



SeaDataCloud

SeaDataNet, What's up in the cloud?

Michele Fichaut (IFREMER, FR - Coordinator),

[Dick M.A. Schaap \(MARIS, NL, Technical coordinator\)](#)

IMDIS 2018 Conference, 5 November 2018, Barcelona - Spain

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



European Open Science Cloud (EOSC)

- **6 May 2015** the EU adopted the Digital Single Markets strategy and announced the launch of a **cloud for research data**
- A **High Level Expert Group** was established to advise on scientific services to be provided on the cloud and its governance structure
- **16 April 2016** the EU published a package of measures for digitising European industry, including a communication on the **European Cloud Initiative**
- **12 June 2017** the **European Open Science Cloud Summit** took place bringing together key players from across Europe, to make the **'EOSC a reality by 2020'**
- **14 March 2018** the EU adopted the **Implementation Roadmap for the European Science Cloud**

EOSC should to give the EU a global lead in research data management and ensure that European scientists reap the full benefits of data-driven science. It also foresees setting up a European Data Infrastructure, with high-capacity cloud solutions with super-computing capacity.

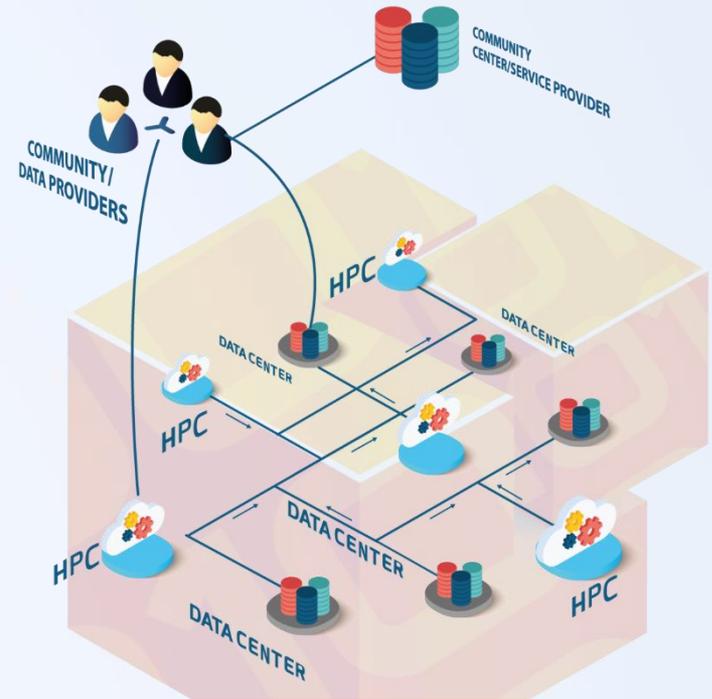
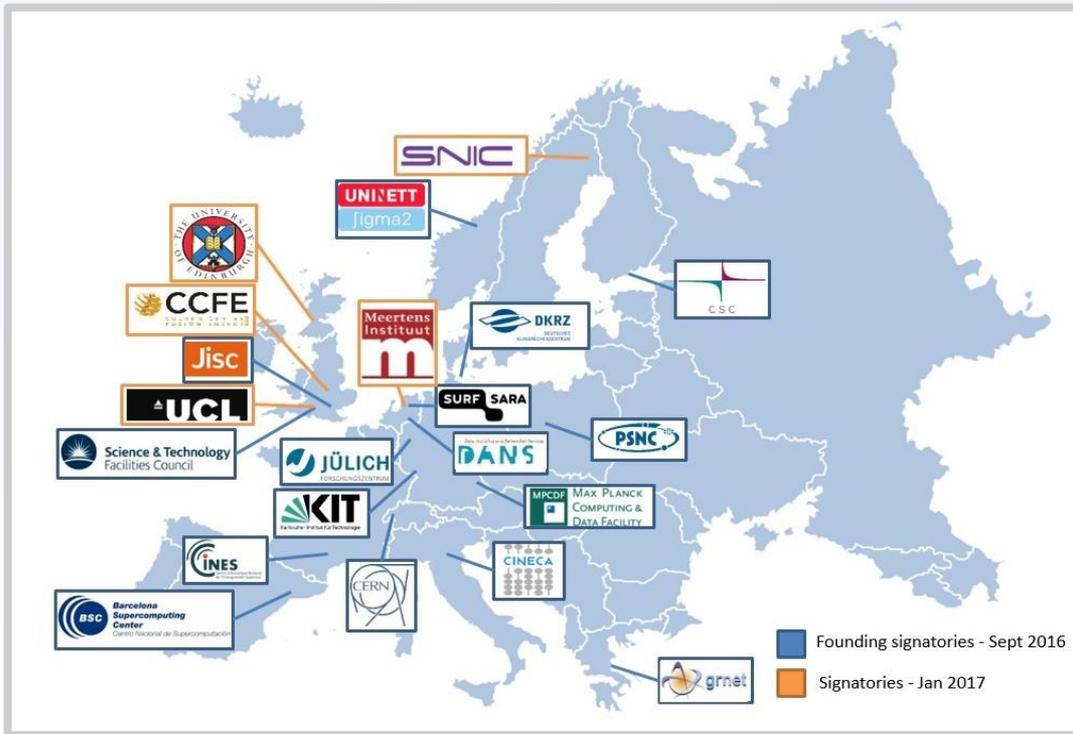


