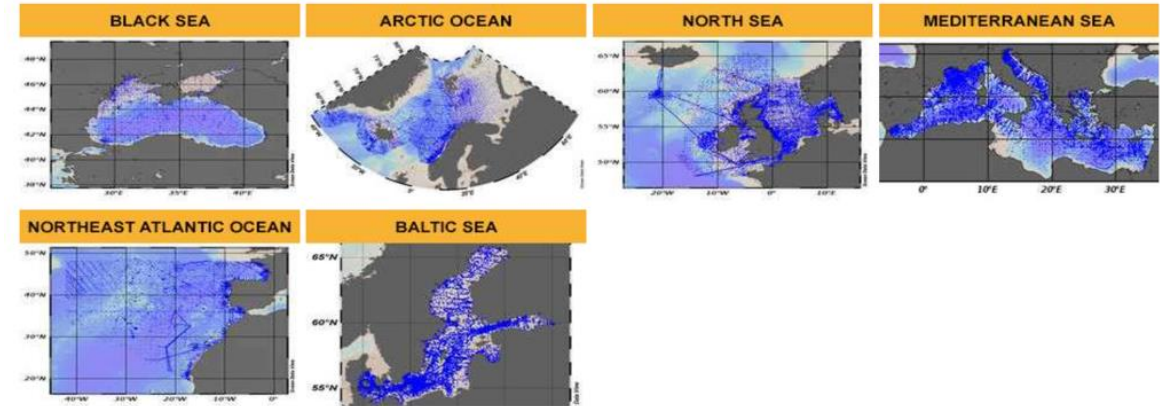
Eutrophication (nutrients, chlorophyll and oxygen) and Ocean Acidification (alkalinity and pH) in seawater



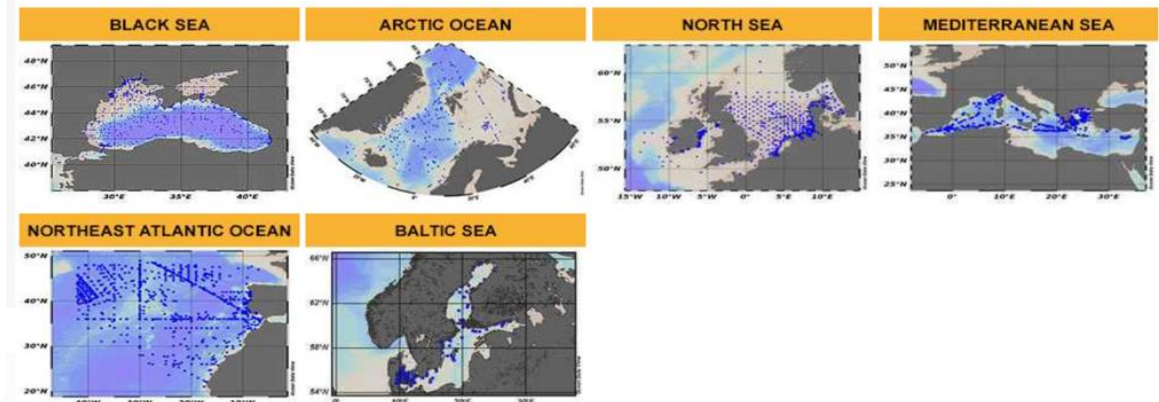
Free and open access to **over 1 million chemistry data sets** in all EU sea basins



QA-QC & Harmonisation by Regional experts



Contaminants in seawater, biota and sediment




Data collections

Data are harmonized, standardized, validated and made available as regional and global data collections.



