

neosc | Blue-Cloud2026

# Third Federation Workshop

---

5-6 November 2025, Brussels [Belgium]



Funded by  
the European Union

Target audience: Scientific partners, data repositories, RIs representatives

Chair: Sara Pittonet Gaiarin, Trust-IT Services

14:30 - 15:00	<b>Towards a network of Trusted Digital repositories: opportunities for the Blue Data infrastructures from the EDEN and FIDELIS projects</b> <b>Jan Meijer</b> (Sikt), <b>Roxanne Wyns</b> (KU LEuven)
15:00 - 16:15	<b>Marine thematic services for EOSC</b>  Overview of Blue-Cloud 2026 Virtual Labs & Workbenches, followed by Short demos presenting the Blue-Cloud Virtual Labs, the Workbenches and their services for the audience gathered for the EOSC Symposium, with insights from the real-case scenario ideas and challenges encountered in the recent Blue-Cloud 2026 hackathon, followed by collaborative discussions.  <b>Nydia Catalina Reyes Suarez</b> (OGS), <b>Francesco Palermo</b> (CMCC), <b>Simona Simoncelli</b> (INGV), <b>Joao Paulo do Nascimento Vitorino</b> (IH), <b>Sebastian Mieruch</b> (AWI), <b>Julien Barde</b> (IRD)
16:15 - 16:30	Q&A & discussion

 eosC | Blue-Cloud2026

# Blue-Cloud 2026 as a candidate node for the EOSC Federation

Introduction - Sara Pittonet Gaiarin, Blue-  
Cloud Coordinator

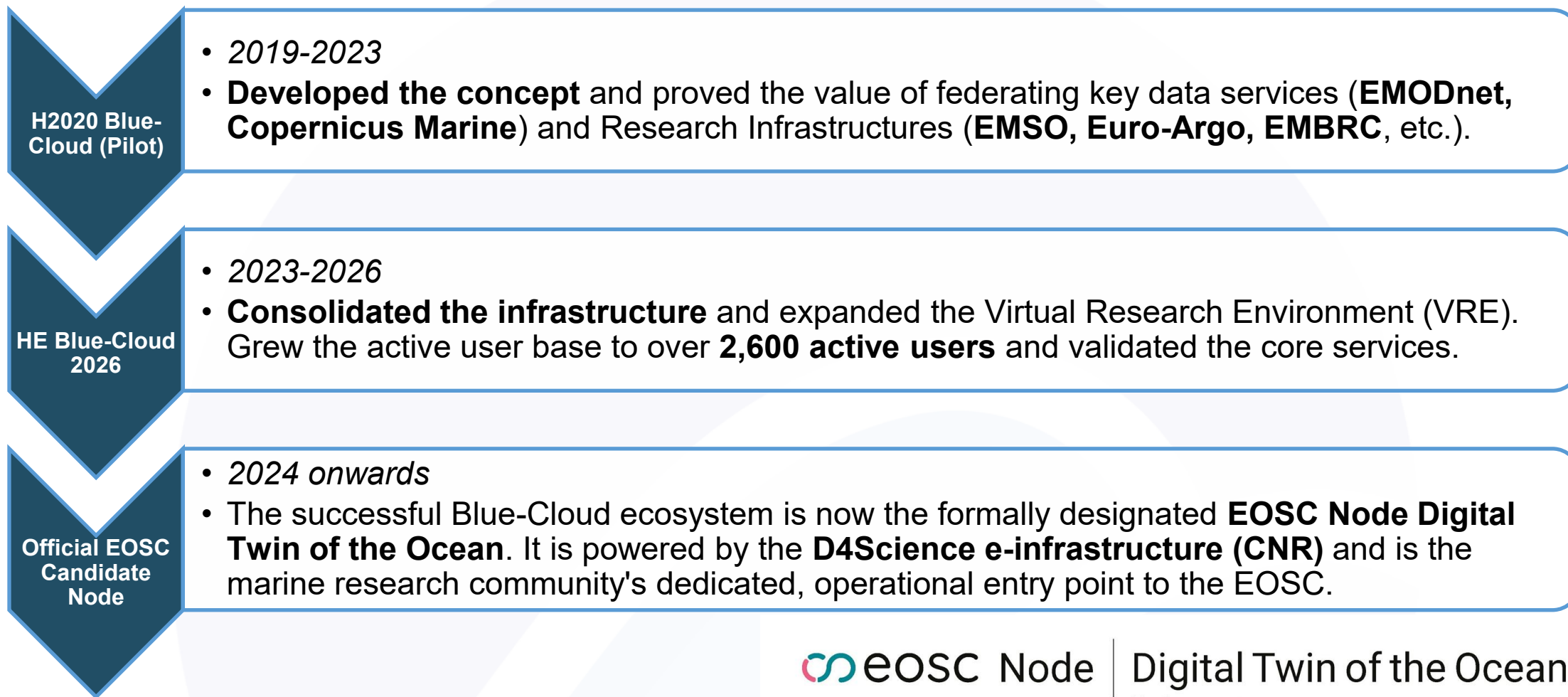
---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union **3**

Our Node is the **evolution of a proven ecosystem** built on six years of strategic investment and strong community engagement.





The Ocean doesn't work in silos — Why should we?" — *Nicolas Segebarth, DG RTD*

## The European Mandate: Unifying Digital Ecosystems

- **EOSC Federation:** The horizontal, trusted **System of Systems** for all European Research.
- **EU Mission: "Restore our Ocean and Waters"** – The clear, policy-driven societal objective.
- **European DTO / EDITO:** The official, operational infrastructure for marine science.

## The Challenge

- Ensuring these powerful strategic pillars are technically and operationally connected to prevent new, larger silos between the DTO Core and the long-tail scientific community.

## Our Position

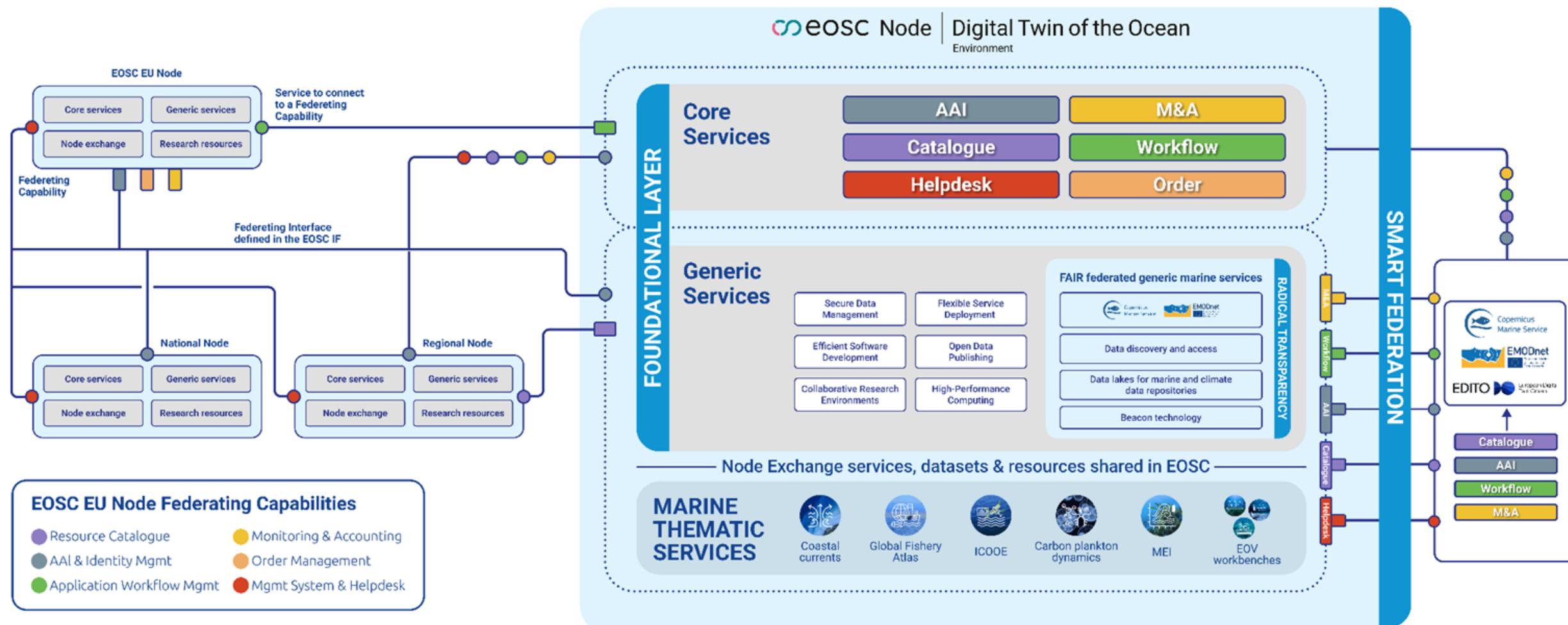
- The **EOSC Node Digital Twin of the Ocean** is the strategic and operational bridge. It promotes dedicated mechanisms to federate marine research capacity into the EOSC ecosystem.