SeaDataNet anticipating EOSC

- In anticipation of these EOSC planning activities, SeaDataNet submitted in **March 2016** the **SeaDataCloud proposal** with the following aims:
 - Further **developing skills, standards, tools and services** for dealing with marine data, such as handling data from new instruments, INSPIRE compliance, interoperability with other data infrastructures, vocabularies governance, and adopting new technical approaches such as ‘Linked Data’, and ‘Sensor Web Enablement’,
 - Exploring the trend towards cloud storage and cloud computing, also taking into account **‘big data’ challenges**, for instance by analysing an architecture for a Virtual Research Environment (VRE) and developing a VRE pilot
 - Improving and expanding services and tools for **data providers** for connecting and ingesting data AND for **users** to make it easier and more performing to find, access, and use datasets
 - ***Better joining in an early stage and riding the EOSC wave than losing ourselves in the undertow***

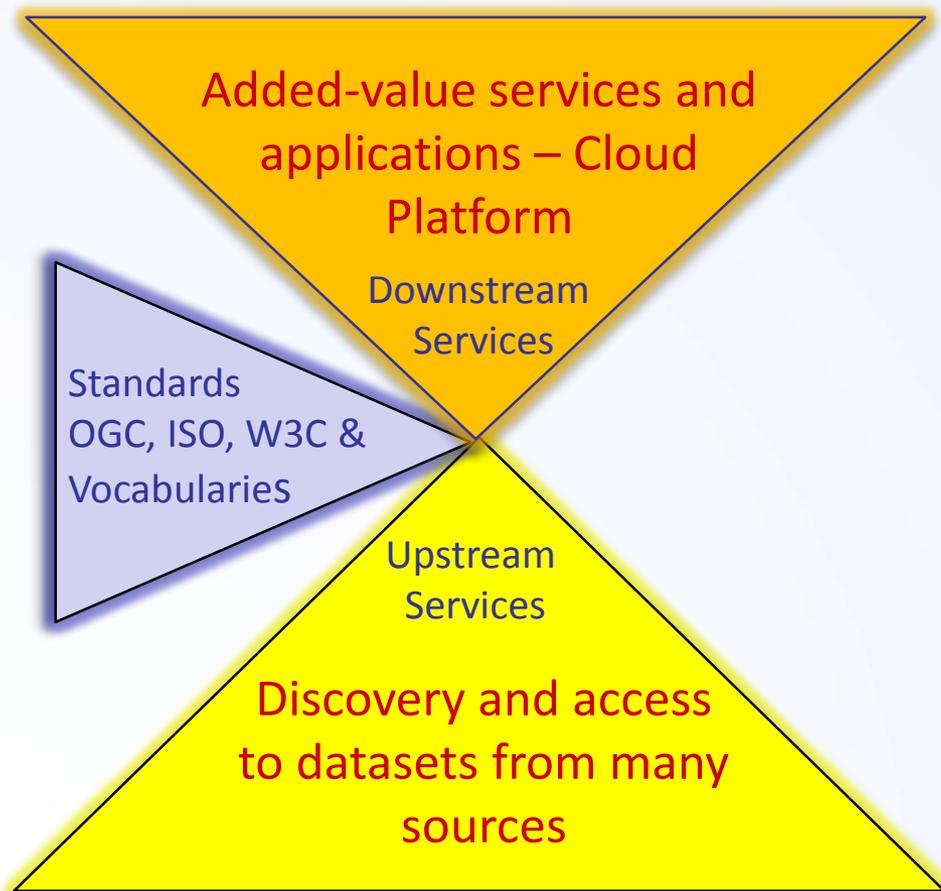


Cooperation with the EUDAT network of e-infrastructure providers



5 EUDAT members are partners of SeaDataCloud :
CINECA, CSC, DKRZ, GRNET and STFC

Leading concept for SeaDataCloud



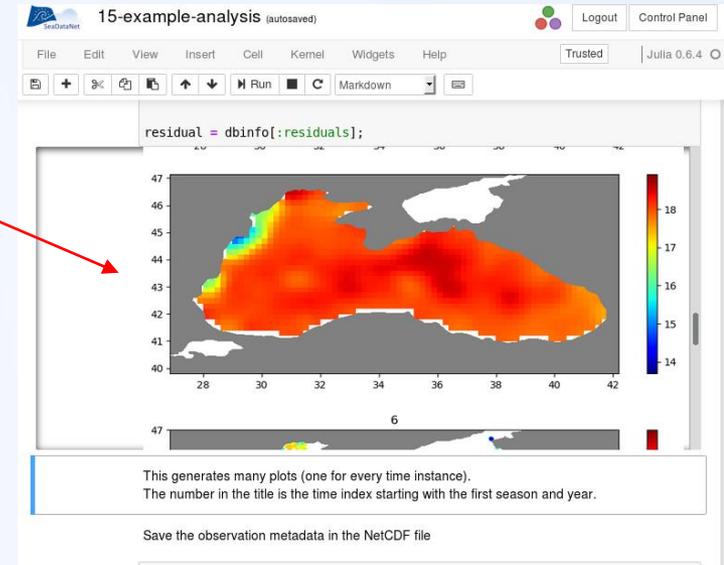
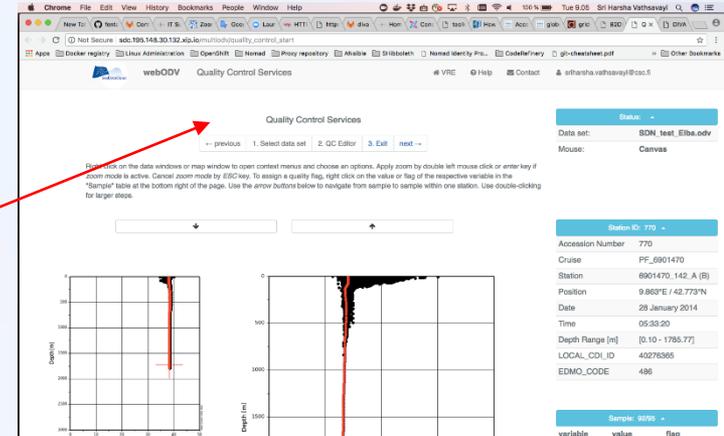
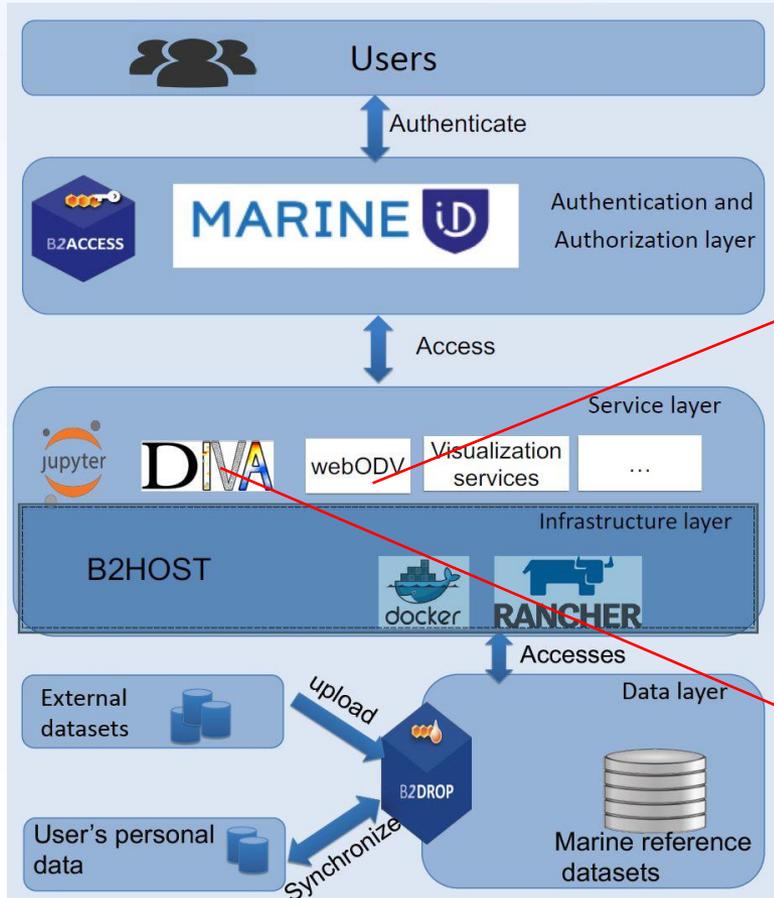
- Providing a cloud platform with common services for data pre-processing, subsetting, analyses, visualizations, publishing, DOIs...
- Applying common standards and interoperability solutions for providing harmonised data and metadata
- Providing harmonised discovery and access to data output from multiple sources, such as European research and monitoring data gathering, but also from other European and international data infrastructures



Virtual Research Environment (VRE)