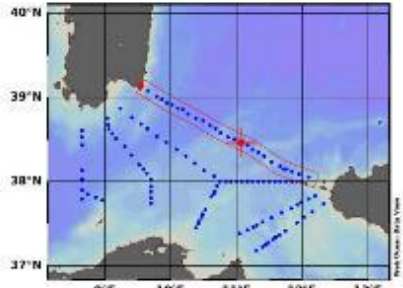
Example: EMODnet Chemistry

**CHEMISTRY**
Data & products on marine water quality

HELPDESK DATA PROTECTION NOTICE

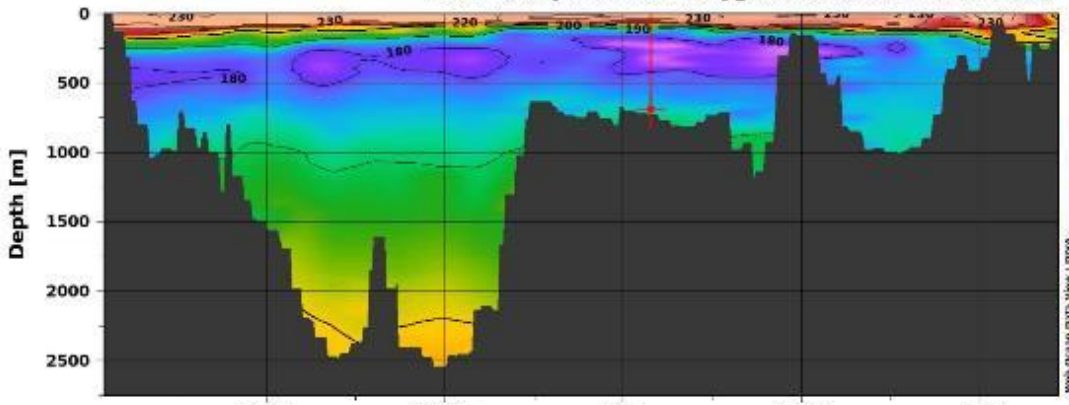
webODV emodnet chemistry > eutrophication > Mediterranean > Mediterranean_Eutrophication_Profiles_DIN_TS.odv

Mediterranean_Eutrophication_Profiles_DIN_TS Collection View



Collection: Mediterranean_Eutrophication_Profiles_DIN_TS User: 77.3.57.153 Date: 2020-11-09T18:12:20 Software: ODV-online v4 | odvres 5.3.5 / 24

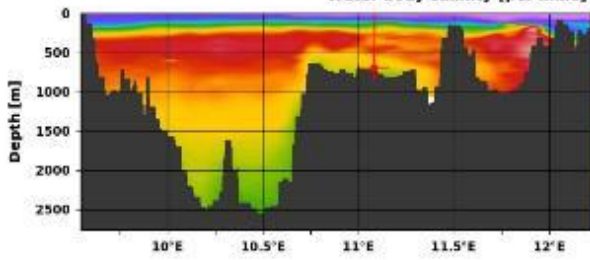
Water body dissolved oxygen concentration [umol/l]



Depth [m] 0 500 1000 1500 2000 2500

Longitude 10°E 10.5°E 11°E 11.5°E 12°E

Water body salinity [per mille]



Depth [m] 0 500 1000 1500 2000 2500

Longitude 10°E 10.5°E 11°E 11.5°E 12°E

Station ID: 67009

Accession Number	67009
Cruise	MTPII-MATER/MA2 JAN9
Station	22300 (C)
Position	11.078°E / 38.467°N
Date	16 January 1997
Time	01:21:58
Depth Range [m]	[2.98 - 834.86]
LOCAL_CDI_ID	321288
Sample: 696 / 841	
1: Depth [m]	691.58
2: ITS-90 water temperature [C]	13.25
3: Water body salinity [per mille]	38.53
4: Water body dissolved oxygen	188.850
5: Water body dissolved oxygen	9
6: Water body nitrate [umol/l]	9
7: Water body nitrate plus nitrite	9
8: Water body nitrite [umol/l]	9
9: Water body phosphate [umol/l]	9
10: Water body silicate [umol/l]	9
11: Water body ammonium [umol/l]	9
12: Water body urea [umol/l]	9
13: Water body chlorophyll-a [C]	9
14: Water body chlorophyll-b [C]	9
15: Water body chlorophyll-c [C]	9
16: Water body phaeopigment	9
17: Water body total alkalinity [C]	9
18: Water body dissolved inorganic carbon dioxide	9
19: Water body carbon dioxide	9
20: Water body total nitrogen [C]	9
21: Water body total phosphorus	9
22: Water body nitrate plus nitrite	9