**A Thematic Node for Empowered Researchers:** Blue-Cloud transforms marine research by providing an integrated environment built on EOSC principles

- **FAIR Federated Data Services:** We provide harmonised data collections with advanced subsetting, semantic enrichment, and provenance tracking, so researchers get exactly the data they need with full trust.
- **Integrated Virtual Research Environment (VRE):** Researchers have direct access to a suite of powerful tools within our Virtual Labs, including:
  - Interactive computing (Jupyter Notebooks, RStudio).
  - Reproducible workflows (Galaxy).
  - Containerization and application management pipelines.
- **EOSC-Ready Capabilities:**
  - Federated AAI: Full integration and alignment with EOSC AAI.
  - Resource Catalogue: Our service catalogue is structured for immediate federation with the EOSC Exchange (EEN).
  - Operations: Professional, FitSM-aligned Helpdesk and Monitoring.



**Researchers are empowered with trusted, interoperable tools and seamless access to data**



 **EOSC** | Blue-Cloud2026

**Towards a network of Trusted  
Digital repositories: opportunities  
for the Blue Data infrastructures  
from the EDEN and FIDELIS projects**

Jan Meijer (Sikt), Roxanne Wyns (KU  
Leuven)

---

**5-6 November 2025, Brussels [Belgium]**



Funded by  
the European Union



# EOSC EDEN

Enhancing Digital preservation strategies at  
European and National level

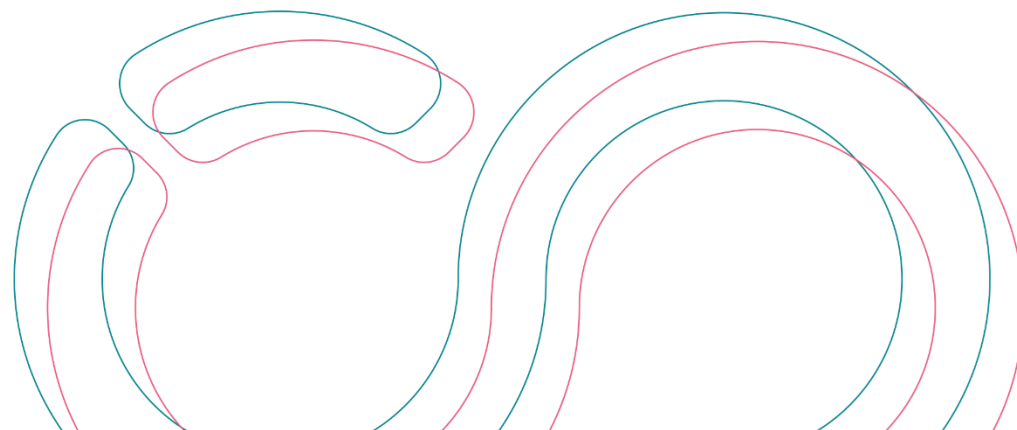
*Presenter: Roxanne Wyns*

*5 November 2025*



**Funded by  
the European Union**

Grant agreement 101188015





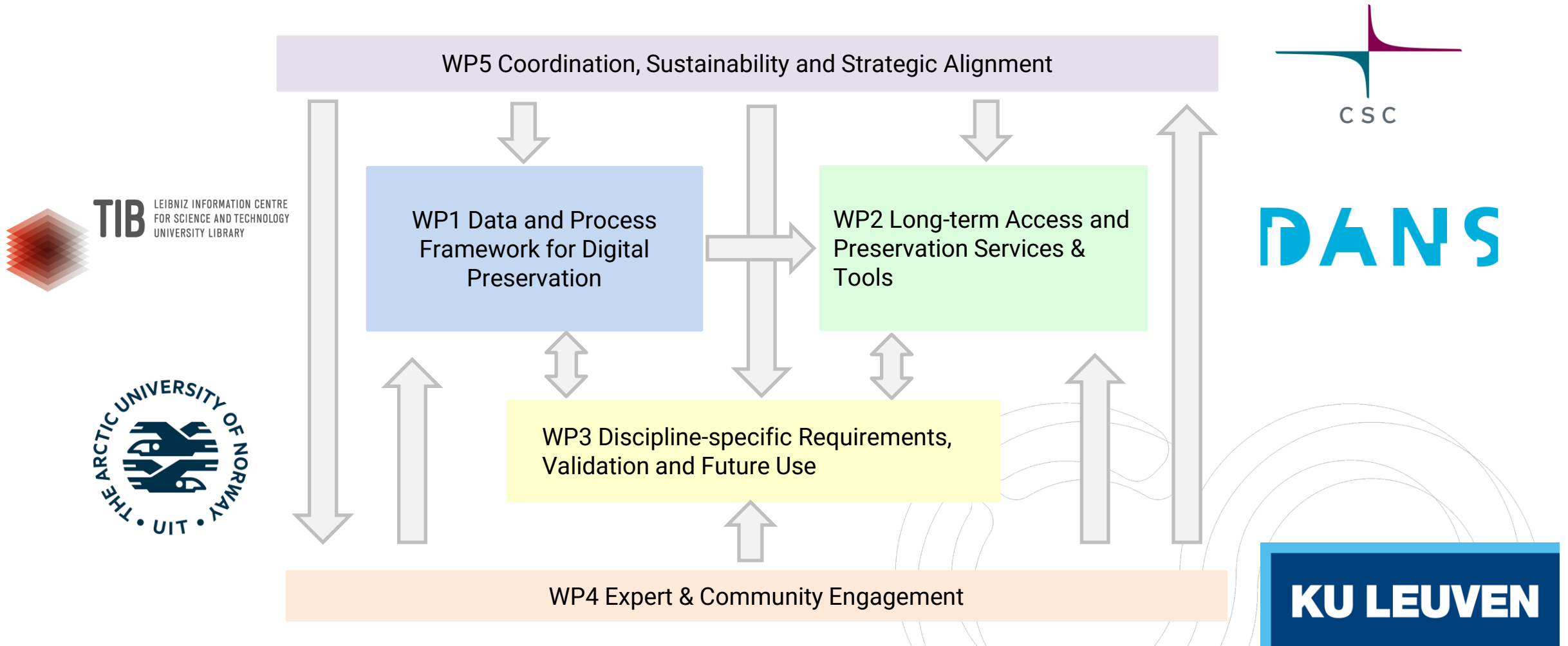
# EOSC EDEN Key Facts

Project full title:	EOSC EDEN - Enhancing Digital preservation strategies at the European and National level
Funding:	Horizon Europe, Grant agreement ID: 101188015
Project type:	Research and Innovation Action (RIA)
Budget:	8M €
Coordinator:	CSC – IT Center for Science, Finland
Project start and end date:	1 Jan 2025 - 31 Dec 2027

# EOSC EDEN Objectives

- **Objective 1:** To establish a general **framework and practices** to support the creation of curation, digital preservation, and access strategies in Europe
- **Objective 2:** To **enrich EOSC with tools** to store and access digital data for long periods, automate and federate certain specialised curation and preservation tasks
- **Objective 3:** To increase **adoption** of curation, digital preservation and access practices within different **scientific disciplines**
- **Objective 4:** To boost the **data curation and quality in Europe**
- **Objective 5:** To identify and consolidate a **network of repositories and archives for digital preservation** within EOSC in collaboration with the FIDELIS project (HORIZON-INFRA-2024-EOSC-01-03)

