

SeaDataCloud – further developing the pan-European SeaDataNet infrastructure for marine and ocean data management

Dick M.A. Schaap (MARIS, The Netherlands), Michèle Fichaut (IFREMER, France),
On behalf of SeaDataNet communities

EGU 2017 – ESSI session

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 35 countries bordering the European seas

90s	Metadata directories MEDAR/MedAtlas
2002-2005	Sea-Search (FP5)
2006-2011	SeaDataNet (FP6)
2011-2015	SeaDataNet II (FP7)
2016-2020	SeaDataCloud (H2020)



PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

SeaDataNet

Portal with standards, tools, and services, both for users and data centres

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

PARTNERS USERS

ABOUT US METADATA DATA ACCESS STANDARDS SOFTWARE PRODUCTS EVENTS PUBLICATIONS

SERVICES

SEARCH DATA

Enter your own criteria to search and download data from SeaDataNet CDI catalogue

SEARCH DATA	BROWSE DATA	DOWNLOAD SOFTWARE
LOOK-UP VOCABULARIES	ACCESS PRODUCTS	ACCESS METADATA CATALOGUES

www.seadatanet.org



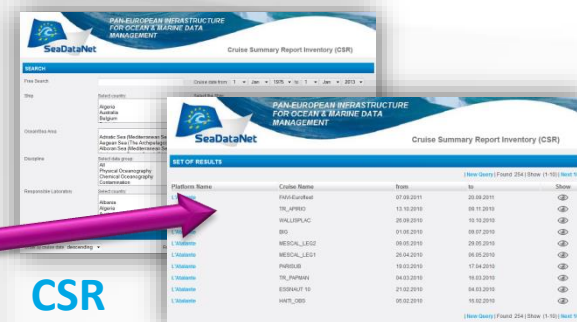
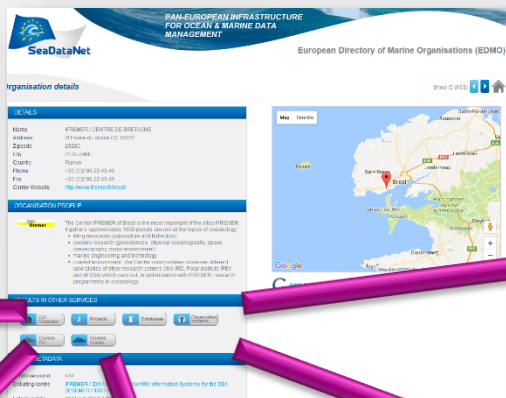
SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT



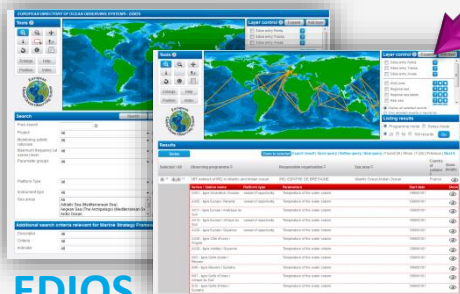
SeaDataNet metadata directories

EDMO Organisations

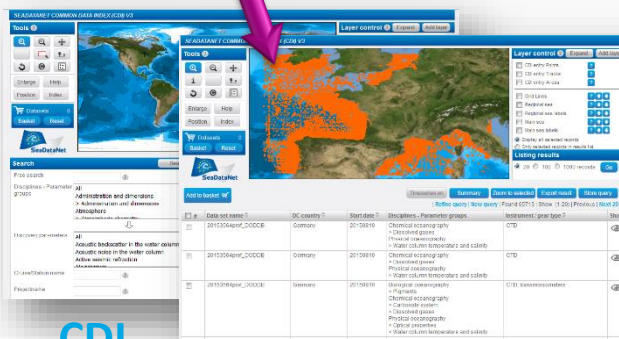


CSR Research cruises

EDMERP Projects

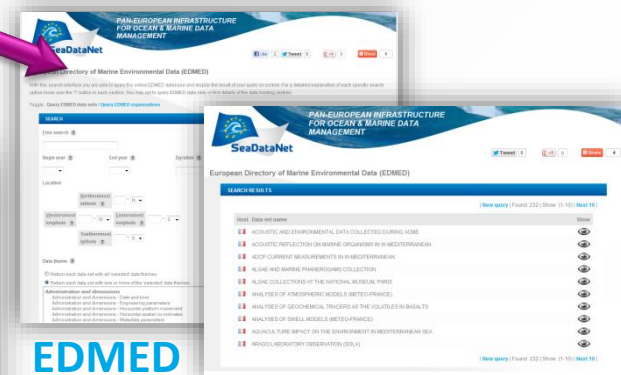


EDIOS Observing programmes



CDI

Data index



EDMED Data sets

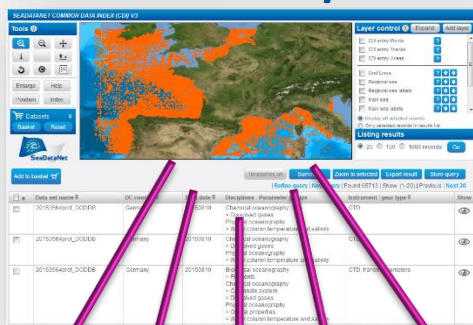


SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

CDI service for discovery and unified data access

SeaDataNet portal