Isosurface Values

Longitude	11.078
Latitude	38.467
Time [yr]	1997.047
Day of Year	18

Station ID: 34270

Accession Number	34278
Cruise	352
Station	194 (B)
Position	23.8°W / 61.616°N
Date	27 July 1968
Time	15:59:59
Depth Range [m]	[0.00 - 1510.00]
LOCAL_CDI_ID	RNODC_Bottle_352_194
EDMO code	681
Ref. Depth [m]	1520
Sample: 1 / 17	
1: Depth [m]	0.00
2: ITS-90 water temperature [C]	12.50
3: Water body salinity [per mille]	35.19
4: Water body dissolved oxygen	298.55
5: Water body dissolved oxygen	0
6: Water body nitrate [umol/l]	0
7: Water body nitrate plus nitrite	0
8: Water body nitrite [umol/l]	0
9: Water body phosphate [umol/l]	0
10: Water body silicate [umol/l]	0
11: Water body ammonium [umol/l]	0
12: Water body chlorophyll-a [C]	0
13: Water body phaeopigment	0
14: Water body total alkalinity	0

Isosurface Values

Longitude	-23.800
Latitude	61.616
Time [yr]	1968.570
Day of Year	208
Water body dissolved oxygen concentration	275.47
Water body salinity [per mille] @ Depth [m]	35.19
Water body dissolved oxygen concentration	298.55
Water body dissolved oxygen saturation [%]	



Blue-Cloud initiative

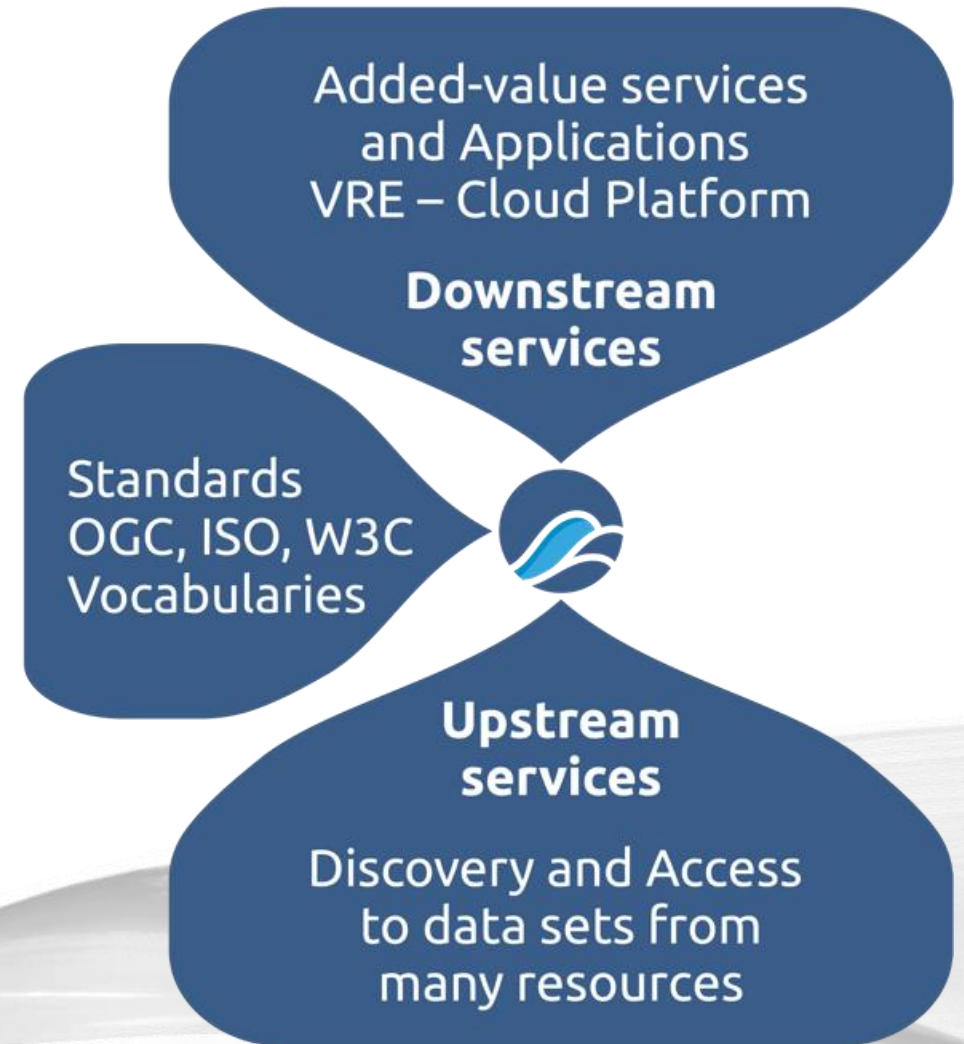
- To explore and demonstrate the potential of **cloud based open science** supporting research for ocean sustainability, and UN Decade of the Oceans and G7 Future of the Oceans
- To deploy a **cyber platform with smart federation** of multidisciplinary data repositories, analytical tools, and computing facilities
- To develop a **marine thematic European Open Science Cloud (EOSC)** serving the blue economy, marine environment & marine knowledge agendas





