



# SeaDataNet

History of the building and upgrade of an European  
oceanographic data infrastructure

Download the presentation:

<https://cloud.ifremer.fr/index.php/s/vKeMb4bD3pPXaoD>

*Michèle Fichaut, IFREMER, SISMER (NODC)*

Innovation platform "Sustainable Sea and ocean Solutions", 23 February 2022

sdn-userdesk@seadatanet.org – [www.seadatanet.org](http://www.seadatanet.org)

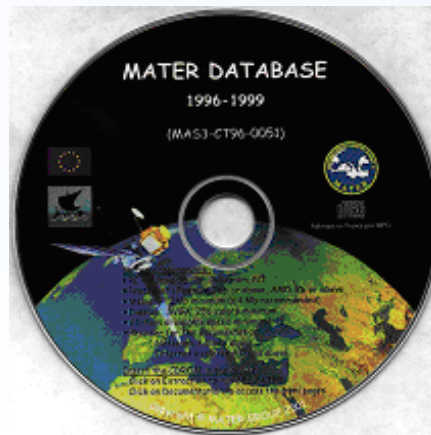
## Part 2

### The SeaDataNet infrastructure

- History
- Objective and principle
- Roadmap
- Method applied to reach interoperability of data and metadata
- Content and example of services

# Need for standardisation

- Over the last 30 years, development of European projects
  - Bringing together NODCs from different countries in joint oceanographic data projects
  - Need for standardisation, to disseminate homogeneous and coherent data sets: CD-ROMs at the end of the 1990s, beginning of the 2000s



## The SeaDataNet infrastructure

- Stop projects requiring the centralisation of homogeneous data
- Data remain in NODCs but are accessible from a single location: the SeaDataNet portal = virtual data centre, based on the European research infrastructure (RI) SeaDataNet



EU – MAST

EU – MAST II

EU-FP5

EU-FP6

EU-FP7

H2020



**90s**

*EDMED*

*Euronodim*

*MEDATLAS*

**2002-2005**

*Sea-Search*

**2006-2011**

*SeaDataNet*

**2011-2015**

*SeaDataNet II*

**2016-2021**

*SeaDataCloud*

Partners : 35 countries, bordering European seas

Coordonnated by IFREMER

## SeaDataNet - Objectives

- Federation of the the oceanographic data centres of 35 countries bordering the European seas
- Creation of a single virtual data centre, allowing a user searching for physics/chemistry data to connect to all 35 countries from a single user interface
- Distribute complete datasets in specific sea basins to privileged users (modellers)
- Creation and dissemination of products (climatologies and aggregated datasets) made from the data put into the infrastructure



## SeaDataNet Infrastructure - Principle

- semi-distributed system that aggregates **NODCs** and enhances the existing NODC network.
- The technical developments implemented allow the NODCs connected to the SeaDataNet system to be seen as a **single virtual data centre** able to deliver **quality controlled data**, metadata and products through a **single web portal**


## SeaDataNet – Road map

- SeaDataNet (2006-2011) : 10 M€
    - System implementation in 2 steps
      - Connection of the 10 most technologically advanced data centres to implement and test the system
      - Connecting the other 29 data centres in a progressive way with assistance from the 10 already connected
  - SeaDataNet 2 (2011-2015) : 6 M€
    - Make the system more reliable (monitoring), more automated (machine-to-machine data exchange) and sustainable (infrastructure funded outside the European project)
    - Connect more data centres
    - Add more data, and more types of data (biology)
- ➔ Bigger, Better, Faster





## SeaDataNet – Road map

- SeaDataCloud (2016-2021) : 10 M€The logo for SeaDataCloud, featuring a blue square with a white wave and yellow stars, and the text "SeaDataCloud" below it.
  - Improving access to data
  - Take into account the evolution of technologies => Cloud, HPC
    - More data processing capacity
    - Improved response times
  - Give a central role to the users
    - Provide the user with tools in a virtual environment (VRE, Virtual Research Environment) in which he/she will be able to work on his/her own data + data from SeaDataNet
    - To store his/her working environment: MySeaDataCloud

## SeaDataNet : Method

- Development of standards
  - Common vocabulary for metadata
  - Common protocol for data and metadata control (Have comparable data)
  - Common file formats
- Definition of common catalogues
- Definition of rules for making data available
- Use of common software developed in the framework of SeaDataNet and made available to all partners (and more)

# Speak the same language: Common vocabularies → Interoperability



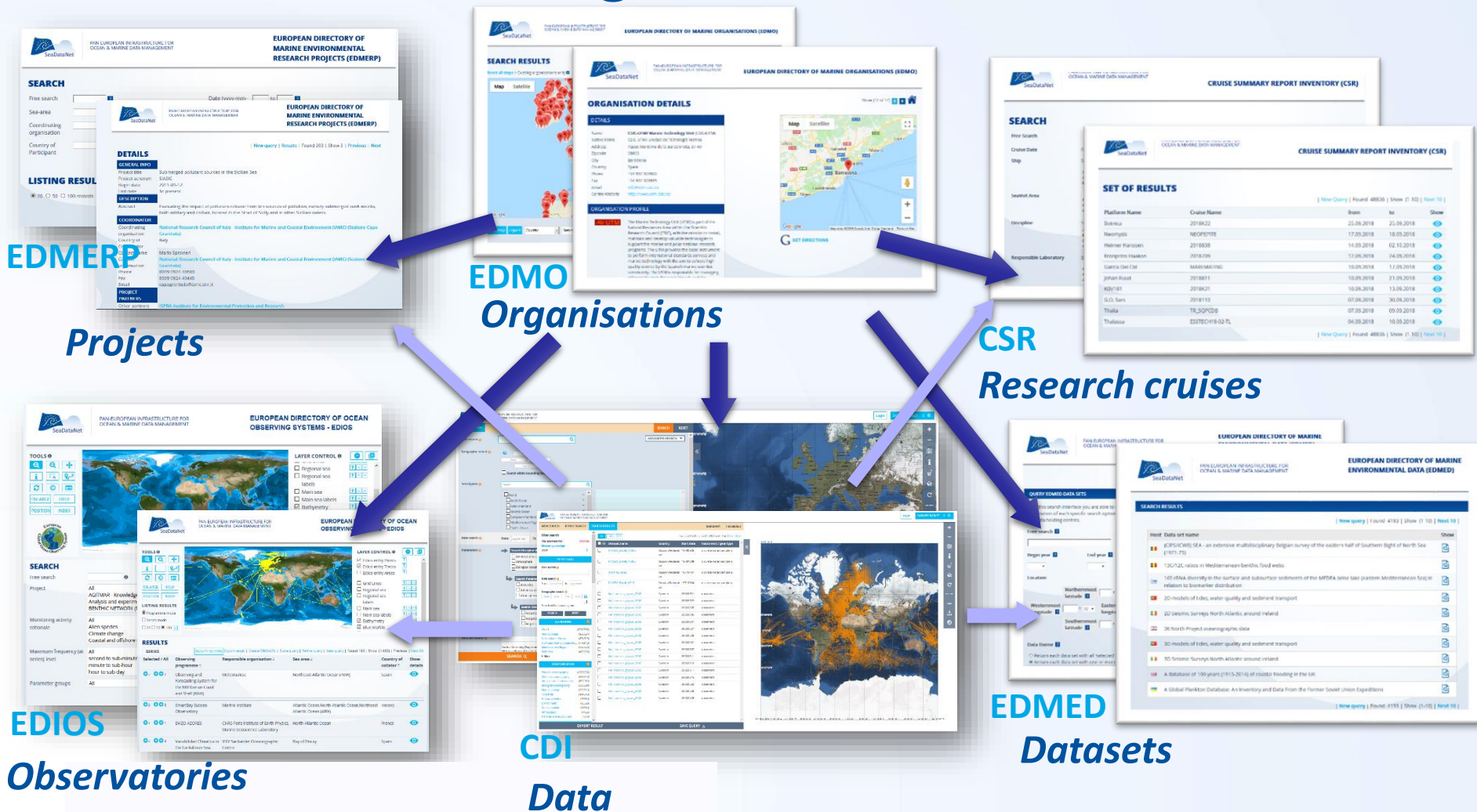
- Vocabulary lists maintained by the British Oceanographic Data Centre (NercVocabularyServer/BODC)
- 90,000 terms in over 110 vocabulary lists
  - Geographical area, ships, ports, scientific disciplines, data types, parameters, measurement units, instruments, positioning systems.....
- On-line through
  - Web site : <https://www.seadatanet.org> → Look-up vocabularies
  - Web services: [http://www.bodc.ac.uk/products/web\\_services/vocab/](http://www.bodc.ac.uk/products/web_services/vocab/)