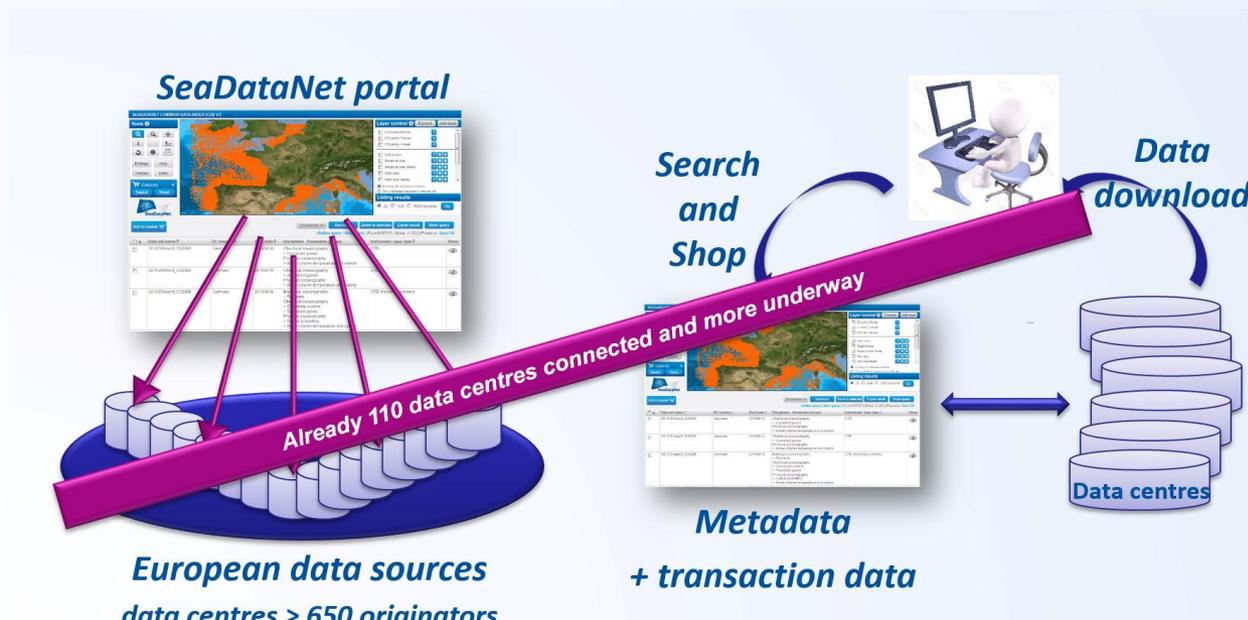
- Reviewing of existing VRE developments in Australia, USA and Europe, also as part of the Ocean Data Interoperability Platform (ODIP) activities
- Formulating a ToR for the VRE overall specifications
- Analysing and defining an overall architecture for the SDC VRE, considering EUDAT's generic services and embedding dedicated SDN services, custom and generic
- Analysing the workflow and specific functions required for supporting the SDC pilot: **generating a SeaDataNet Temperature & Salinity climatology**
- Developing VRE components and integrating these into the SDC VRE pilot