Blue-Cloud overarching concept

- Developing and deploying Virtual Research Environment (VRE) with an array of services for configuring and running Virtual Labs for specific analytical workflows, use cases, and demonstrators
- Applying common standards and interoperability solutions for providing harmonized metadata and data
- Developing and deploying harmonized discovery and access to established European marine data management and processing infrastructures





Blue-Cloud federation of major infrastructures

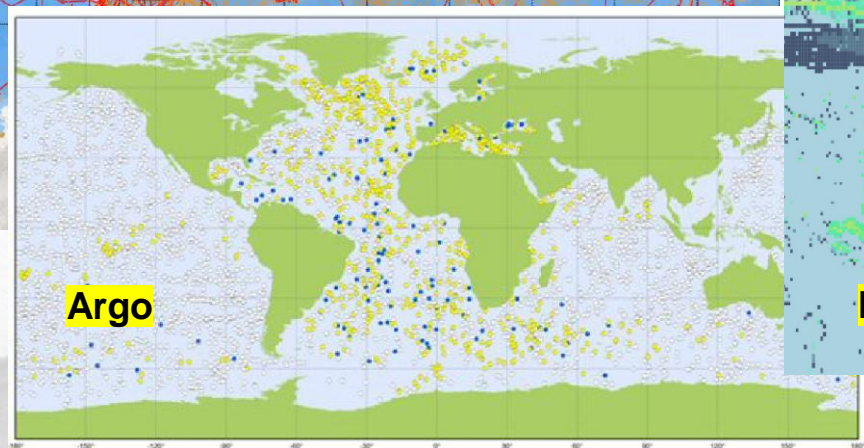
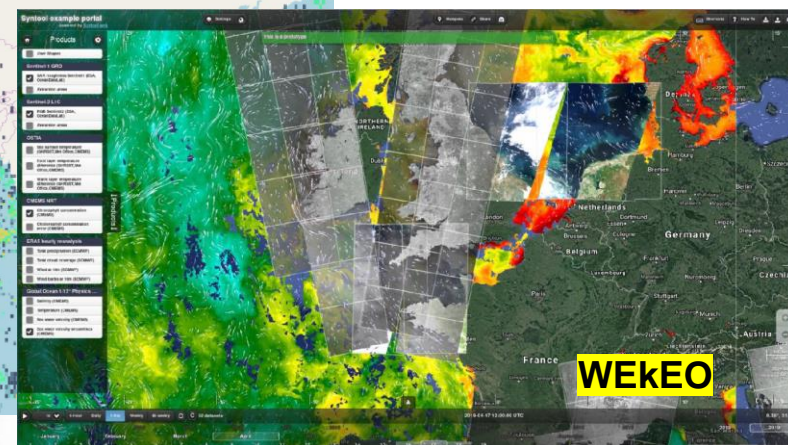
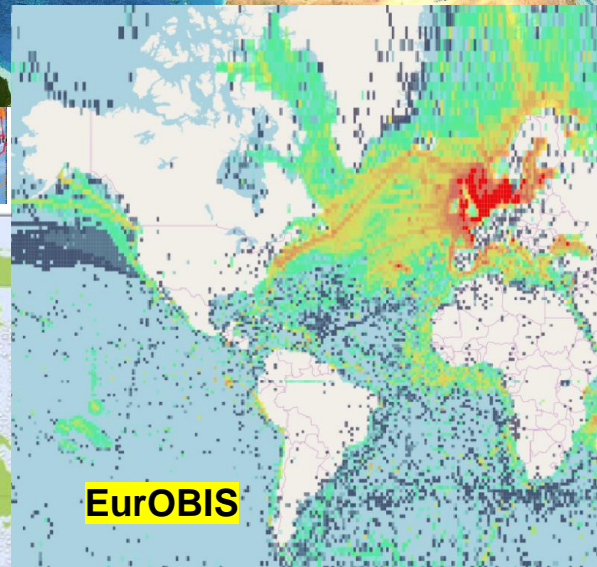
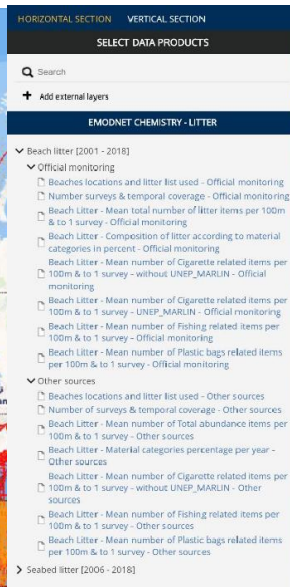