# To speak the same language : Common metadata descriptions → interoperability

- Following International/European standards
  - Metadata descriptions based upon ISO-19115 and ISO-19139 for the compliancy to INSPIRE
  - available on-line on SeaDataNet website (<https://www.seadatanet.org/Standards/Metadata-formats>) and on the Ocean best practices repository (<https://www.oceanbestpractices.org/>)



## 6 common catalogues of metadata

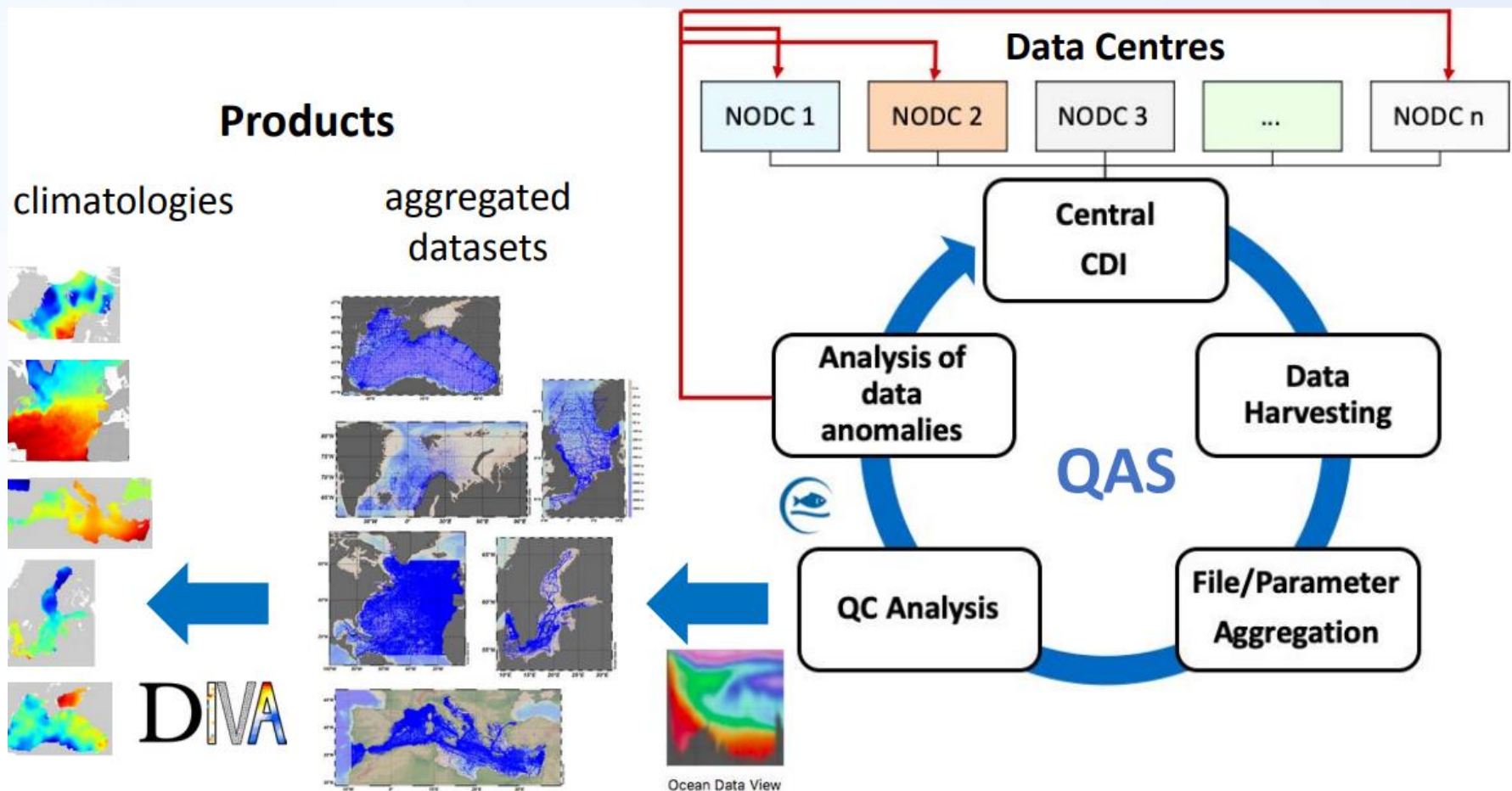


## Have comparable data → interoperability

- Same file formats relying on the common vocabularies
  - ODV - Ocean Data view (ASCII)
  - MEDATLAS (ASCII)
  - NETCDF – CF (Binaire)
- Quality check protocol based on international (IOC/IODE) recommendations applied by all data centres
  - With automatic and manual checks
- Same quality flags on all measurements (part of the common vocabularies)
- Quality Assurance strategy, implementing a QC -Loop

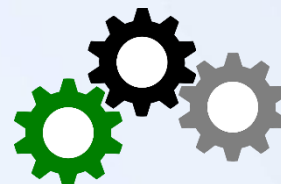


# QC-Loop with feedback to data centres





# Have common tools for data and metadata preparation



- **Tools are distributed to all SeaDataNet partners**



- **MIKADO** : To generate the metadata descriptions of the SDN catalogues



- **NEMO** To convert files to SeaDataNet formats



- **OCTOPUS** : To convert from One SDN format to another

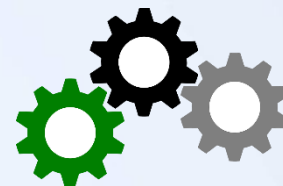


- **Ocean Data View (ODV)** : To visualise and QC the data



- **DIVA** : For data analysis and product generations (climatology)
- **Download Manager (DM) and Replication Manager (RM)** : to send datafiles from one data centre to the users or to the cloud