Overall architecture for the SDC VRE

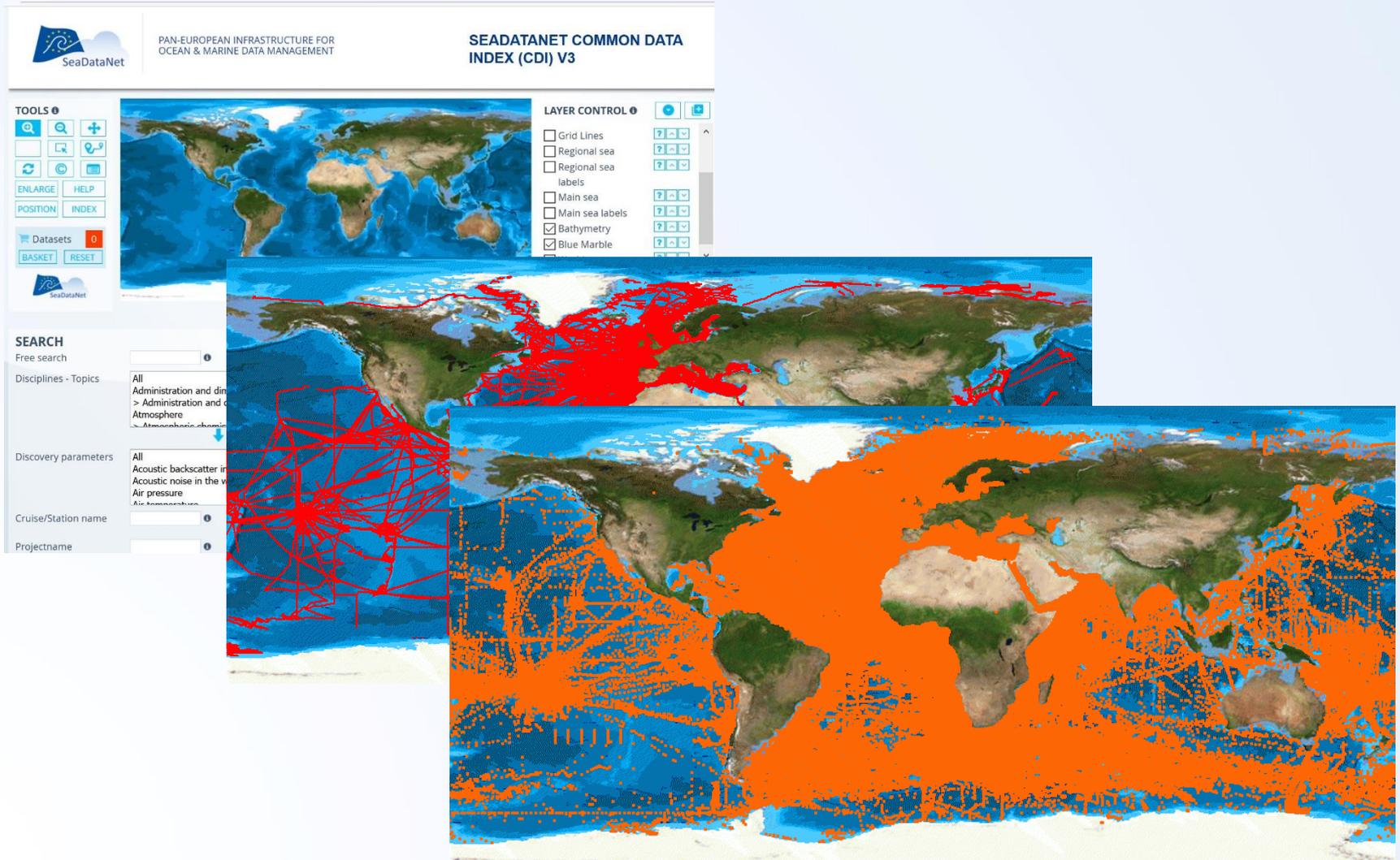


CDI Data Discovery and Access service

- It is one of the core services of the SeaDataNet infrastructure
- It provides a highly detailed insight and unified access to the large volumes of marine and oceanographic data sets managed by the distributed data centres
- It is a fine-grained index (ISO 19115 – ISO 19139) to individual data measurements (such as a CTD cast or moored instrument record)



CDI service with global coverage



PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

SEADATANET COMMON DATA INDEX (CDI) V3

TOOLS

ENLARGE HELP
POSITION INDEX

Datasets 0
BASKET RESET

SEARCH

Free search

Disciplines - Topics

All
Administration and d...
> Administration and...
Atmosphere
Atmospheric observ...

Discovery parameters

All
Acoustic backscatter in...
Acoustic noise in the v...
Air pressure
Air temperature

Cruise/Station name

Projectname

LAYER CONTROL

Grid Lines
 Regional sea
 Regional sea labels
 Main sea
 Main sea labels
 Bathymetry
 Blue Marble

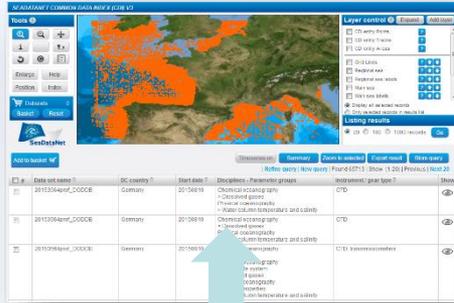
> 2.15 Million CDI entries for physics, chemistry, biology, geology and geophysics



SeaDataNet cooperation

- **Copernicus Marine Environmental Monitoring Services (CMEMS):** providing long-term archives and standards
- **Marine Strategy Framework Directive (MSFD):** providing infrastructure, standards and data collections for several indicators
- **Large ocean monitoring systems and their projects (EuroGOOS, AtlantOS, Euro-ARGO, JERICO-Next, ..):** providing standards and validation + long-term archiving services
- **EU projects, such as Upgrade BlackSeaScene, CaspInfo, Geo-Seas, Eurofleets ...:** adopting and adapting SeaDataNet standards and services for developing marine data management capabilities
- **Ocean Data Interoperability Platform (ODIP):** exploring and demonstrating common standards and interoperability with leading data management infrastructures in USA and Australia
- **GEOSS - EuroGEOSS:** Maintaining the GEOSS portal with SeaDataNet in-situ data collections from large community of European data holders (> 100 data centres; >600 data originators)
- **European Marine Observation and Data Network (EMODnet):** developing a network of thematic portals for data and data products