D4SCIENCE
INFRASTRUCTURE



Blue Data infrastructures

E-infrastructures



Key products and services

- **Blue-Cloud Data Discovery & Access service**, federating key European data management infrastructures, to facilitate users in finding and retrieving multi-disciplinary datasets from multiple repositories
- **Blue-Cloud Virtual Research Environment infrastructure** to provide a range of services and to facilitate orchestration of computing and analytical services for constructing, hosting and operating Virtual Labs for specific applications
- **Blue-Cloud Virtual Labs**, configured with specific analytical workflows to serve as **Demonstrators**, which can be adopted and adapted for other inputs and analyses

www.blue-cloud.org



Biodiversity
Zoo and Phytoplankton EOV
products



Genomics
Plankton Genomics



Environment
Marine Environmental
Indicators



Fisheries
Global Record of Stocks and
Fisheries



Aquaculture
Aquaculture Monitor



DATA DISCOVERY & ACCESS SERVICE

The image displays a collage of screenshots of the Blue-Cloud Data Discovery & Access Service interface. The interface is designed for searching and accessing marine data. Key components include:

- Header:** Blue-Cloud logo, "DATA DISCOVERY & ACCESS SERVICE", user name "WELCOME DICK M.A. SCHAAAP", and a "DATASET BASKET 0" indicator.
- Search Filters:** A sidebar with "Filter search" options including "Free search", "Date search" (From, To), and "Geographic search" (North, South, West, East).
- Search Results:** A central area showing "Results found: 296 in 0.07 Seconds" and a "VIEW RESULTS" button. Below this, a list of search results is displayed, including "SOLO" and "EuroArgo - Argo".
- Map View:** A map showing the search results as red dots, primarily located in the Southern Ocean. The map includes a coordinate grid and a "POSITION" indicator.
- Variables:** A table listing the variables available for the search results:

Variables	Count
SUBSURFACE PRESSURE	296
SUBSURFACE SALINITY	296
SUBSURFACE TEMPERATURE	296

<https://data.blue-cloud.org>



BLUE-CLOUD VIRTUAL RESEARCH ENVIRONMENT



Collaboration



Sharing



Reuse



Reproducibility

Blue Cloud VRE promotes Open Science



Conclusions

- It is very important that ocean and marine data as collected by science, government, industry, and citizens are shared and become available for wider use in many applications and derived data products
- Europe over time has established an extensive landscape for marine data management, featuring several research data infrastructures for specific data types, complemented with EU initiatives such as EMODnet, Copernicus Marine Service, and European Open Science Cloud.
- There is an ongoing move to more integration between the infrastructures, making data and services more FAIR, and developing more cloud based analytical and visualisation capabilities, supported by powerful computing and storage resources. This is explored as part of ENVRI-FAIR, Blue-Cloud and EOSC-FUTURE projects.
- SeaDataNet is a major network of NODCs and it handles the data management for many data collectors, including data from most cruises of the European marine research vessel fleet.

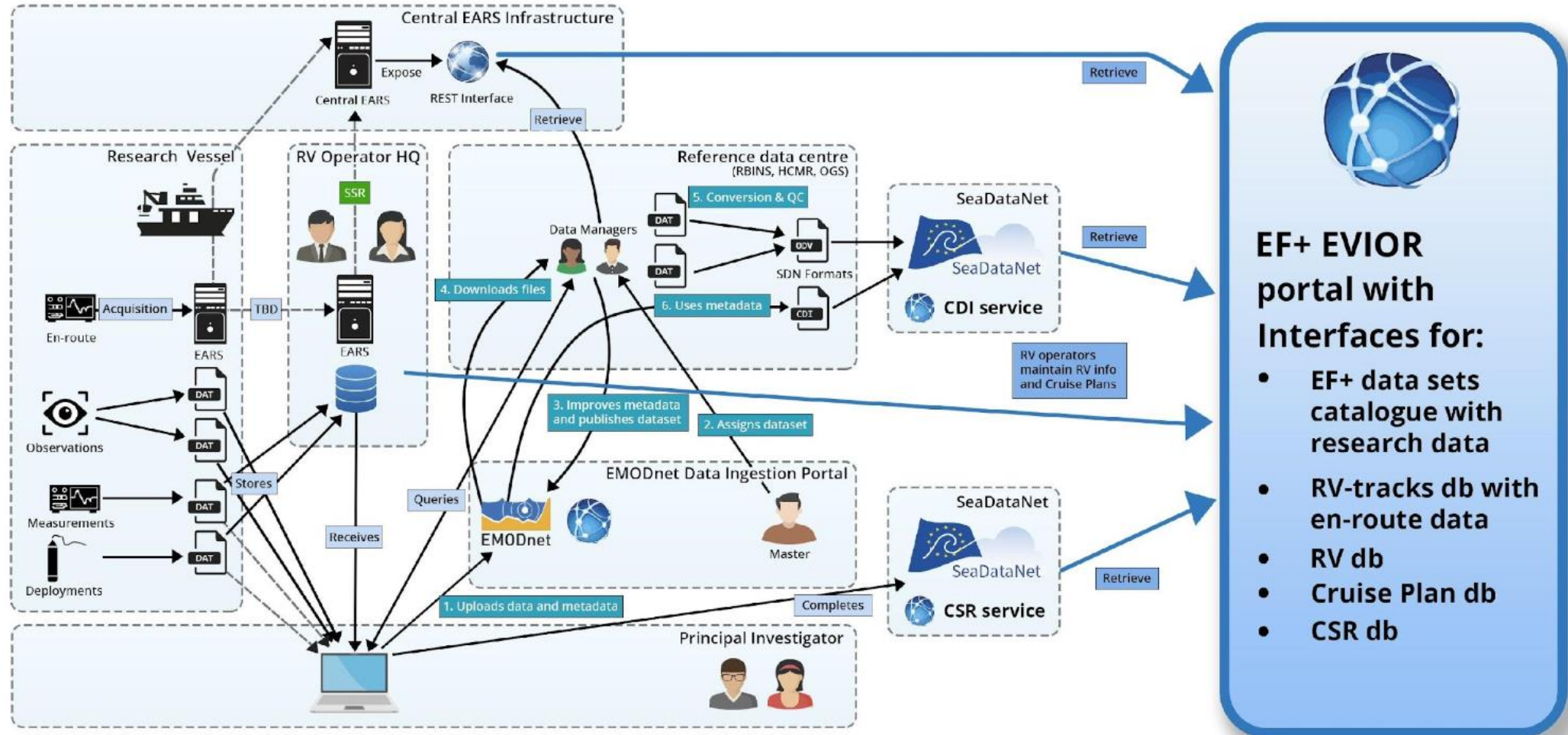


Eurofleets+ open data management strategy

- To ensure that the research data collected during the funded TA cruises, and the en-route data collected by the research vessels are made widely available in line with FAIR and Open Research Data principles
- To make use of existing European standards and services of SeaDataNet for managing and publishing collected cruise data sets, which ensures also distribution towards EMODnet and the wider community of potential users
- To establish and populate also the EVIOR ((European Virtual Infrastructure in Ocean Research) platform, embedded in Eurofleets website, with all information relevant for following the Eurofleets+ cruises and outcomes



Eurofleets+ data management workflow



Capturing and publishing metadata and data from cruise plan to underway to post cruise

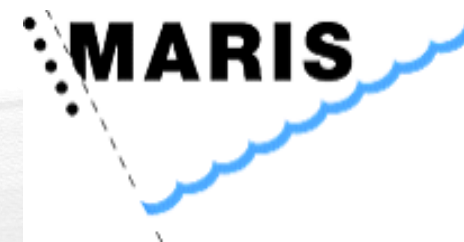


OGS

National Institute
of Oceanography
and Applied
Geophysics

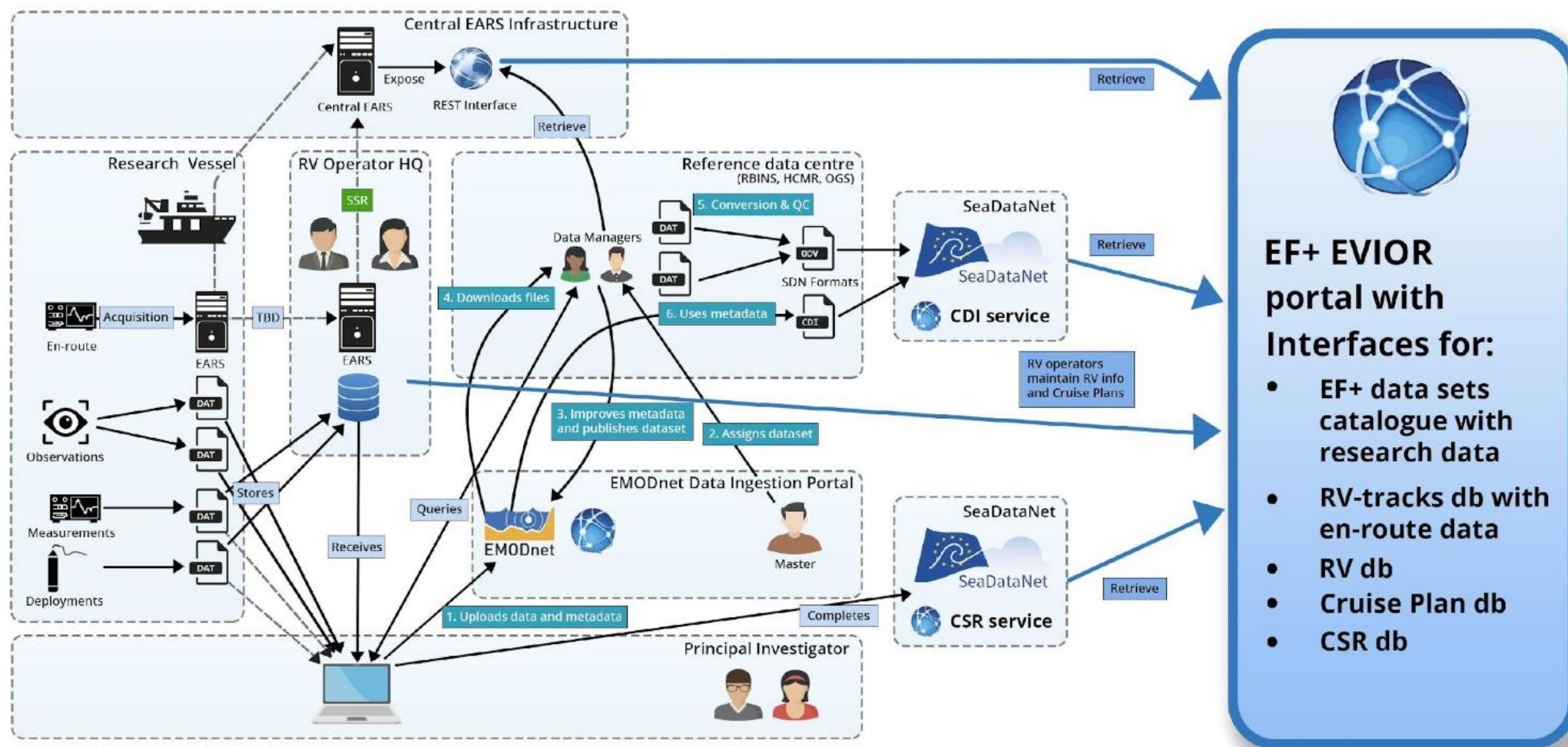
EVIOR (European Virtual Infrastructure in Ocean Research) platform