# Have common tools for data and metadata access



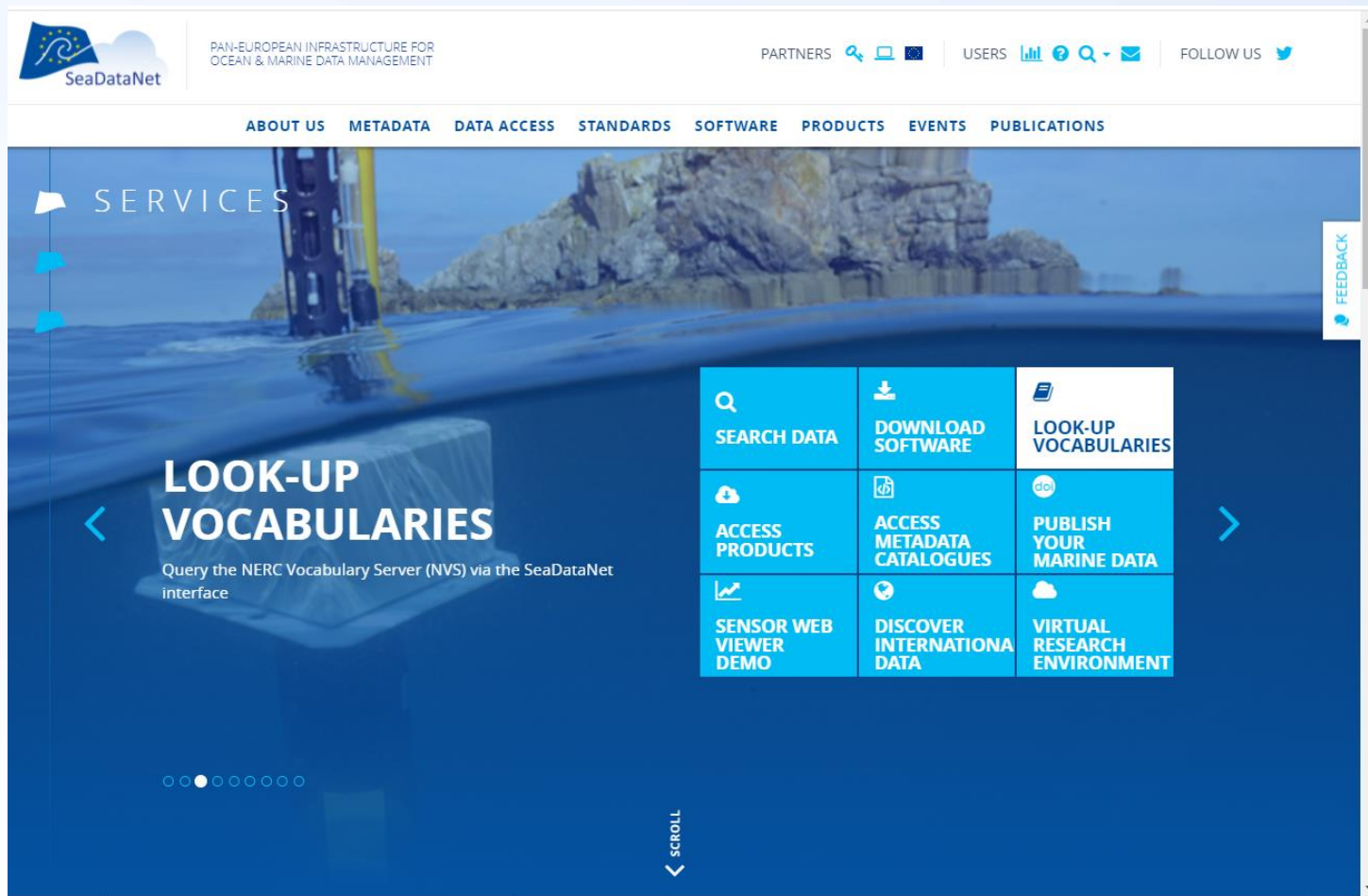
- **Installed on central servers: :**
  - **Central catalogues and the corresponding CMS** = Content Management System for updating (for projects and cruises and datasets)
  - **Web interfaces** for querying the various catalogues and searching for data
  - **Request Status Manager (RSM):** so that a user can track the progress of his or her data request
  - **MARINE ID** for identification in case of data downloading

## Define the data policy

- Most of the SeaDataNet data are publicly available (CC-by licence) for download (91%)
- 9% of restricted data are managed too : restriction like moratorium or other access restrictions
  - The metadata of restricted data are available but
  - Their distribution is under the responsibility of each data centre that have included them in the SeaDataNet infrastructure : case per case negotiation with the data user