CDI service contributes to many portals



Data discovery and access



> 110 data centres



NODCs; HO; GEOs; BIOs; ICES; PANGAEA

≈ 650 European data originators

CDI Data Discovery and Access service

GEOSS portal

IODE ODP portal

Aggregated collection

Black Sea portal

Caspian portal

Regional subsets

Geo-Seas portal

Thematic portals



Bathymetry

Physics

Chemistry

Geology

Biology

Installed base of CDI nodes (> 110)



Issues with current CDI service

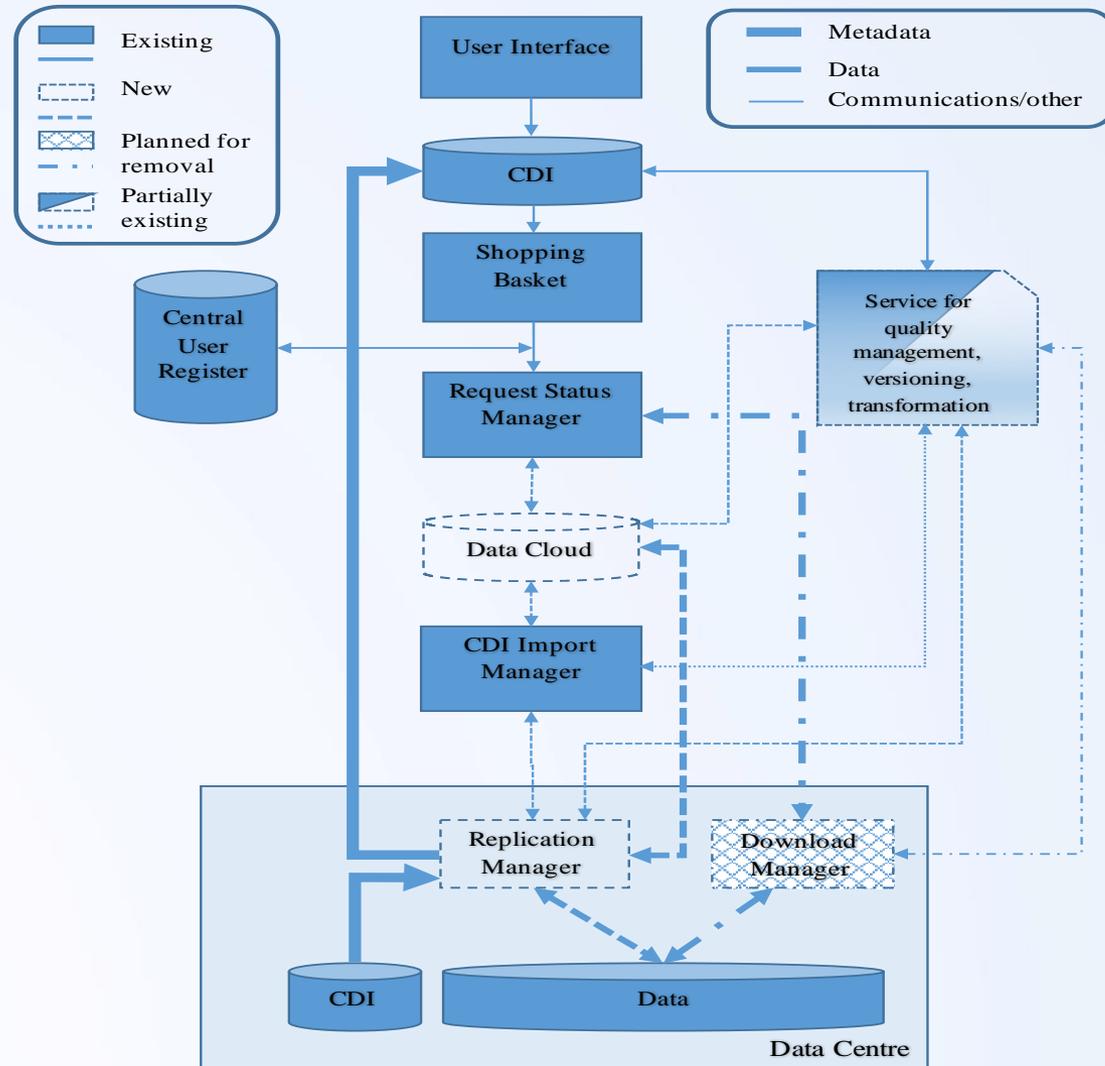
- **performance for users:** CDI data access service interacts with the distributed data collections and databases at the connected data centres.
 - user can submit a shopping basket with requests for data from multiple data centres.
 - user must await the automatic data preparation by each of these data centres
 - user must download resulting data sets through the RSM as packages directly from each data centre, which implicates multiple download transactions
- **performance for users:** data centres are not always online, operational and have different machine capacities which might give extra delays
- **quality issues:** concerning formats of data files (ODV + NetCDF) and their consistency with CDI metadata.
- **installation and configuration** of the Download Manager software can be challenging due to different configurations, firewalls etc., which in practice results in having different versions installed