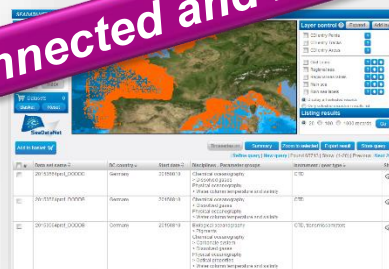


Search
and
Shop



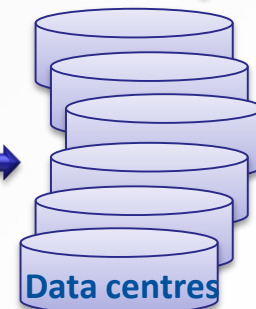
Data
add

Already 102 data centres connected and more underway



Metadata

+ transaction data

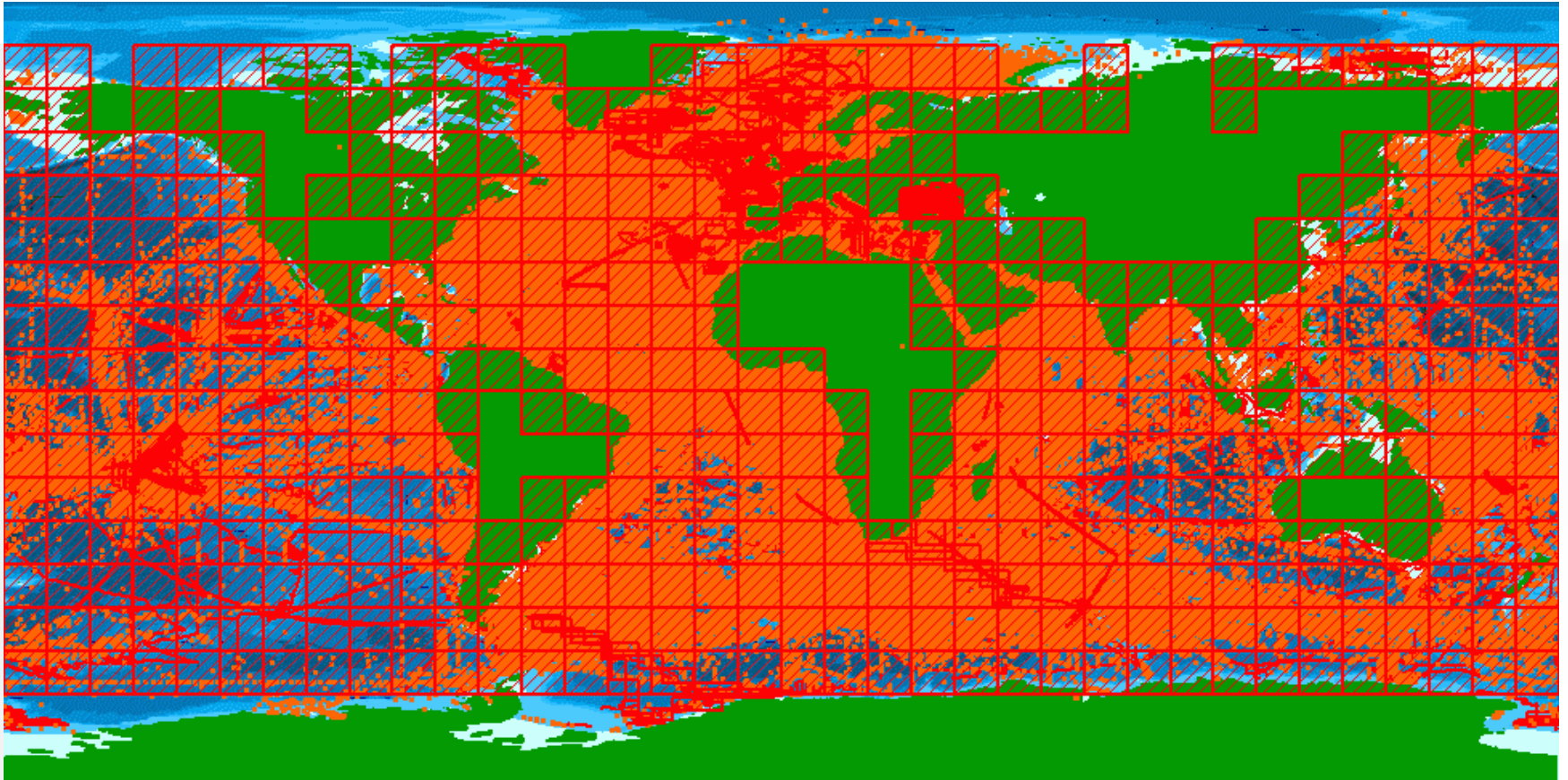


European data sources
data centres > 600 originators



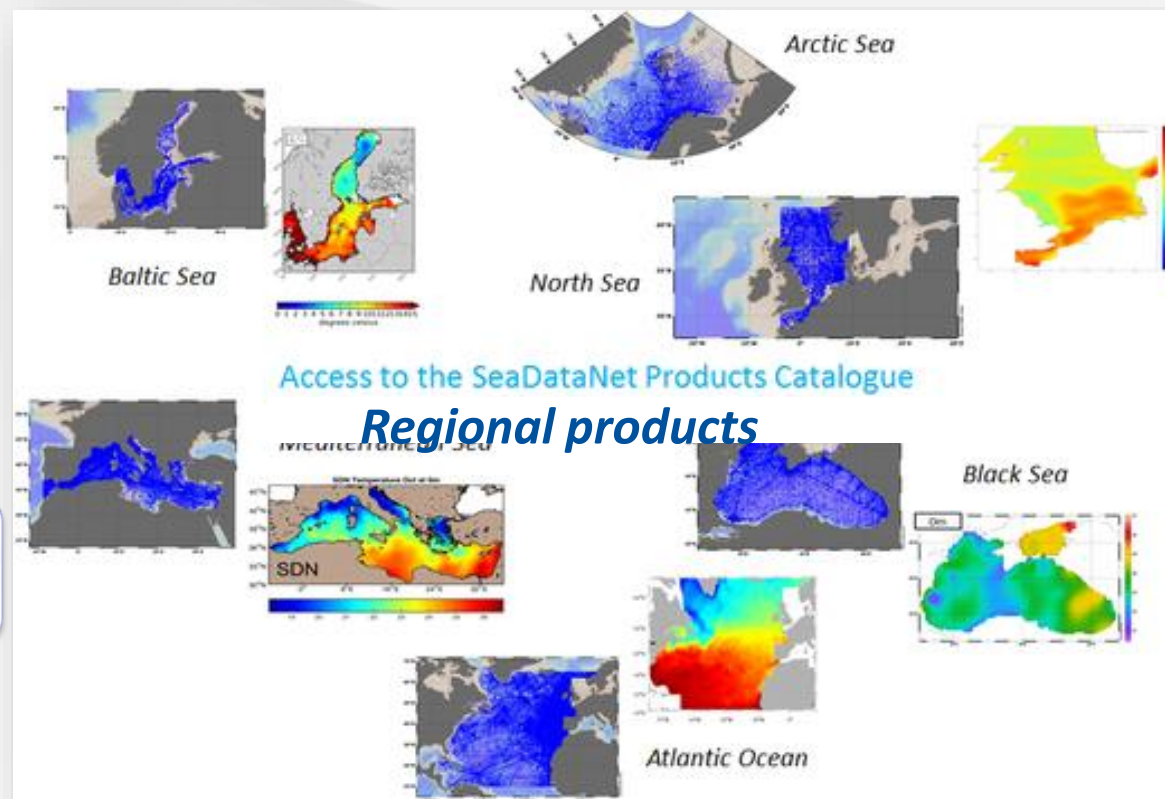
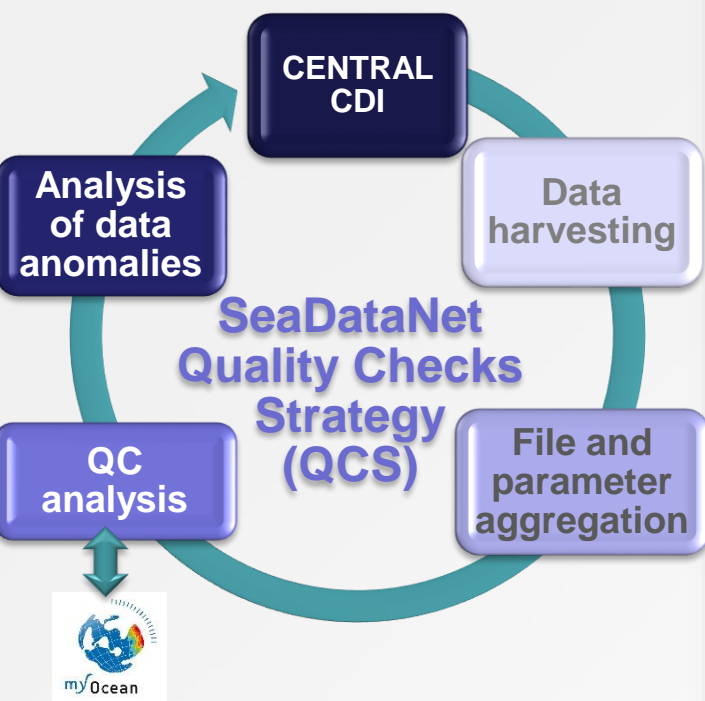
SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT



1.95 million CDI entries from ***34*** countries, ***102*** data centres and ***612*** originators for physics, chemistry, geology, geophysics, bathymetry and biology; from ***1805 to 2017***; ***87.6%*** unrestricted or under SDN License

SeaDataNet products

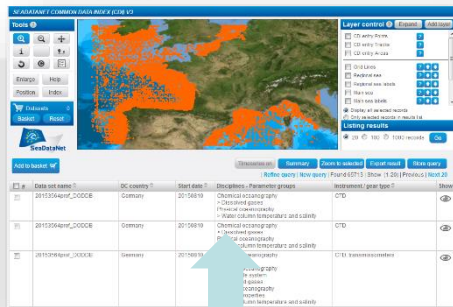


Aggregated datasets and climatologies
Improvement of the data quality



SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT



Data discovery
and access



> 100 data centres



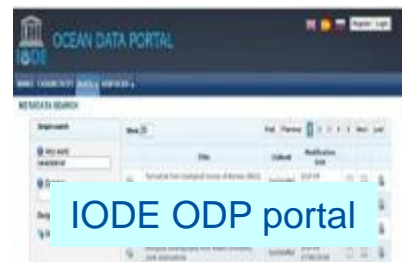
NODCs; HO; GEOs; BIOS; ICES; PANGAEA

≈ 600 European
data originators

CDI Data Discovery and
Access service



GEOSS portal



IODE ODP portal

Aggregated collection



Black Sea portal



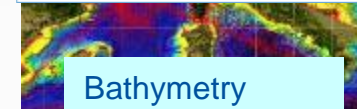
Caspian portal

Regional subsets



Geo-Seas portal

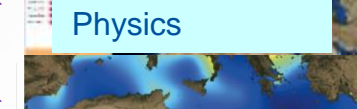
Thematic portals



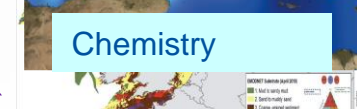
Bathymetry



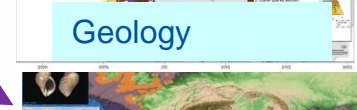
Physics



Chemistry



Geology



Biology

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
- November 2016 start of H2020 **SeaDataCloud** project for further developing SeaDataNet infrastructure and associated standards: 10 Meuro, 61 members, 32 countries, 4 years



SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

SeaDataCloud – general challenges

- It is about updating and further developing standards
- It is about improving and innovating services & products
- It is about adopting and elaborating new technologies
- It is about giving more attention to users and putting the user experience in a central position
- Moreover, it is about implementing a strategic and operational cooperation between the SeaDataNet consortium of marine and ocean data centres and the EUDAT consortium of e-infrastructure service providers



SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

SeaDataCloud – cooperation with EUDAT



European Computing Infrastructure

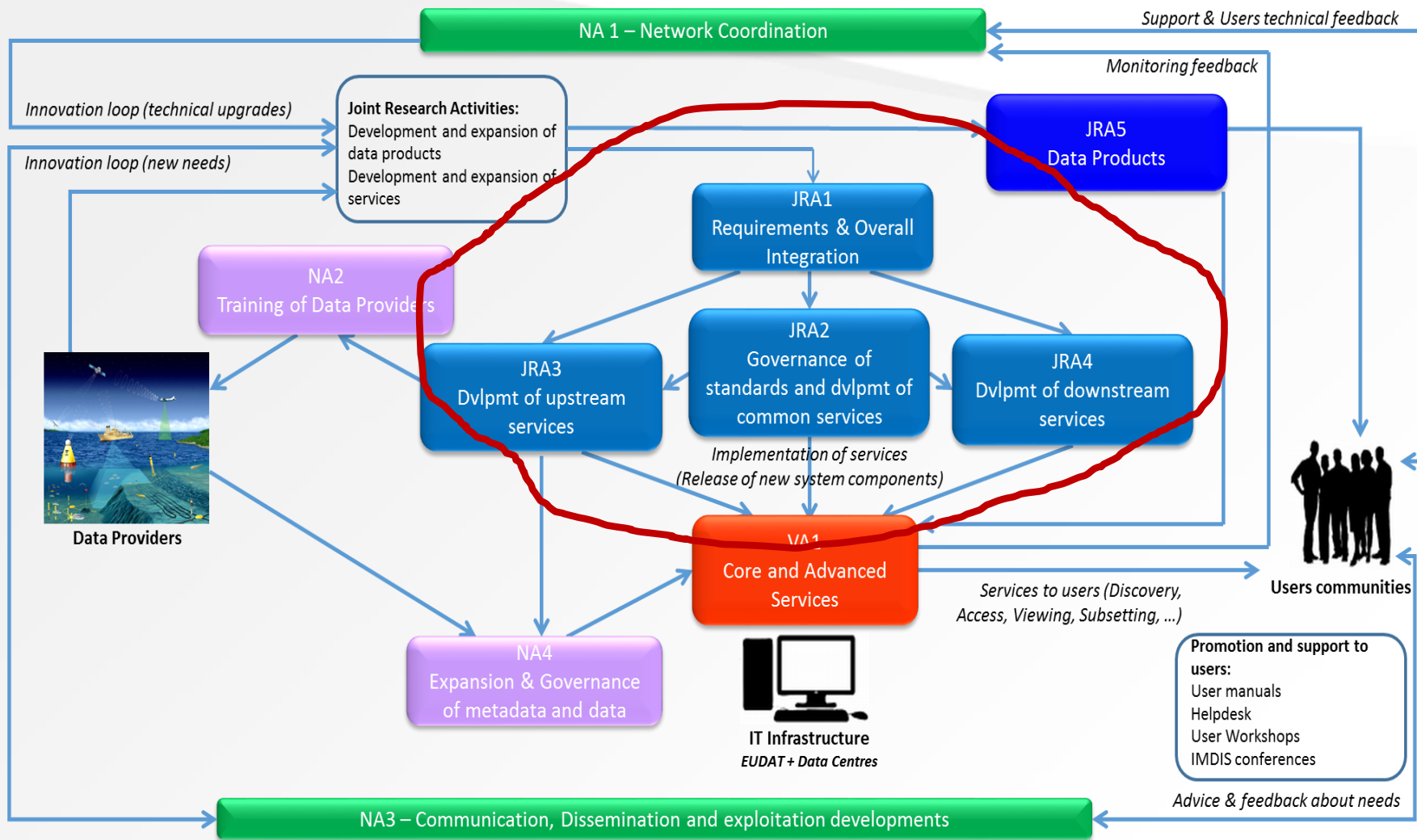


A consortium of high performance computing (HPC) / data centres, libraries, scientific communities, data scientists



SeaDataCloud Pert diagram

WP1 – Project Management



WP8 - Governance of standards and development of common services

- To develop further the SeaDataNet controlled vocabularies and related services,
- To analyse and deploy a pilot for adopting the Linked Data principle for SeaDataNet directories,
- To review and expand the SeaDataNet data formats for achieving INSPIRE compliance,
- To integrate the SeaDataNet AAI services with GEANT/eduGAIN and social networks,
- To upgrade the SeaDataCloud monitoring service.