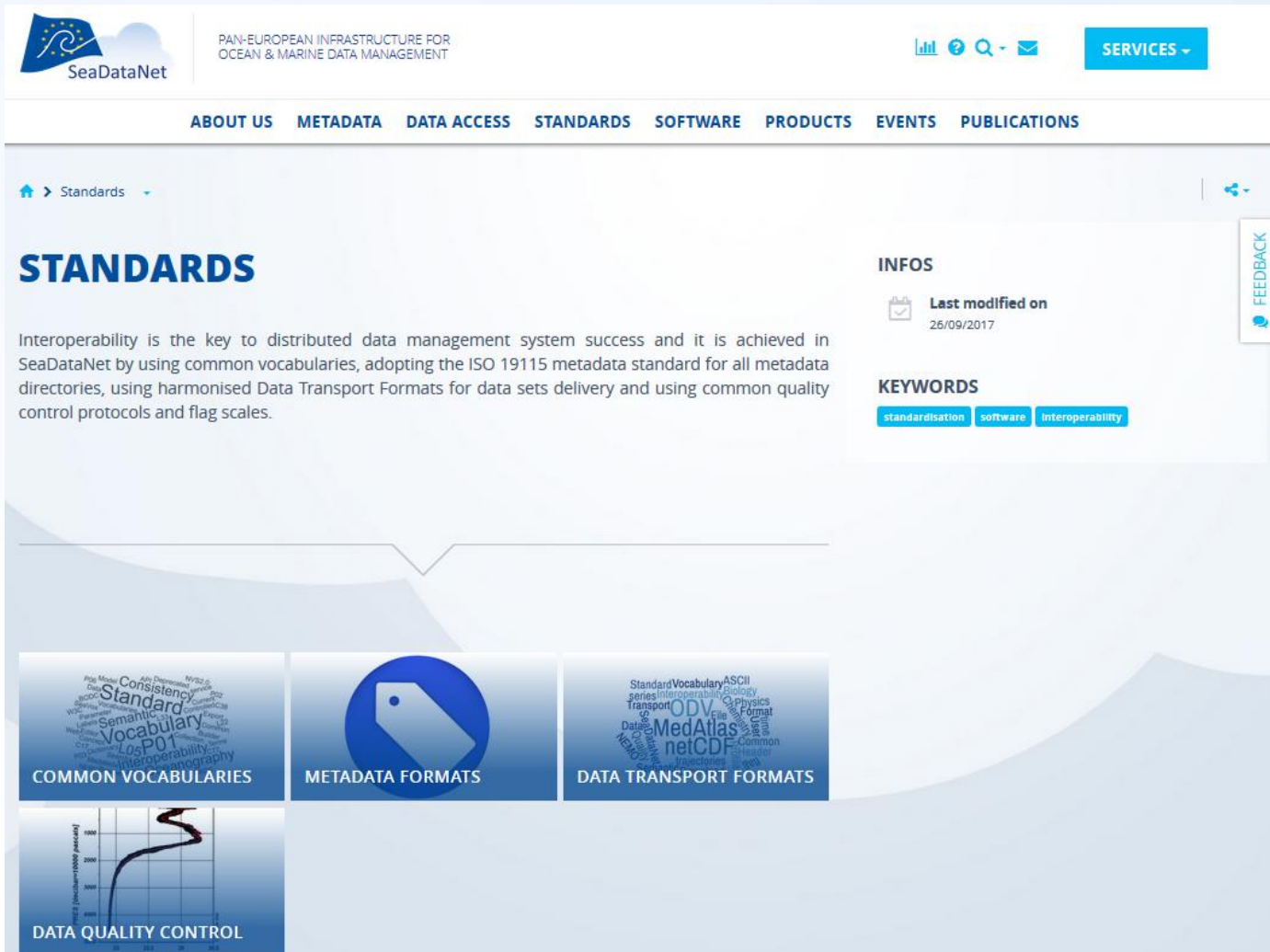


# SeaDataNet portal- <https://www.seadatanet.org/>



- Metadata
- Data
- Standards
- Tools
- Products

# Standards



The screenshot shows the SeaDataNet website's 'Standards' page. The header includes the SeaDataNet logo, the text 'PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT', and navigation icons. A main navigation bar lists: ABOUT US, METADATA, DATA ACCESS, STANDARDS, SOFTWARE, PRODUCTS, EVENTS, and PUBLICATIONS. The 'Standards' page content includes a breadcrumb 'Home > Standards', a large 'STANDARDS' heading, and a paragraph explaining interoperability. A right sidebar contains 'INFOS' (Last modified on 26/09/2017) and 'KEYWORDS' (standardisation, software, interoperability). A vertical 'FEEDBACK' button is on the far right. The footer features four tiles: 'COMMON VOCABULARIES' (word cloud), 'METADATA FORMATS' (tag icon), 'DATA TRANSPORT FORMATS' (word cloud), and 'DATA QUALITY CONTROL' (line graph).

**SeaDataNet**  
PAN-EUROPEAN INFRASTRUCTURE FOR  
OCEAN & MARINE DATA MANAGEMENT

**STANDARDS**

Interoperability is the key to distributed data management system success and it is achieved in SeaDataNet by using common vocabularies, adopting the ISO 19115 metadata standard for all metadata directories, using harmonised Data Transport Formats for data sets delivery and using common quality control protocols and flag scales.

**INFOS**  
Last modified on 26/09/2017

**KEYWORDS**  
standardisation software interoperability

**COMMON VOCABULARIES**

**METADATA FORMATS**

**DATA TRANSPORT FORMATS**

**DATA QUALITY CONTROL**



# Tools



DOWNLOAD SOFTWARE

## SOFTWARE

### INFOS



Last modified on  
16/09/2019

A major objective and challenge in SeaDataNet is to provide an integrated and harmonised overview and access to data resources, managed by distributed data centres. Moreover it is an objective to provide users common means for analysing and presenting data and data products. Therefore the Technical Task Team of SeaDataNet has designed an overall system architecture, and is developing common software tools for data centres and users.

Common software tools are being developed and freely made available to **Data Centres and/or End Users** for:

- Editing and generating XML metadata entries: [MIKADO javatool](#)
- Tool for the generation of spatial objects from vessel navigation during observations: [EndsAndBends](#)
- SeaDataNet file format converter : [OCTOPUS](#)
- Conversion of any ASCII format to the SeaDataNet ODV4 ASCII format: [NEMO javatool](#)
- Connecting systems of Data Centres to the SeaDataNet portal for data access: [Replication Manager javatool](#)
- Analysing and visualising of data sets: [Ocean Data View \(ODV\) software package](#)
- Interpolation and variational analysis of data sets: [DIVA software package](#)



# Capacity building

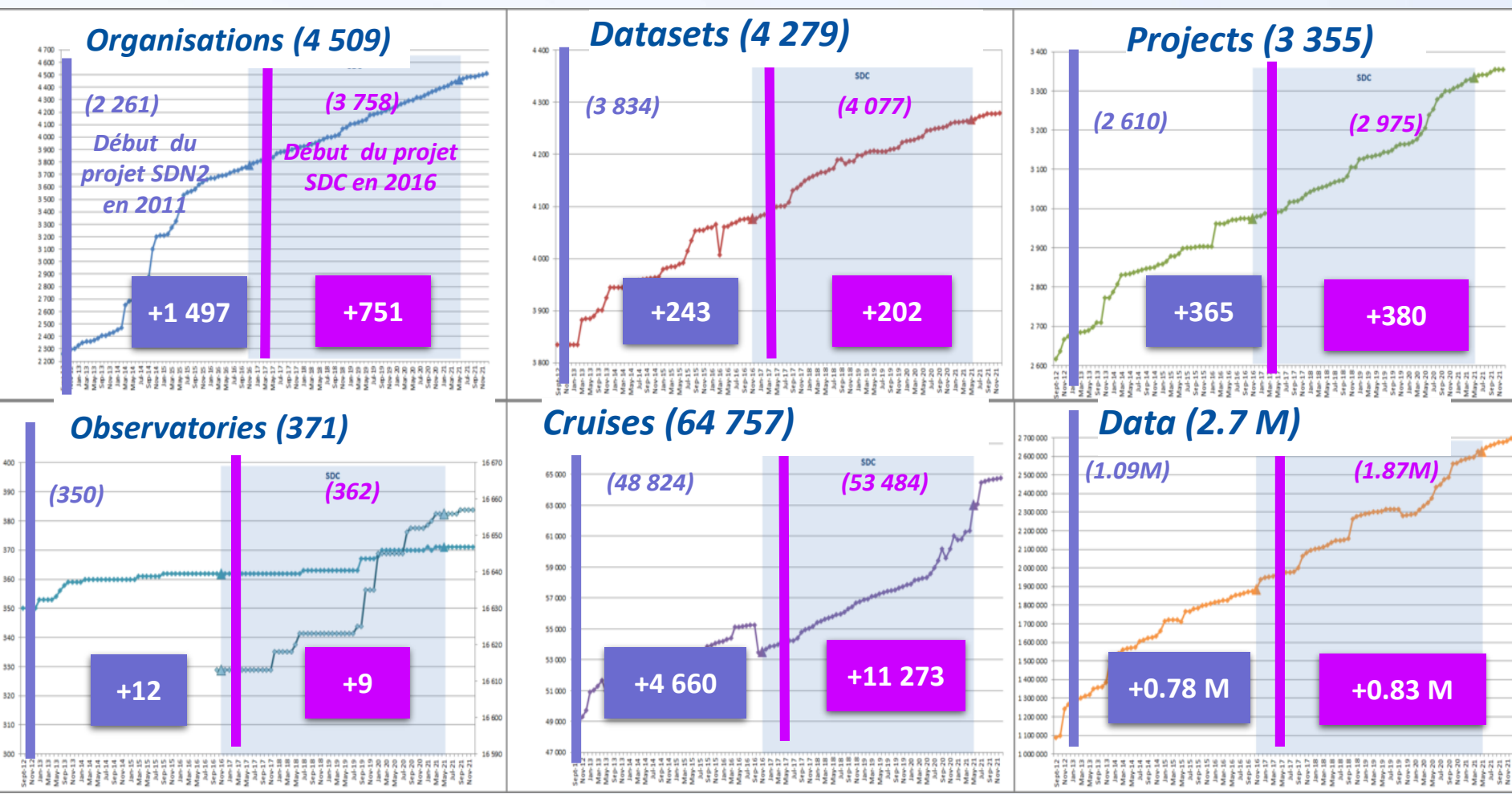
- Around every 2 years: Training for all data centres connected to SDN infrastructure