# Project Structure and Work Packages



# Iterative Development

## Year 1

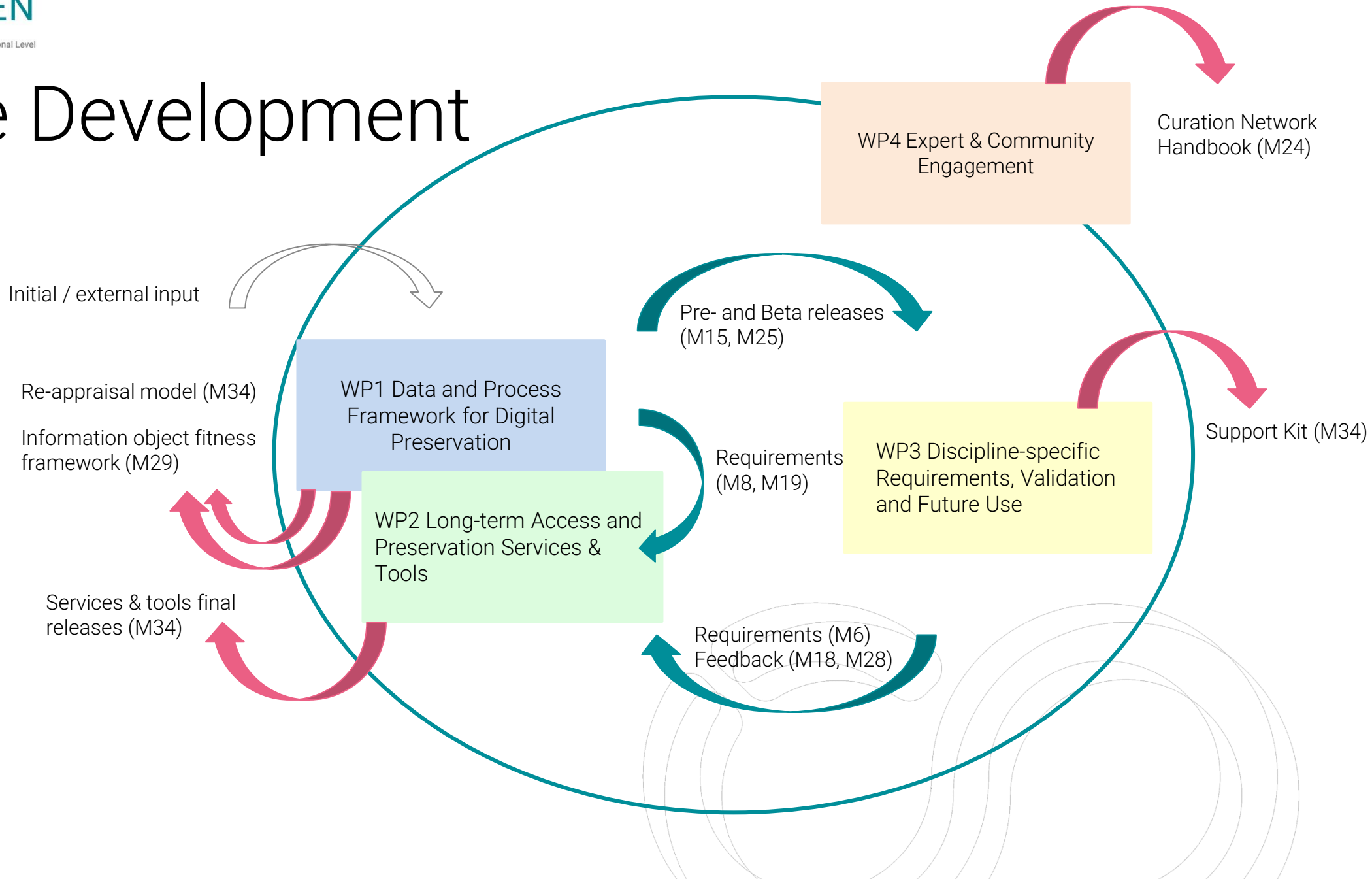
M1 Jan 2025  
M6 Jun 2025  
M12 Dec 2025

## Year 2

M13 Jan 2026  
M18 Jun 2026  
M24 Dec 2026

## Year 3

M25 Jan 2027  
M30 Jun 2027  
M36 Dec 2027

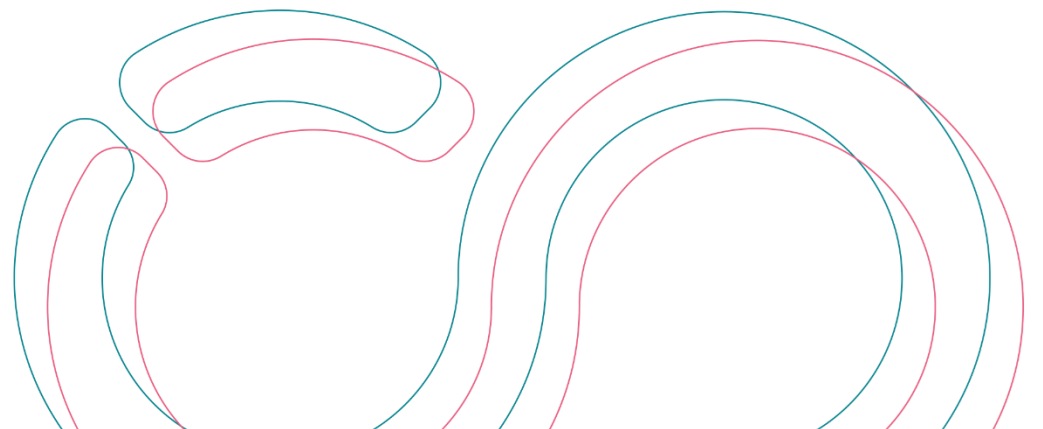


# EOSC EDEN

## Work Packages and Key Output

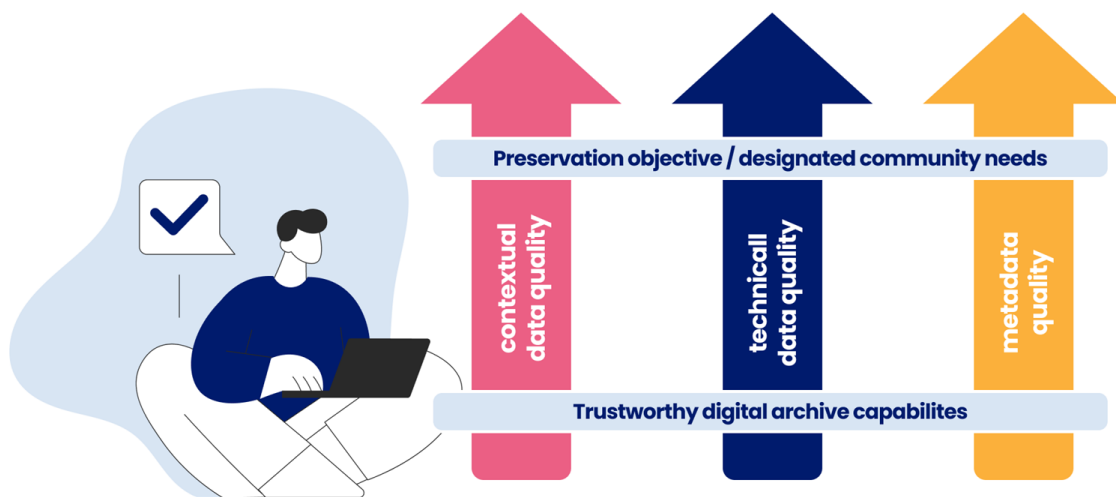


Funded by  
the European Union





# WP1 Data and Process Framework for Digital Preservation



Develops processes as requirements for implementation (WP2) and community adoption (WP3, WP4)



Existing practices for identification, selection and appraisal of data for digital preservation



Requirements in digital preservation processes for **re-use fitness** of digital objects



Framework to identify candidates to digital preservation based on use, benefit and quality



Model for re-appraisal points along data lifecycle

# Core Preservation Processes

**A Core Preservation Process (CPP)** is a specific action that every Trustworthy Digital Archive should undertake adequately - either directly or through its associated parties or services, in order to fulfil its digital preservation missions as evidenced in its preservation policy.

**Visualisation tool for CPP relationships:**  
<https://cpp.fd-dev.csc.fi/>

For more information: <https://eden-fidelis.eu/core-preservation-processes>

Leave your feedback via the [GitHub repository](#) or [webform](#).



Visualisation tool for CPP relationships: <https://cpp.fd-dev.csc.fi/>

Classification: Logical/Strategic ▾

View: Graph ▾

Select All

Deselect All

☒ Preservation Planning ☒ Dissemination ☒ Bit-level Preservation ☒ Generation of New Files ☒ Other Activities ☒ Lifecycle Management ☒ Characterisation

☐ Category Source Relations Only

Select All

Deselect All

Total Visible: 513

Procedural

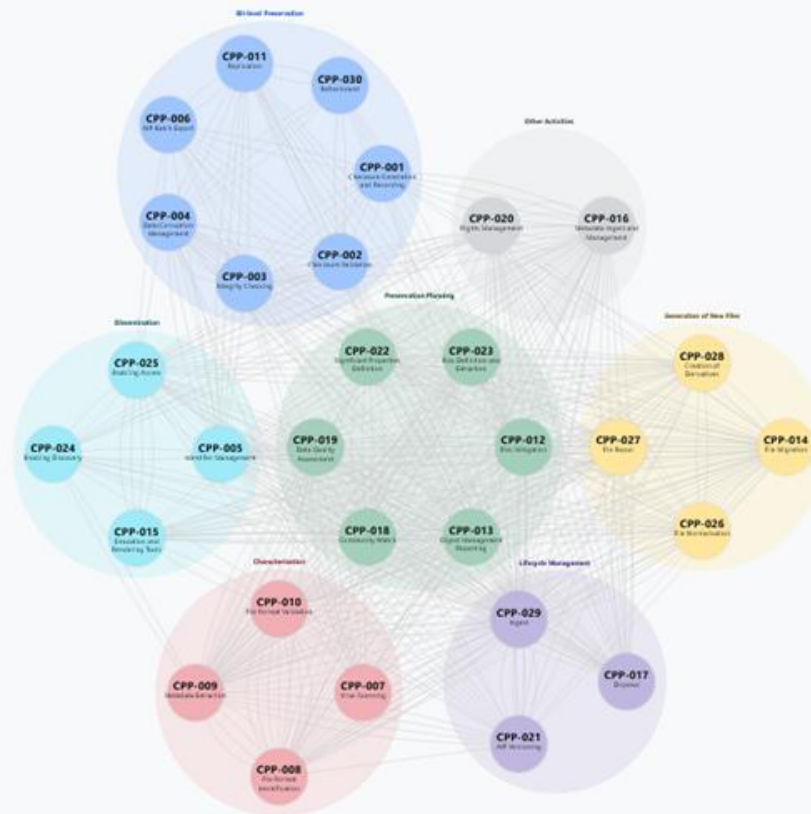
☒ Triggered By (85) ☒ Triggers (8)  
☒ Supplier (91) ☒ Customer (67)  
☒ Alternative To (2)