Dick M.A. Schaap





Data Management Workflow



Capturing and publishing metadata and data from cruise plan to underway to post cruise



EVIOR platform

Eurofleets+ Cruise Data Sets Catalogue

[RETURN TO EVIOR HOME](#) | [BACK](#)

[Next](#)

Record No.	1
Cruise ID	iMAR
Cruise Information	Cruise Narrative
Research Vessel	Pelagia
Date start	2021-05-18
Date end	2021-06-02
Cruise Plan	
Cruise Summary Report	Cruise Summary Report
En-route Data	En-route Data
Original Research data in EMODnet Ingestion	1 2 3 4 5
Original Research data in SeaDataNet SEANOE	1 2 3 4 5
Elaborated Research data in SeaDataNet CDI service	1 2 3
Elaborated Research data in EMODnet service	1 2

h (EVIOR)

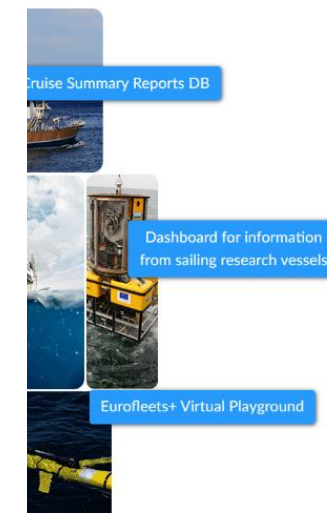


Figure: Virtual Playground system structure





NODCs for Data Management support

- TA cruise teams make an EF+ Data Management Plan (DMP)
- DMPs are reviewed by 3 NODCs in EF+: HCMR, OGS, RBINS
- One NODC is assigned per TA Cruise
- Guidance and support by NODC to PI and scientific teams:
 - Keeping index of data and sample acquisition during cruise, preferably using the EF+ EARS system on board
 - Preparing a Cruise Summary Report (CSR) after cruise
 - Ensuring transfer of data sets after scientific embargo with sufficient documentation for uptake by NODCs in SeaDataNet, also for wider distribution to EMODnet, CMEMS, and EF+ EVIOR portal