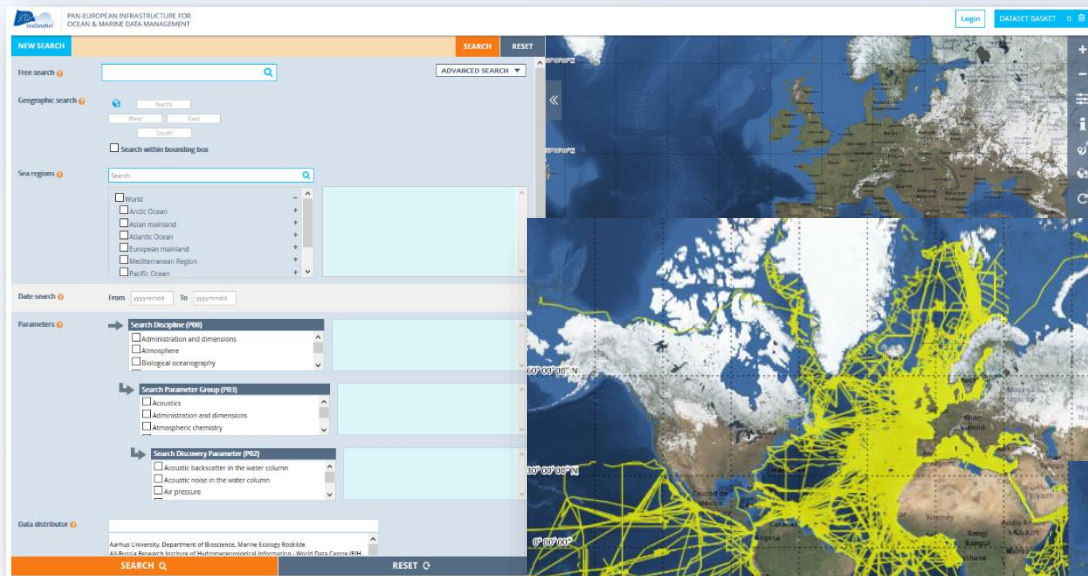




# Current content of the SDN infrastructure

- Content of the catalogues





Search interface

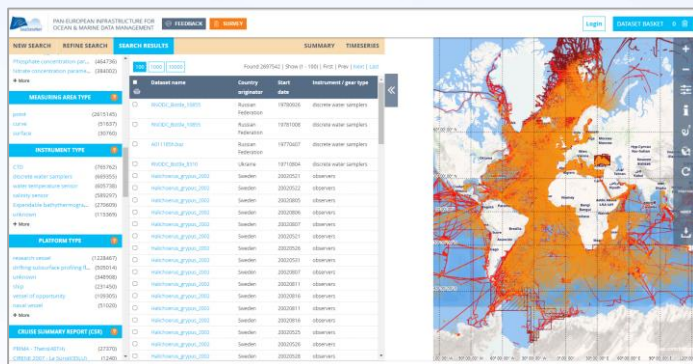
[cdi.seadatanet.org/search](https://cdi.seadatanet.org/search)

Trajectories

Vertical profiles or time series

- since 1800 → 2022
- 2.69 Millions of CDI for physics, chemistry, biology, geology and geophysics
- 91 % free access after Marine-ID connection

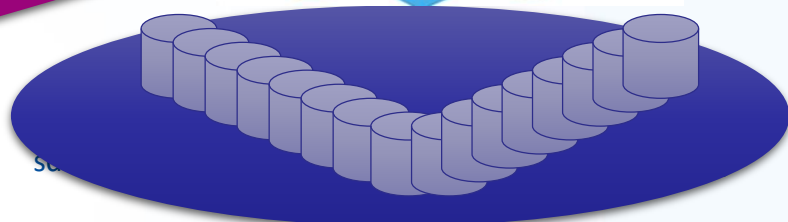
## Service to users: Discovery and Access to data



*Data  
discovery*



- At the end of the 1<sup>st</sup> project (2011): 40 data centres, data distributed in NODCs
- 2<sup>nd</sup> project (2016): 80 data centres, data distributed in NODCs
- Currently (2022): 113 data centres, public data duplicated in the Cloud

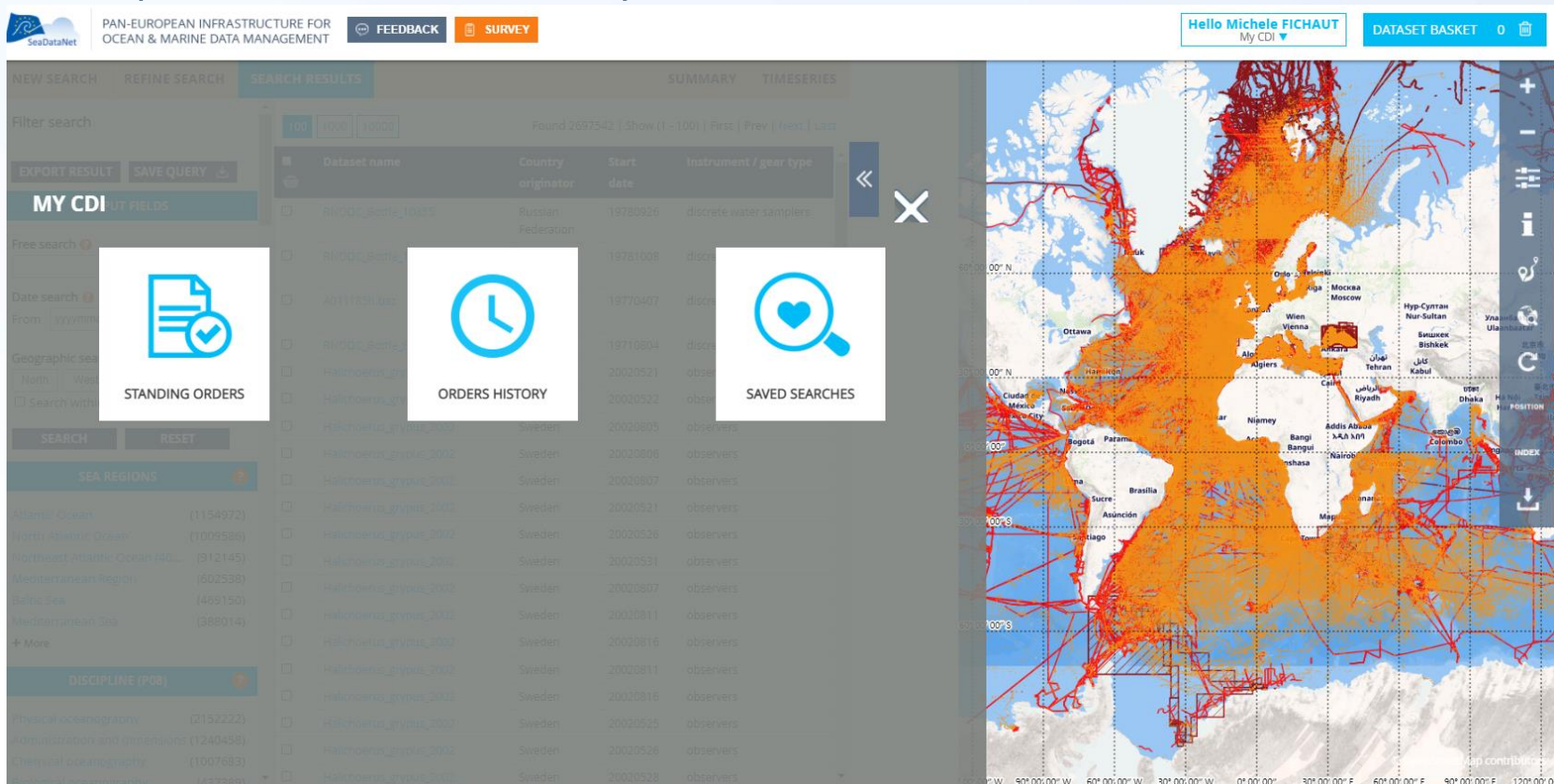


*Data sources =  
Data centres fed by  
more than 800 collecting laboratories*



# User personal environment: MySeaDataCloud

- To be able to save and share searches, to follow the status of data requests, to have an history of data downloads




The screenshot displays the MySeaDataCloud user interface. At the top, there is a navigation bar with the SeaDataNet logo, the text "PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT", and buttons for "FEEDBACK" and "SURVEY". On the right, a user profile for "Hello Michele FICHAUT" and a "DATASET BASKET" with 0 items are visible.