Dependencies

☒ Requires (84) ☒ Required By (80)  
☒ May Require (16) ☒ May Be Req. By (16)

Logical

☒ Affects (2) ☒ Affected By (2)  
☒ Facilitates (3) ☒ Facilitated By (3)  
☒ Affinity (36)  
☒ Not to be confused with (18)



Visualisation tool for CPP relationships: <https://cpp.f-d-dev.csc.fi/>

**Classification:** Logical/Strategic ▾ **View:** Graph ▾

Select All Deselect All ☒ Preservation Planning ☒ Dissemination ☒ Bit-level Preservation

Select All Deselect All

**Total Visible: 513**

**Procedural**

☒ Triggered By (85) ☒ Triggers (8)

☒ Supplier (91) ☒ Customer (67)

☒ Alternative To (2)

**Dependencies**

☒ Requires (84) ☒ Required By (80)

☒ May Require (16) ☒ May Be Req. By (16)


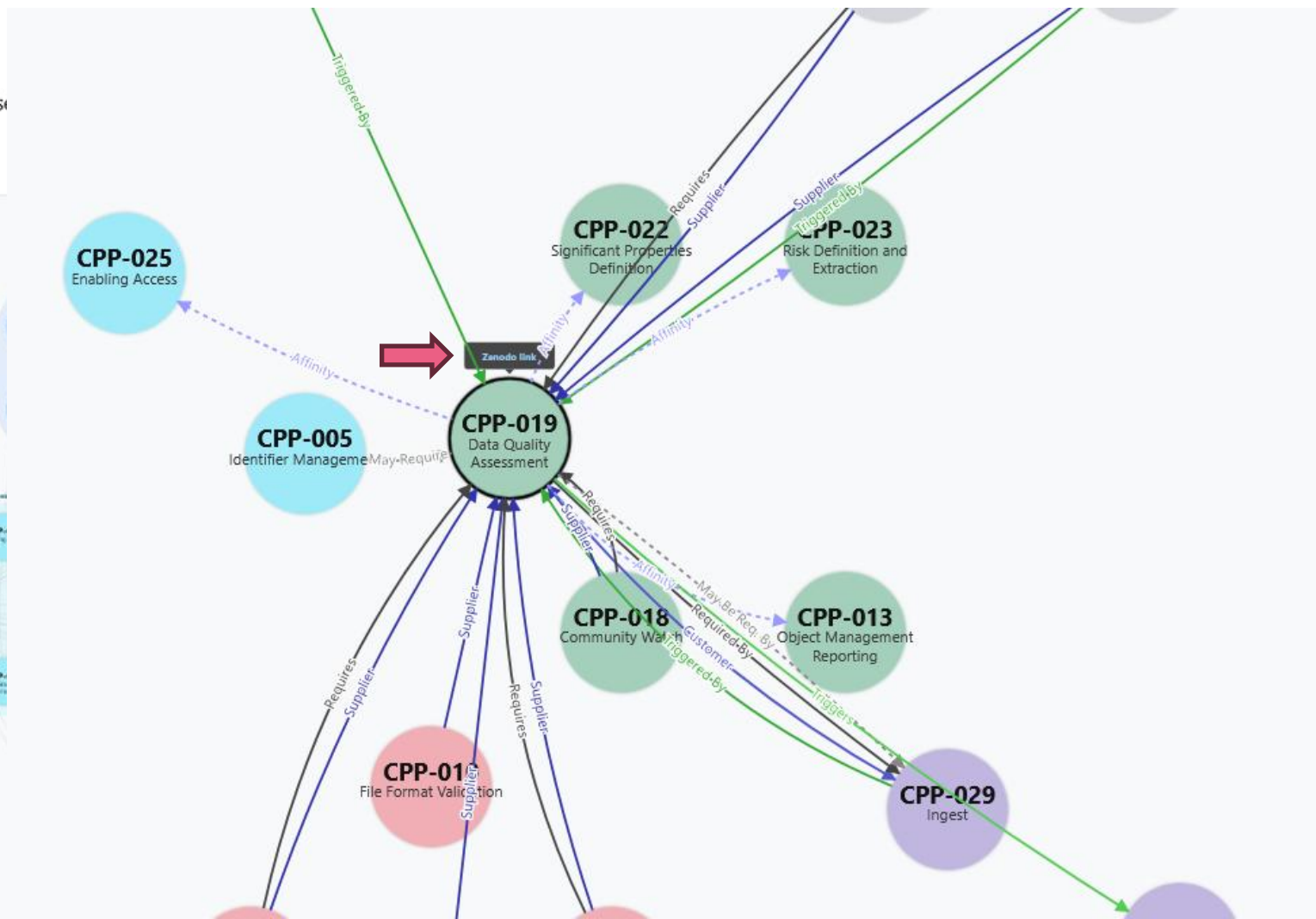
**Logical**

☒ Affects (2) ☒ Affected By (2)

☒ Facilitates (3) ☒ Facilitated By (3)

☒ Affinity (36)

☒ Not to be confused with (18)

# CPPs are described as a sequence of implementable steps, either by humans or by automation.

## Definition and scope

Data Quality Assessment refers to the systematic evaluation of *Objects* and their associated *Metadata* against predefined measures to ensure they meet the standards necessary for consumers' needs and continued access. The assessment typically covers several key dimensions, some of these are for example:

- **Authenticity:** The *Object* is what it purports to be (i.e. it has been created, modified and sent by the person purported to have done it at the date and time purported). The designated community must be able to trust that the data is real and credible and is managed by a trustworthy TDA. Sufficient information must exist to understand the *Object's* creation circumstances, provenance, and relationship to other content. In addition to integrity checks, the authenticity of the data is ensured by controlled changes through preservation actions and the *Provenance metadata*.

Page 2 of 12

- **Completeness:** The *Object* and the *Metadata* are complete. They do not have missing parts or links to targets outside the preserved *Object* which should remain accessible.
- **Consistency:** The *Object* is presented in applicable file formats or *Representations* with applicable metadata formats. Conflicting values in the *Metadata* should be avoided.
- **Relevance:** The data preservation is based on a predefined collection development policy (i.e. has a purpose of being preserved).
- **Structured:** The structure of the *Object* is described in the *Metadata*. Complex *Objects* are organised, including relationships between *Files*, proper sequencing of multi-part *Objects*, and the integrity of any embedded *Metadata* or links.
- **Understandability:** The information is understandable and meaningful for the designated community.
- **Validity:** The *Object* and *Metadata* are valid against the *File* and metadata format specifications and standards, and comply with all other predefined profiles and rules.

## Step-by-step description

No	Supplier	Input	Steps	Output	Customer
1	CPP-018 (Community Watch)	Preservation objectives	Based on preservation intent as defined by Community Watch, derive quality properties that will be extracted by other CPPs	Quality properties	
2		Quality properties	The TDA receives a defined set of quality properties and determines what data is required to create a quality assessment report. This triggers steps 3 to 8)	Specification of the data required for the assessment.	
3	CPP-008 (File Format)	Specification of the data required for the assessment	If quality properties concern file formats:	Technical quality report	

Page 4 of 12

	Identification)	File	Assess the file format against the preferred formats policy		
		File format identifier			
		Format policy - preferred formats			
4	CPP-010 (File Format Validation)	Specification of the data required for the assessment	If quality properties concern the validity of formats:  Assess the validity status.	Technical quality report	
		File			
		Validity status			



# Webinar on CPPs on 1/12/2025



 **EDEN**

**Exploring Core Preservation Processes (CPPs)**

Join the 1st EOSC EDEN Seminar!

01 Dec 2025  
11:00-12:00 CET

[Register now](#)

 **EDEN**

**What every Trustworthy Digital Archive should be doing: 30 Core Preservation Process Descriptions**

Exploring Core Preservation Processes (CPPs) – Join the 1st EOSC EDEN Seminar...