Upgrading the CDI service using the cloud

- To configure and maintain a **CLOUD buffer** to host **copies of unrestricted data resources**
- Exchange by dynamic **replication** from the individual data centres, following their updating of the CDI catalogue service
- In the cloud buffer:
 - checking overall quality of metadata and data, as extra check on top of local QA-QC by data centres
 - checking integrity of data files and metadata relations.
 - results of checks to be reported back to data centres for corrections
- Include transformation services for converting data sets to SeaDataNet ODV and NetCDF formats and relevant INSPIRE data models.
- Introduce versioning of metadata and data as part of provenance

New CDI service architecture





New GUI on top of a new system

- Faster, using **Elastic Search** for search and indexing, and **GeoServer** for latest mapping technology
- Easier, including full text search next to controlled terms
- Modern design with large map and sliding windows
- **MySeaDataNet** integrated in GUI for customized services, such as SSO, RSM access, search profiles, prepared for VRE data pooling
- Developed and refined as prototype (<http://sdc.maris.nl>) interacting with users and data managers, following the SeaDataCloud Training Workshops in June 2018



PAN-EUROPEAN INFRASTRUCTURE FOR
OCEAN & MARINE DATA MANAGEMENT

DOWNLOAD THE ORDER ✕

Order ready to download

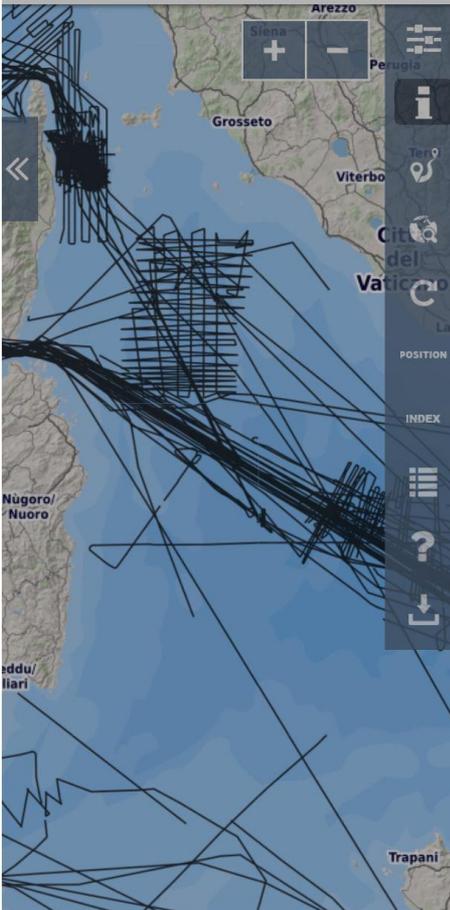
Order number 17501
 Last update 11/4/2018 11:20:59 PM
 Input date 11/4/2018 11:20:59 PM
 Organisation MARIENE INFORMATIE SERVICE MARIS B.V.
 Contact name D.m.a. SCHAAP

Unrestricted

[↓ RECORDS \(1.05 MB\)](#)

[↓ METADATA CSV FILE \(1.19 KB\)](#)