The main interface is divided into several sections:

- NEW SEARCH / REFINE SEARCH / SEARCH RESULTS:** The "SEARCH RESULTS" tab is active, showing a table of datasets. The table has columns for "Dataset name", "Country originator", "Start date", and "Instrument / gear type". A list of datasets is displayed, including "H2O2\_2002\_1000", "H2O2\_2002\_1001", "H2O2\_2002\_1002", etc.
- MY CDI IT FIELDS:** A section for filtering search results, including "Free search", "Date search", and "Geographic search".
- STANDING ORDERS:** A button with a document icon and a checkmark.
- ORDERS HISTORY:** A button with a clock icon.
- SAVED SEARCHES:** A button with a magnifying glass icon and a heart.
- SEA REGIONS:** A list of sea regions with their respective counts, such as "Atlantic Ocean (1154972)", "North Atlantic Ocean (1009586)", etc.
- DISCIPLINE (P08):** A list of disciplines with their respective counts, such as "Physical oceanography (215222)", "Biology and zoology (1240458)", etc.
- Map:** A map of the Atlantic Ocean showing data points and search results. The map includes a grid of latitude and longitude lines and a legend for "POSITION".

# User personal environment: MySeaDataCloud

- Standing orders




PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

FEEDBACK

Hello Michele FICHAUT  
My CDI ▼

DATASET BASKET 0



PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

FEEDBACK

SURVEY

STANDING ORDERS

ORDERS HISTORY

SAVED SEARCHES

CLOSE

**MY STANDING ORDERS**

View the status of your orders below (the colours indicate in which stage data is). Unrestricted data will normally be ready for delivery in several minutes (or a bit more when your order is large). Restricted data depends on the release by the datacenter and will take more time. You will be notified per email when the status has changed.

Waiting for processing

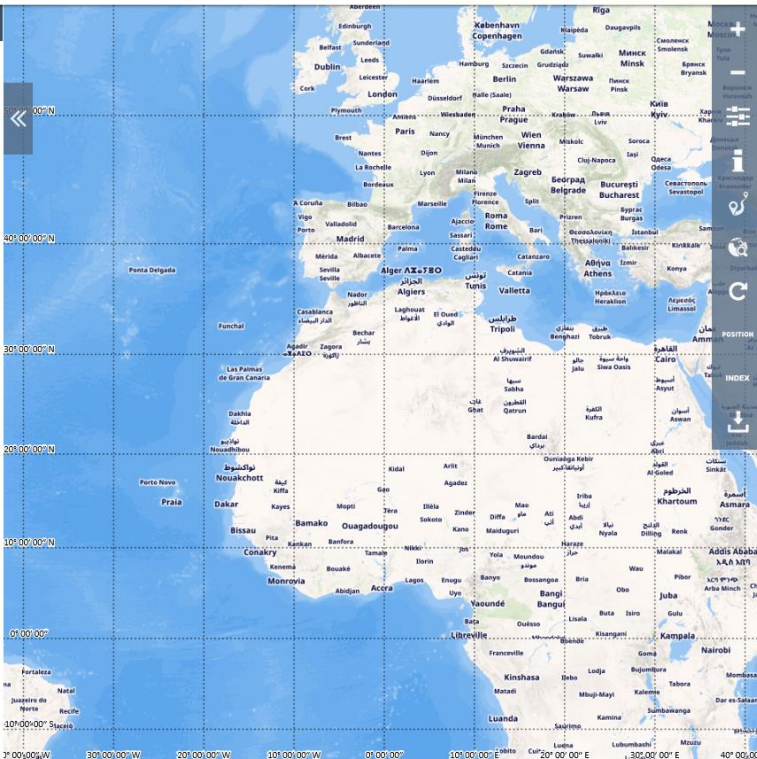
Approval pending

Ready for user action

User action completed

Access denied


Order number	Name	Unrestricted	Restricted	Count	Date created	Action
55425	CTD Mer IROISE	2520		2520	29-03-2021	
55256	PCB in biota	3		3	02-03-2021	





18:09

## User personal environment: MySeaDataCloud

- Orders history


PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT


PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT


FEEDBACK

Hello Michele FICHAUT  
My CDI

DATASET BASKET 0

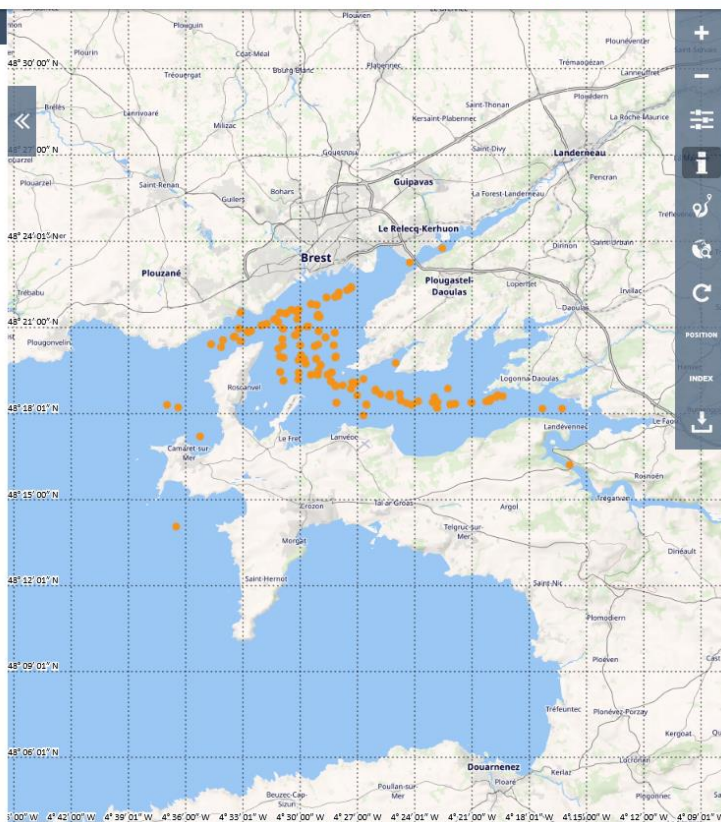
STANDING ORDERS
ORDERS HISTORY
SAVED SEARCHES


CLOSE

### MY ORDER HISTORY

Found 38 | Show (1 - 10) | First | Prev | Next | Last

Data centre	Country	Action complete	Access denied	All
Swedish Meteorological and Hydrological Institute	Sweden	3		3
Sinop University, Fisheries Faculty	Turkey	60		60
Shom	France	1852	1	1853
Scientific - Research Firm "GAMMA"	Georgia	12		12
Royal Netherlands Meteorological Institute	Netherlands	25		25
PANGAEA - Data Publisher for Earth & Environmental Science	Germany	1		1
ORION	Cyprus	772		772
OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale), Division of Oceanography	Italy	1868	4	1872
Odesa National I.I.Mechnikov University	Ukraine	25		25
OceanWise Limited	United Kingdom		3	3