 1 December 2025

[Read more](#) 



Registration: <https://eden-fidelis.eu/form/eosc-eden-exploring-core-preserv>

# WP2 Long-term Access and Preservation Services & Tools



**Delivers reference implementations and supporting tools to integrate repositories and services into the EOSC Federation**



Registry of digital preservation services and tools



Publish machine-interoperable services for curation and preservation actions

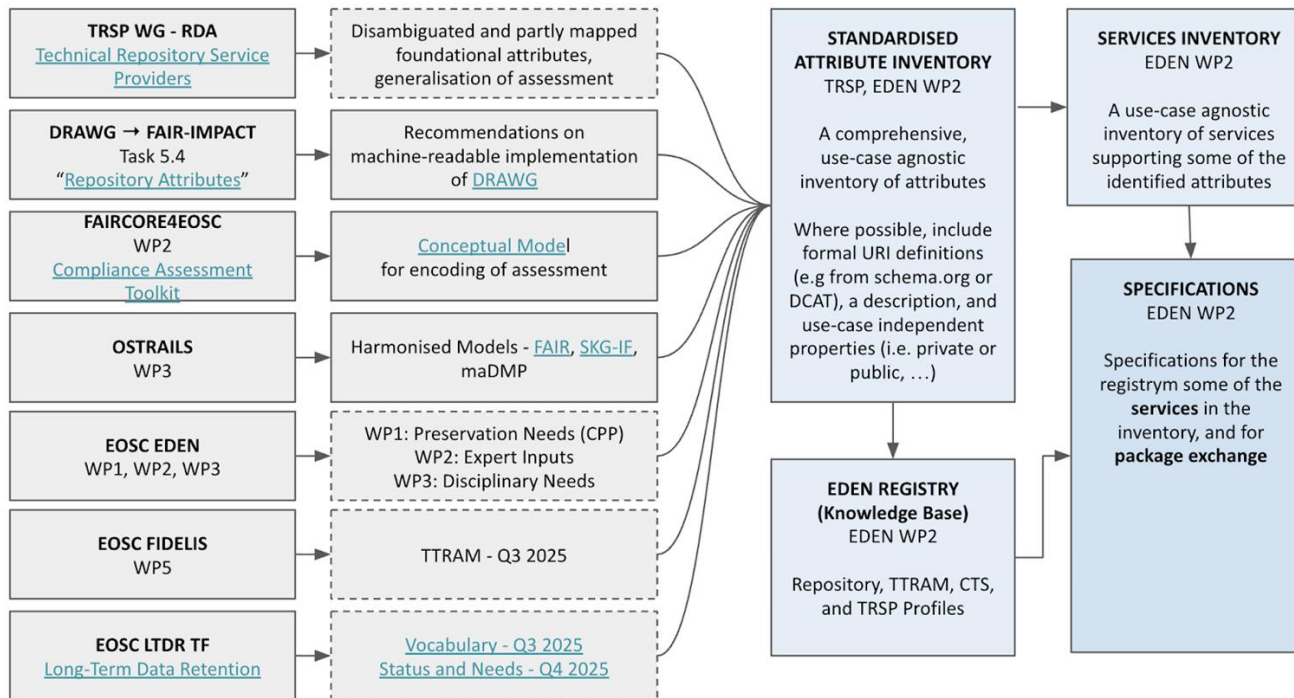


Identify standards and protocols for submit and exchange Information packages



Use case implementation and testing of services and tools

# Specifications and architecture for long-term access and preservation services



Overview of WP2 Methodology

- Interoperability
- Quality assurance of research outputs
- Rights & Ethics
- Services
- Attributes
- Requirements
- Specifications

For more information: <https://zenodo.org/records/17232536>

# WP3 Discipline-specific Requirements, Validation and Future Use



**Links the project to discipline and data-type specific needs and requirements**



**Iteratively provide discipline-specific, and cross-disciplinary / data type-specific requirements for digital preservation and data quality to WP1 and WP2**



**Validate and enhance the new digital preservation framework (WP1) and tools (WP2) via pilot testing**



**Provide a support-kit to empower discipline and data type-oriented communities to adopt, extend and use the new digital preservation framework (WP1) and fit-for-purpose tools (WP2)**

# Discipline Requirements and Needs

Applied several methods including:

- *desk-based mapping*
- *stand-alone interviews*
- *survey*

The main results is organized in 5 main components of the Re-use Fitness model:

- Contextual Data Quality
- Technical Data Quality
- Metadata Quality
- Trustworthy Digital Archive Capabilities
- Preservation Objective/Designated Community Needs.

For more information: <https://zenodo.org/records/15789261>



Social Sciences



UK Data Service



Earth & Environmental Sciences



Universität  
Bremen



Food Sciences



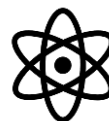
PREMOTEC



Climate Simulations



DKRZ  
DEUTSCHES  
KLIMARECHENZENTRUM



High-Energy Physics



Life Sciences and Bioinformatics



Swiss Institute of  
Bioinformatics



Linguistics





# WP4 Expert & Community Engagement



Support the other WPs in effectively engaging with communities and experts



Gather input and feedback during the development, testing and validation cycles



Increase awareness, knowledge and uptake of digital preservation actions and processes in the EOSC ecosystem



Identify and consolidate a network of service providers

# European Expert Curation and Digital Preservation Network

The EOSC EDEN Project will establish a **European Expert Curation and Digital Preservation Network** with representation from organisations, repositories (both generalist and specialist), collections, catalogues, and at various digital object type levels.

During the EOSC EDEN project, two curation network events will be organised for intensive learning and discussions with curation and digital preservation specialists.



*For more information on the 1<sup>st</sup> workshop in Leuven, Belgium:*

<https://eden-fidelis.eu/blog/1st-eosc-eden-curation-workshop-leuven-belgium-gathering-experts-build-european-curation-and>

# WP5 Coordination, Sustainability and Strategic Alignment



Project coordination and management incl. content work, finance & administration



Liaison with the External Advisory Board



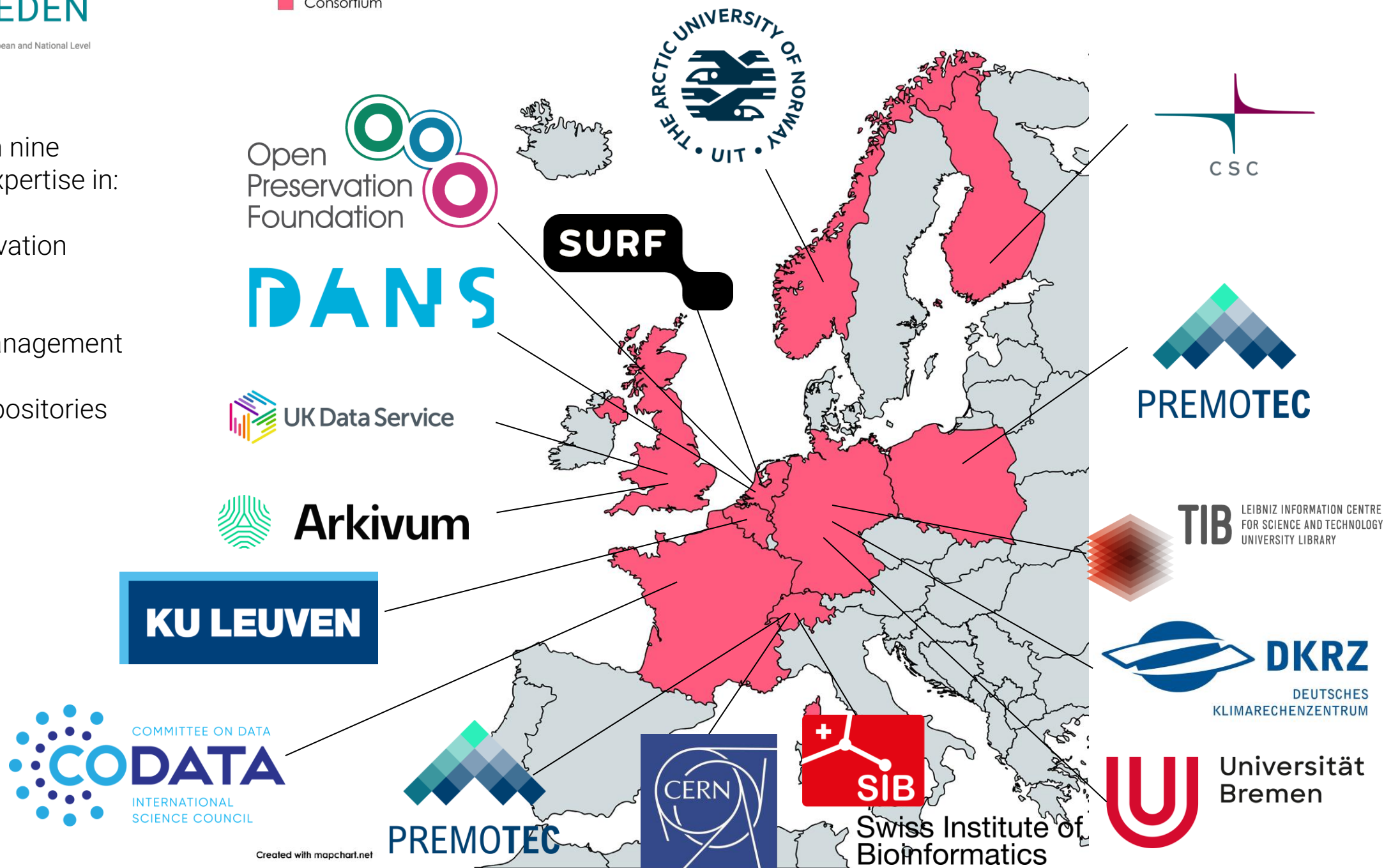
Strategic alignment with the EOSC Partnership and other relevant strategic initiatives