WP9 - Developments of upstream services

- To upgrade the CDI Data Discovery and Access service making use of the cloud,
- To develop an online SWE ingestion service for operational observing systems,
- To expand SeaDataNet capability for handling different data types,
- To integrate external datasets from international programmes and organisations,
- To develop a solution for a coordinated distributed DataCite DOI minting service.

WP10 - Developments of downstream services

To expand the range of services of the SeaDataNet infrastructure by specifying, developing and deploying a **Virtual Research Environment (VRE)** with advanced e-services to facilitate individual and collaborative research by using, handling, curating, quality controlling, transforming and processing marine and ocean data into value-added analyses, harmonised data collections, and data products which can be integrated, visualised and published using OGC and high level visualisation services.

Added-value services
and applications

WP10
Downstream
Services

WP8
Standards &
Vocabularies

make it work!

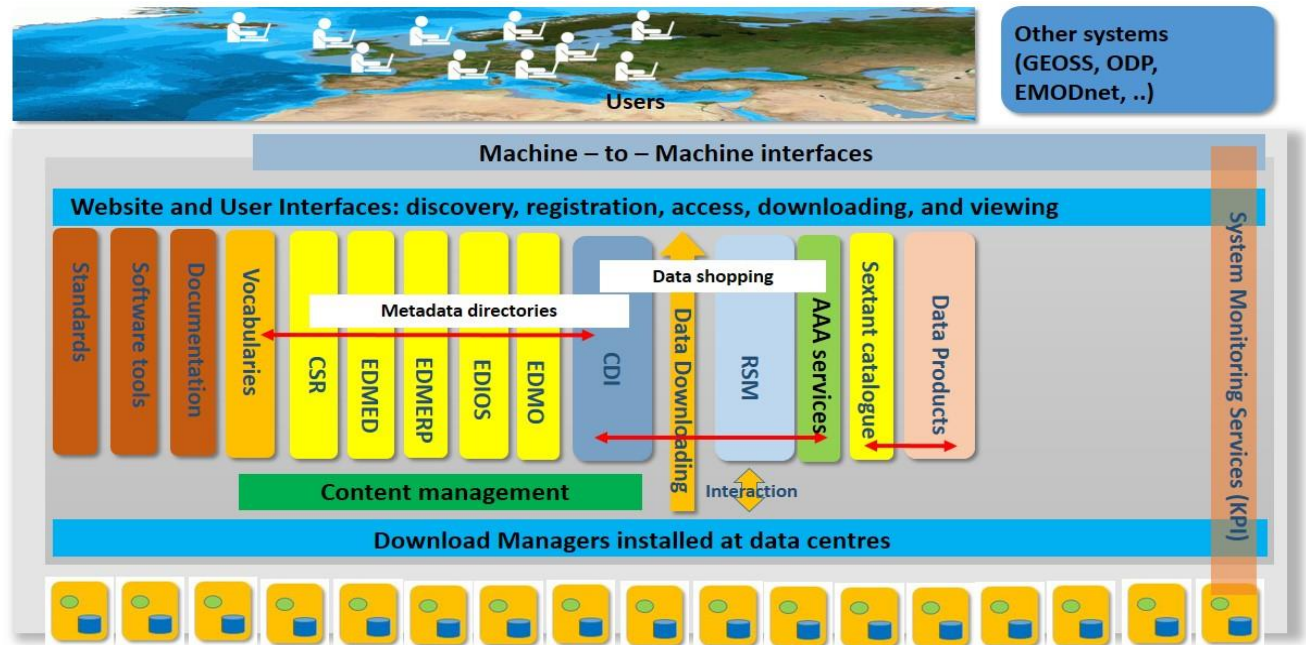
WP9
Upstream
Services

Discovery and access
to more datasets and
information

Upgrading the CDI service using the cloud

- To configure and maintain a cloud environment to host **copies of data resources**
- Exchange by dynamic **replication** from the individual data centres, following their updating of the CDI catalogue service
- In the cloud buffer:
 - checking possible duplicates
 - Checking overall quality of formats
 - Checking integrity of data files and metadata relations.
 - Results of checks to be reported back to data centres for amendments of their submissions and/or local configurations for mapping data and metadata.
- Include **transformation services** for converting data sets to other required output formats such as SeaDataNet NetCDF and relevant INSPIRE data models.

Present SeaDataNet architecture



Proposed upgraded architecture with data replication, advance services and VRE in the cloud