STANDING ORDERS

NEW SEARCH
REFINE SEARCH

Filter search  
You searched for:  
Reset all

MY CDI


STANDING ORDERS

Geographic search  
Search within bounding box  
SEARCH
RESET

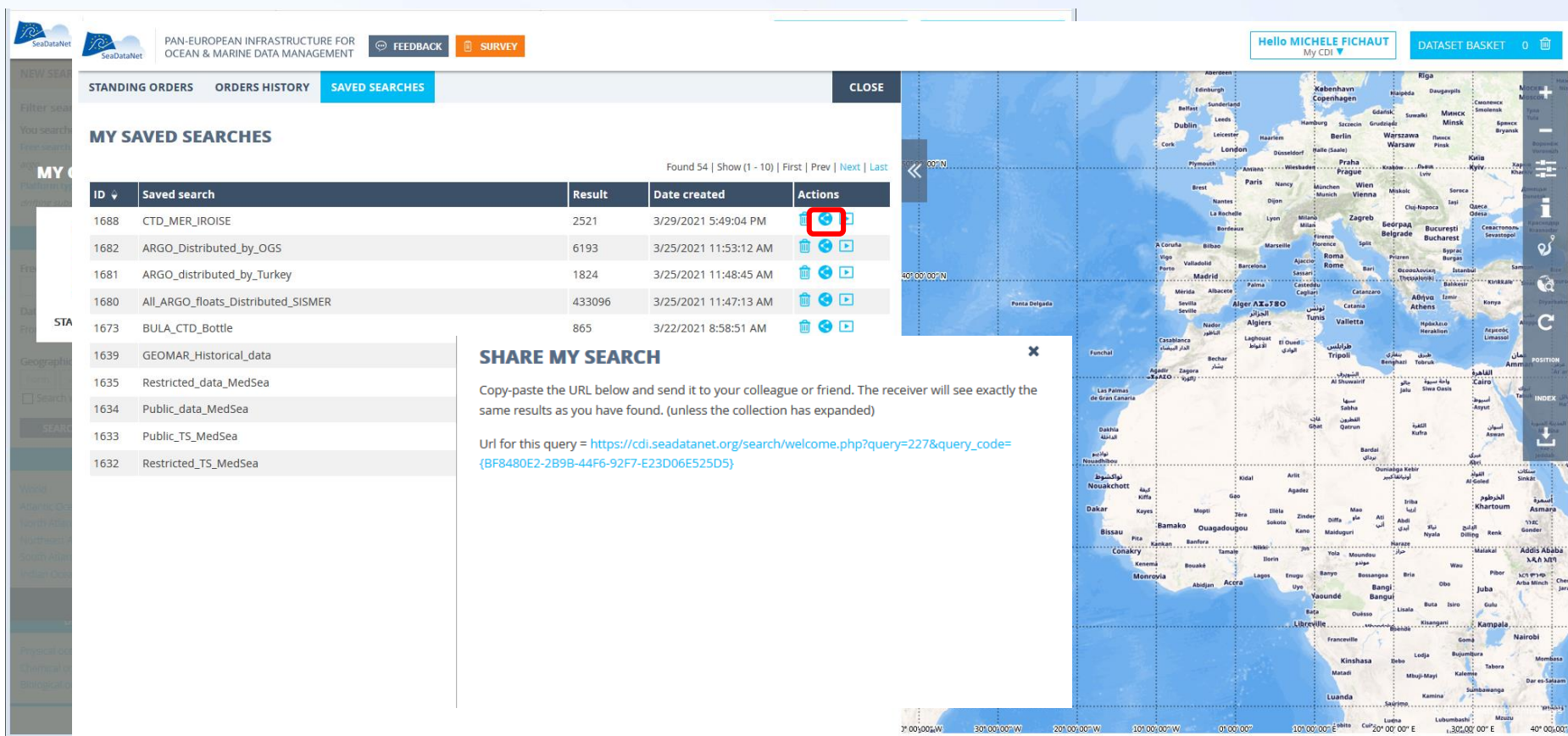
SEA REGIONS

EXPORT RESULTS








# User personal environment: MySeaDataCloud

- Saved searches



The screenshot displays the MySeaDataCloud user interface. At the top, there's a header with the SeaDataNet logo, navigation tabs (STANDING ORDERS, ORDERS HISTORY, SAVED SEARCHES), and a user profile section for 'Hello MICHELE FICHAUT'. Below the tabs, the 'MY SAVED SEARCHES' section shows a table of saved searches. A red box highlights the 'Actions' column for the first search. To the right, a map of the Mediterranean and surrounding regions is visible. A 'SHARE MY SEARCH' dialog box is open, providing a URL to share the search results.

**MY SAVED SEARCHES**

ID	Saved search	Result	Date created	Actions
1688	CTD_MER_IROISE	2521	3/29/2021 5:49:04 PM	
1682	ARGO_Distributed_by_OGS	6193	3/25/2021 11:53:12 AM	
1681	ARGO_distributed_by_Turkey	1824	3/25/2021 11:48:45 AM	
1680	All_ARGO_floats_Distributed_SISMER	433096	3/25/2021 11:47:13 AM	
1673	BULA_CTD_Bottle	865	3/22/2021 8:58:51 AM	
1639	GEOMAR_Historical_data			
1635	Restricted_data_MedSea			
1634	Public_data_MedSea			
1633	Public_TS_MedSea			
1632	Restricted_TS_MedSea			

**SHARE MY SEARCH**

Copy-paste the URL below and send it to your colleague or friend. The receiver will see exactly the same results as you have found. (unless the collection has expanded)

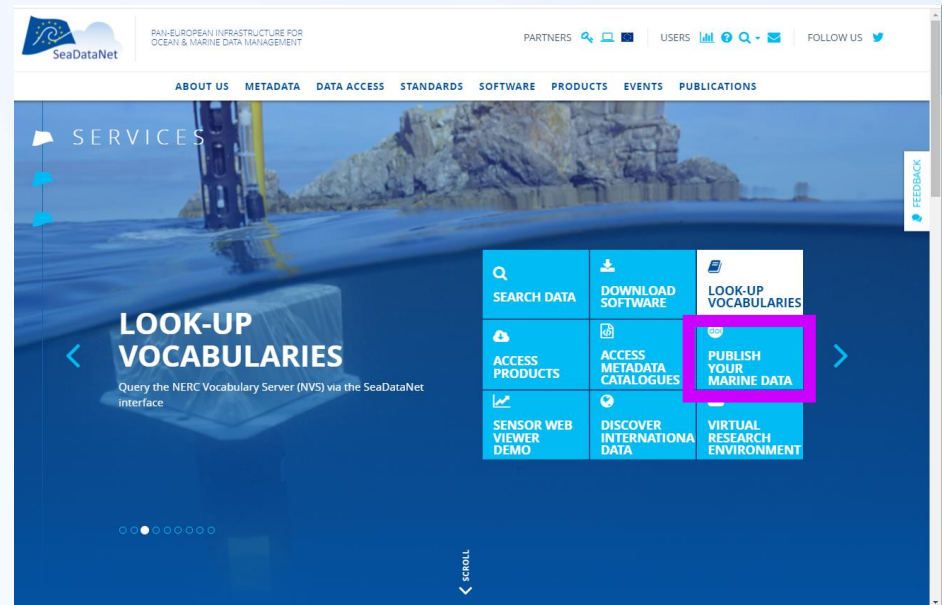
Url for this query = [https://cdi.seadatanet.org/search/welcome.php?query=227&query\\_code={BF4840E2-2B9B-44F6-92F7-E23D06E525D5}](https://cdi.seadatanet.org/search/welcome.php?query=227&query_code={BF4840E2-2B9B-44F6-92F7-E23D06E525D5})