Sustainability and exploitation of the project results

16 partners from nine countries with expertise in:

- digital preservation
- data quality
- curation
- FAIR data management and services
- hosting of repositories and archives







[eden-fidelis.eu](https://eden-fidelis.eu)



[linkedin.com/company/eosc-eden](https://linkedin.com/company/eosc-eden)



[@eosc-eden.bsky.social](https://eden-fidelis.eu)



[@EOSC-EDEN](https://eden-fidelis.eu)



<https://eden-fidelis.eu/#newsletter>



[github.com/EOSC-EDEN](https://github.com/EOSC-EDEN)



[EOSC EDEN Zenodo Community](https://zenodo.org/communities/eosc-eden)

#EOSCEDEN



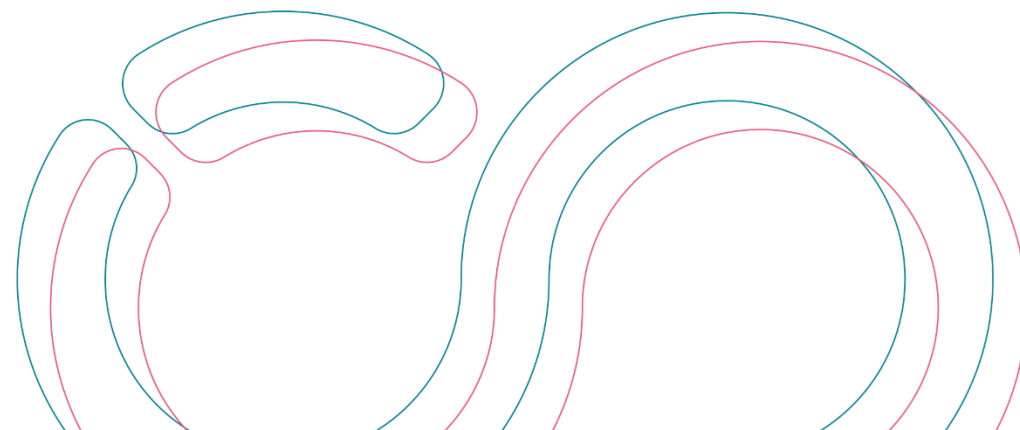
Funded by  
the European Union



The banner features the EOSC EDEN logo on the left. The main text reads 'Exploring Core Preservation Processes (CPPs)' and 'Join the 1st EOSC EDEN Seminar!'. A date box indicates '01 Dec 2025' and '11:00-12:00 CET'. A 'Register now' button is present. On the right, a section titled 'EDEN' contains the text 'What every Trustworthy Digital Archive should be doing: 30 Core Preservation Process Descriptions' and 'Exploring Core Preservation Processes (CPPs) - Join the 1st EOSC EDEN Seminar...'. A date '1 December 2025' and a 'Read more' link with an arrow are at the bottom right. An illustration of a laptop with a video call interface is in the center.

Registration: <https://eden-fidelis.eu/form/eosc-eden-exploring-core-preserv>

# Thank You



# FIDELIS, towards a network of TDRs





# FIDELIS the Project

Project full title	FIDELIS: Establishing A European Network of Trustworthy Digital Repositories
Acronym	FIDELIS
Call identifier	HORIZON-INFRA-2024-EOSC-01-03
Type of action	Coordination and Support Action (CSA)
Coordinator	CSC - IT Centre for Science

In June 2022, a group of renowned European repository owners and infrastructure providers published a working paper titled *Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDRs)*, articulating a shared community vision to launch “a European network of FAIR-enabling Trustworthy Digital Repositories which will support the development and growth of TDRs”<sup>1</sup>.

The paper, which received broad input from the community and was endorsed by the EOSC Task Force on Long-Term Data Preservation, argues that such a network “can unite the repository community beyond finite project efforts, and offer a coordination and networking mechanism for addressing common challenges, like the need for long-term sustainability resources.”

**The FIDELIS project aims to further develop and implement the vision set by the community in this paper and establish, by the end of a three-year project, a healthy, vibrant and self-sustaining network of TDRs**



# Project Consortium

Coordinator



Swiss Institute of Bioinformatics

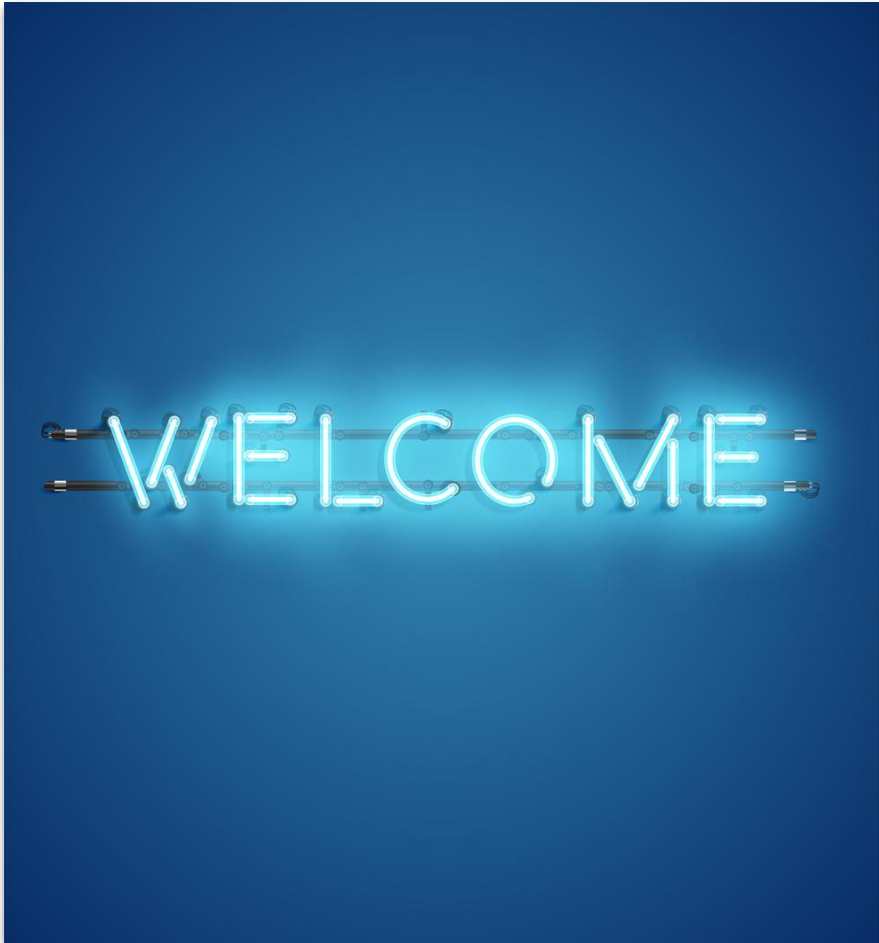


UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



# FIDELIS: help us shape a future-proof network!





- Which repository do you represent?
- Would your organisation consider to join the network?
- What are your expectations of the network?
  - useful topics for the network to discuss

eosc | Blue-Cloud2026

# Any questions?

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union



eOSC | Blue-Cloud2026

# Overview of Blue-Cloud 2026 Virtual Labs & Workbenches

Presented by Cyrielle Delvenne (VLIZ) &  
Marine Vernet (IFREMER)