ello D.m.a. SCHAAP My CDI ▾ DATASET BASKET 0 



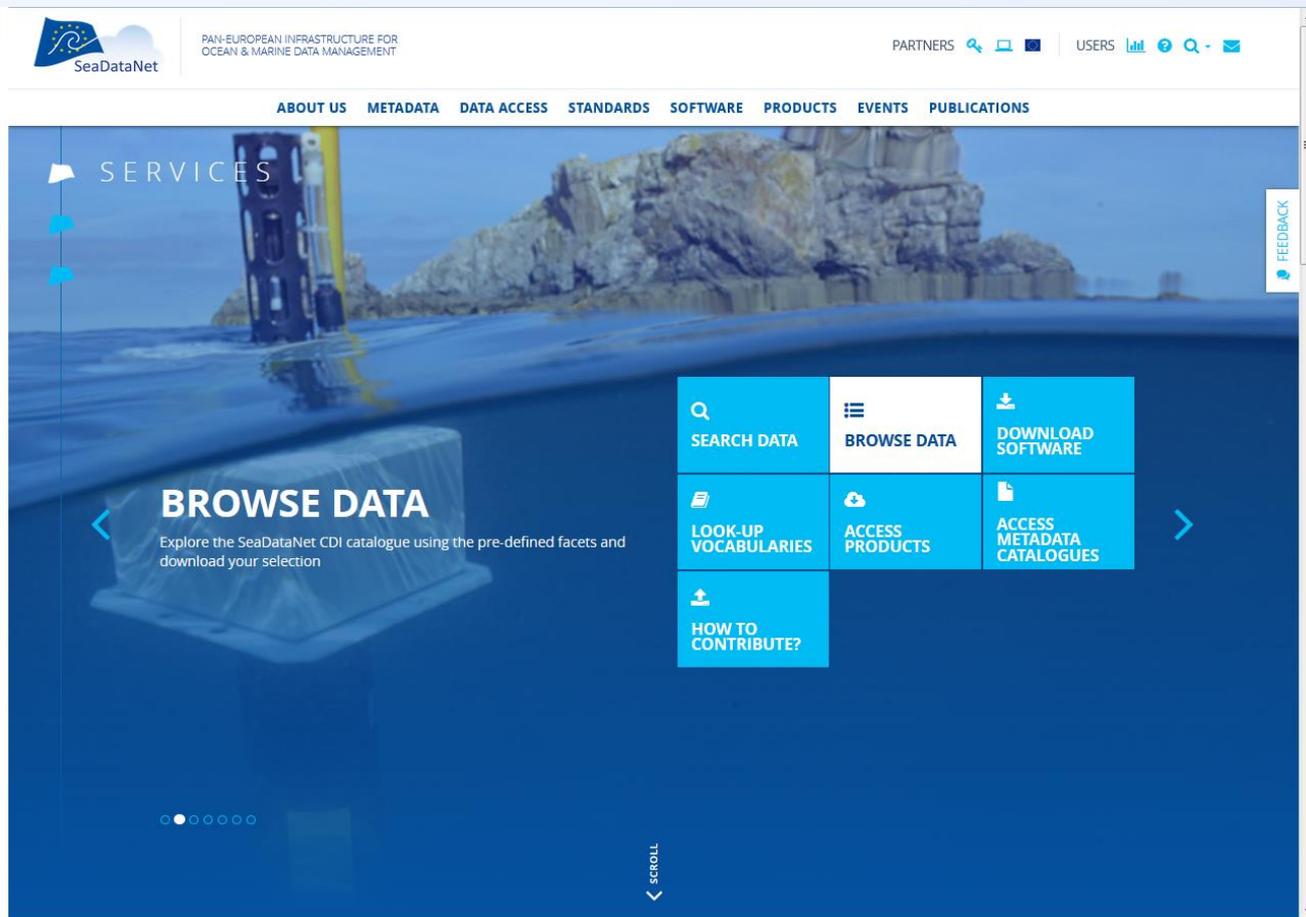
STANDING ORDERS | ORDERS HISTORY | SAV

ORDER DETAILS

Stukje tekst over de betekenis van de kleuren conse
 auctor. Phasellus sagittis purus nec augue malesua
 dolor quis volutpat. Nulla pharetra lectus ipsum, sit

Waiting for processing  Approval pending

ID	Name	Unrestricted
17501	IMDIS2018 test	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div> 3
17500	ARGO floats	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div> 25
17499	IMDIS2018 test CTDs	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div> 100
17498	IMDIS2018 test	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div> 3
17493	Bathymetry order	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div> 2
17492	Test ARGO op NetCDF	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div>
17491	Test ARGO floats	<div style="display: flex; width: 100px; height: 15px; background-color: #ccc;"><div style="width: 30px; background-color: #8bc34a;"></div><div style="width: 30px; background-color: #2196f3;"></div><div style="width: 40px; background-color: #9e9e9e;"></div></div> 7



The screenshot shows the SeaDataNet website homepage. At the top left is the SeaDataNet logo and the text "PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT". To the right are links for "PARTNERS" and "USERS" with various icons. Below this is a navigation menu with items: "ABOUT US", "METADATA", "DATA ACCESS", "STANDARDS", "SOFTWARE", "PRODUCTS", "EVENTS", and "PUBLICATIONS". The main content area features a large blue background image of a boat's deck with a yellow crane. On the left, there's a "SERVICES" section with a white flag icon and two blue arrows. In the center, a "BROWSE DATA" section is highlighted with a white box, containing the text "Explore the SeaDataNet CDI catalogue using the pre-defined facets and download your selection". To the right of this is a grid of six blue buttons: "SEARCH DATA" (with a magnifying glass icon), "BROWSE DATA" (with a list icon), "DOWNLOAD SOFTWARE" (with a download icon), "LOOK-UP VOCABULARIES" (with a document icon), "ACCESS PRODUCTS" (with a cloud icon), and "ACCESS METADATA CATALOGUES" (with a document icon). Below these is a "HOW TO CONTRIBUTE?" button (with an upload icon). On the far right, there's a vertical "FEEDBACK" button. At the bottom center, there's a "SCROLL" button with a downward arrow. A progress indicator with five circles is visible at the bottom left of the main content area.

www.seadatanet.org