## Publishing service

SEANOE Sea scientific open data publication

- To offer scientists a place to publish their datasets and get DOI (Digital Object Identifier)
  - Useful for publications based on datasets, publisher often require a DOI on the studied dataset



# Publishing service : landing page

## □ Cruises PRIMO 1994 (CNR IT): Hydrographic measurements in the Sicily Channel and in the southern Tyrrhenian Sea (spring and fall 1994)

**Date** 2021-12-13

**Temporal extent** 1994-05-20 -1994-10-18

**Author(s)** [Sparnocchia Stefania](#)<sup>1</sup>, Borghini Mireno<sup>2</sup>

**Affiliation(s)** 1 : CNR-ISMAR, Trieste, Italy  
2 : CNR-ISMAR, La Spezia, Italy

**DOI** [10.17882/85185](#)

**Publisher** SEANO

**Keyword(s)** CTD, Sicily Channel, Tyrrhenian Sea, Western Mediterranean

**Abstract** This data set contains the CTD data collected from the RV URANIA of the CNR (Italy) during the PRIMO-94 and PRIMO-94B cruises (20-29 May 1994 and 5-18 October 1994). These cruises were part of an intensive field program in the Sicily Channel and in the southern Tyrrhenian basin conducted by the Stazione Oceanografica of CNR in different periods from 1993 to 1995. Data have been used in several studies (see References).  
CTD profiles were collected using a Neil-Brown MKIII CTD. The probe was calibrated in temperature and conductivity at the SACLANT Center of La Spezia, before and after each cruise, and at sea in salinity and oxygen, against water samples. Declared instrumental precisions were 0.002 °C for temperature and 0.005 for salinity (PSS-78).  
The data set is provided per cruise as ODV Spreadsheet files in TXT format, containing:

- Cruise name
- Station number
- Type of acquisition (here C)
- Date in mon/day/yr and Time in hh:mm:ss
- Coordinates in Longitude [degrees\_east] and Latitude [degrees\_north]
- Bottom depth [m]
- Depth, salt water [m]
- Temperature, IPTS-68 [degC]
- Conductivity [S/m]
- Temperature, ITS-90 [degC]
- Salinity, PSS-78 (Practical Salinity)
- Dissolved oxygen [ml/l]

Click to download the data

DATA

### Utilisation

These data are published without any warranty, express or implied. The user assumes all risk arising from their use. These data are intended to be quality controlled, but it is possible that they contain errors. It is the unique responsibility of the user to assess if the data are appropriate for their use, and to interpret the data, data quality, and data accuracy accordingly. Authors welcome users to ask questions and report problems.

### Acknowledgements

This data set was collected by the group known as the Stazione Oceanografica (Oceanographic Station) of the CNR, Pozzuolo di Lerici, La Spezia, led by Mario Astraldi and Gian Pietro Gasparini. We are grateful to Mr. Carlo Galli, Mr. Eglito Lazzoni and Mr. Domenico Bacciola for their remarkable contribution in the field and in the laboratory work. The research was funded by the EU through the MAST program (Contracts MAS2-CT93-0061 GEODYME and MAS2-CT93-0066 EUROMODEL II-MTP). The experiment in the Sicily Channel was a contribution to the IOC Programmes POEM and PRIMO.

### Sensor metadata

Neil-Brown MK III CTD

### Data

File	Size	Format	Processing	Access
<a href="#">CTD Data from PRIMO-94</a>	9 MB	ODV		<a href="#">Open access</a>
<a href="#">CTD Data from PRIMO-94B</a>	9 MB	ODV		<a href="#">Open access</a>

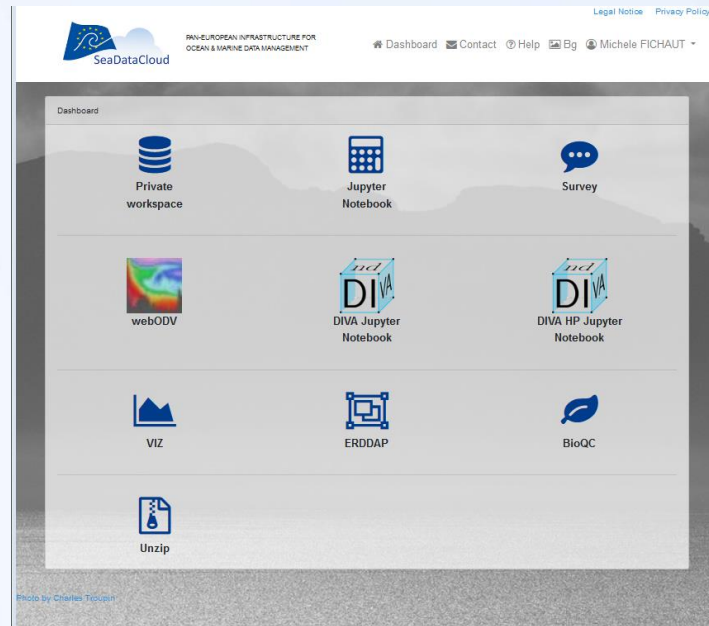
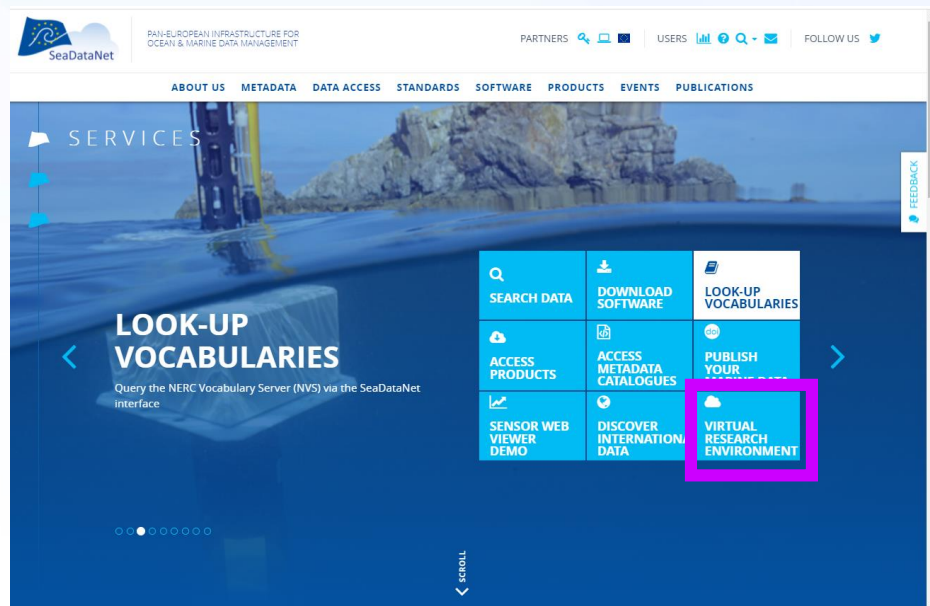


## How to cite

Sparnocchia Stefania, Borghini Mireno (2021). Cruises PRIMO 1994 (CNR IT): Hydrographic measurements in the Sicily Channel and in the southern Tyrrhenian Sea (spring and fall 1994). SEANO. <https://doi.org/10.17882/85185>

## Virtual Research Environment for users

- A virtual environment in the Cloud,
  - With data and tools
  - Possibility to add his/her own data



# European landscape in terms of Marine Data management 3 main components