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union **38**

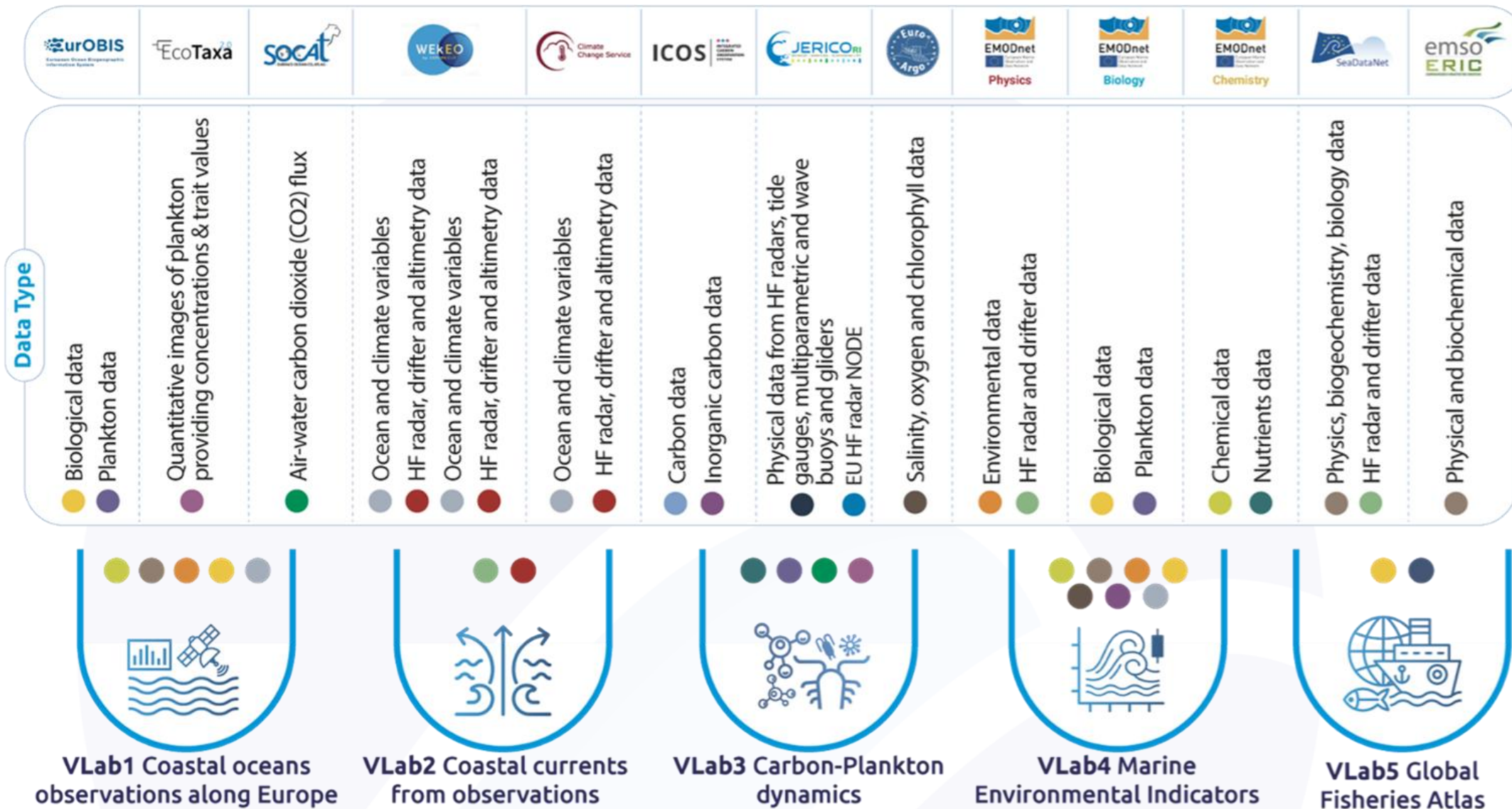


# Blue-Cloud 2026 Virtual Labs



<https://blue-cloud.org/virtual-labs>

## Blue Data Infrastructures



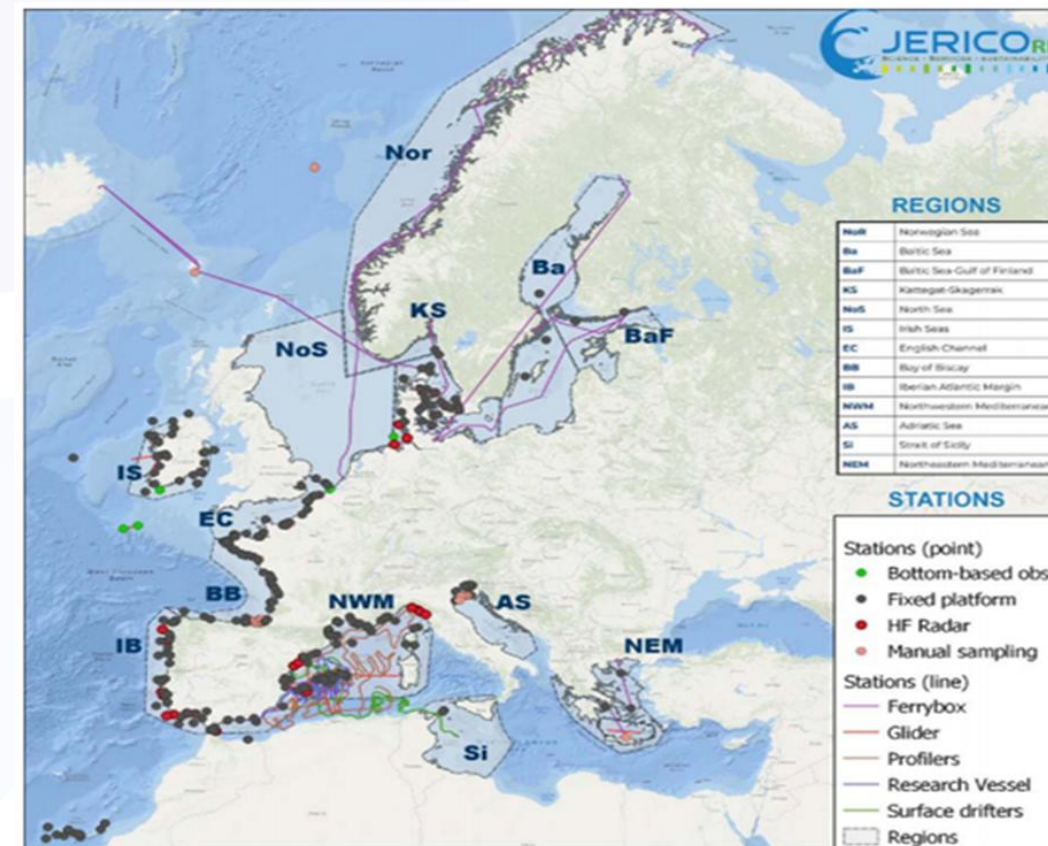




Integration of European coastal observations in 3 thematic services: **Transboundary Processes and Connectivity**, Extreme Events & Ocean Glider



- HF Radar Currents
- Current Profile in MP Buoys
- T in Wave Buoys
- T,S in MP Buoys
- T,S Glider profiles
- SSH at coastal tide gauges
- SST fields
- NEMO 3D T,S, SSH, Current
- ERA5 Surface Meteo Params
- Physics, BGC, Biology
- Physics, Chemistry, Biology, Bathymetry
- Bathymetry





Integration of direct and indirect currents data from different sources, and application to run an oil spill model



In Situ - Global Ocean-Delayed Mode  
Observations of surface (drifters, HFR) and satellite altimetry

Jupyter notebooks  
DIVAnd analysis

Gridded surface currents

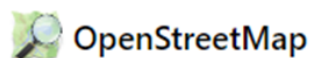
Initialise the oil spill forecasting model (MEDSLIK-II)



EMODnet



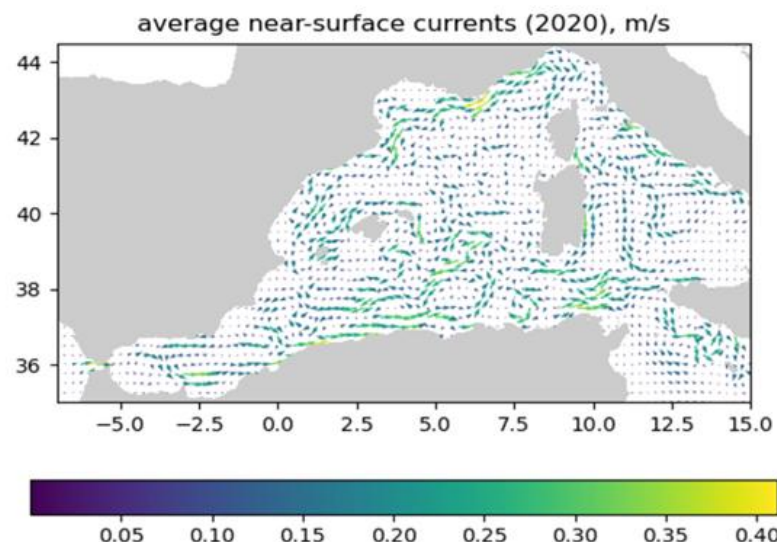
Bathymetry



Coastline

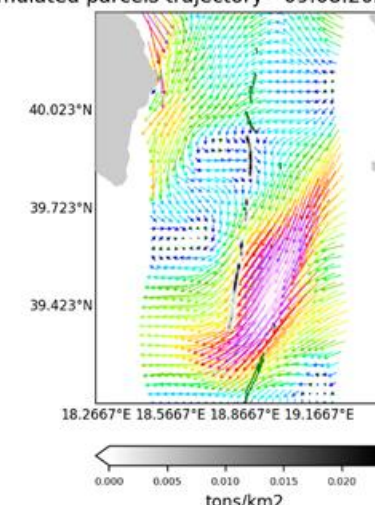


Wind speed



SANIFS + ECMWF (1% wind drift)

Simulated parcels trajectory - 09.08.2020 17:00 UTC



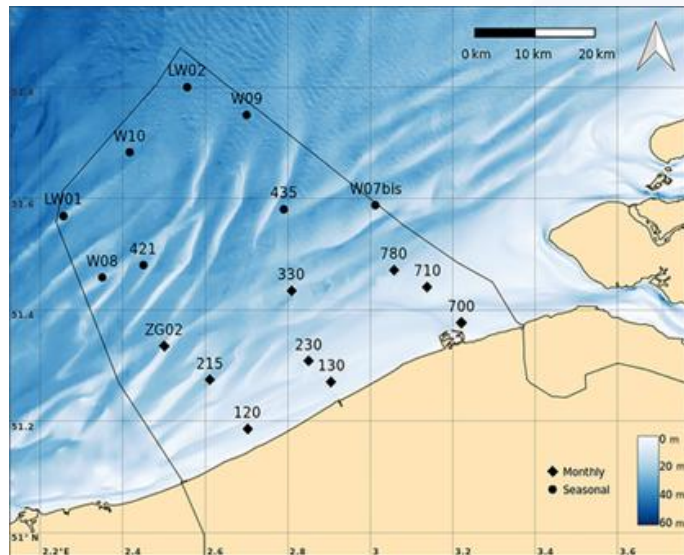


Seasonal and monthly records from 2011-2022

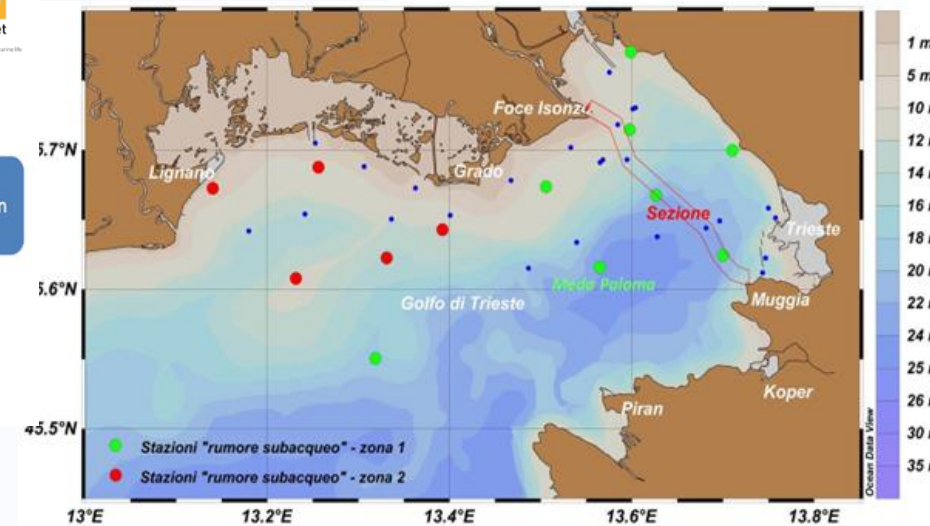
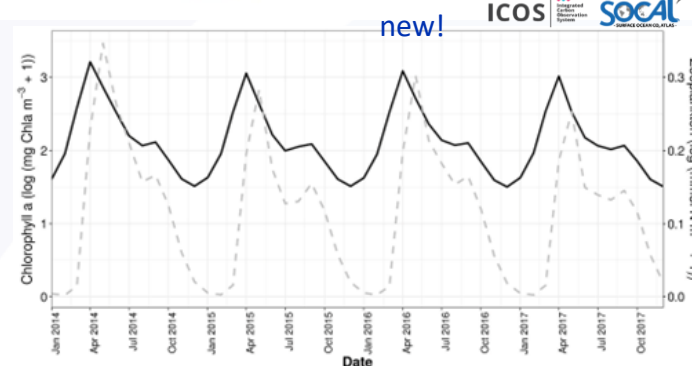
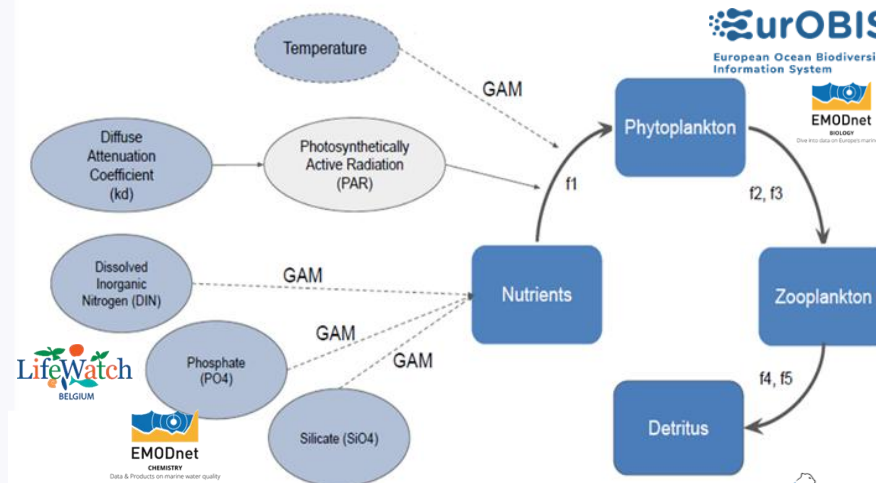
Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) Model, i.e. a mechanistic model, to identify the contribution of the drivers in phytoplankton dynamics and carbon dynamics.



Train and validate the NPZD model for the Adriatic



Mortelmans et al. (2019)



ogs.it







## Marine Environmental Indicators

Web app for cloud computation of new added-value data to monitor the environmental status of marine areas.



## SERVICES

new!

- Marine Environmental indicator (MEI) generator
- Ocean patterns and ocean regimes indicators
- Storm severity index
- Easy access to carbon data
- Ocean heat content
- Enhance Storm Severity Index (SSI v2)
- Trophic Index (TRIX)
- Marine heat wave



- Temperature and salinity Historical Data
- Mediterranean Sea Physics Reanalysis
- Global Ocean Physics Reanalysis
- Wind speed
- Wind (ERA5) reanalysis
- Global in-situ observation
- Other environmental variables



Cloud Computing Platform (CCP)

Available outputs | Generate output | My requests | Help

Method: Ocean Climate

Output type: annual climatology map

Data source: MEDSEA\_MULTYEAR\_PHY\_006\_004\_BC

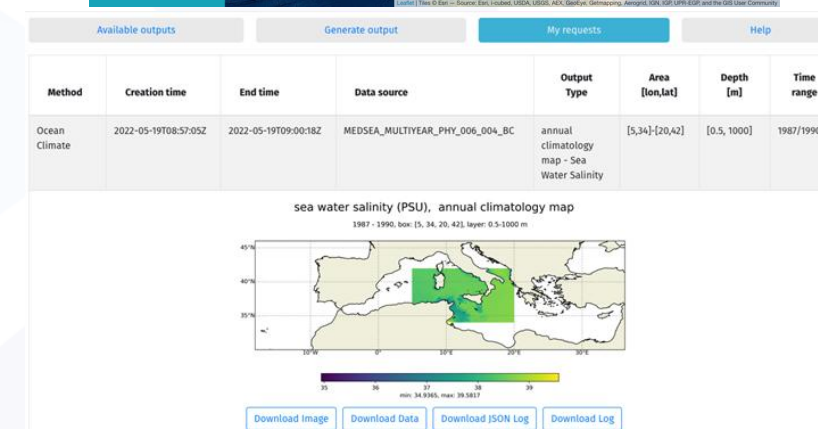
Output field: Sea Water Salinity

Time range: 1987 - 1990

Area: Left: 34, Right: 42, Extension from 30.0 to 48.00

Depth: From: 0.5, To: 1000, Extension from 0 to 1000

Execute process







## Discovery & Access of Global Record of Stocks and Fisheries, and Fisheries Atlas datasets



GRSF Catalog  
(Knowledge level)



Fetch  
Transform

RDF  
Format

GRSF staging  
KB

Data processing: records  
cleaning, standardization,  
dissection & merging

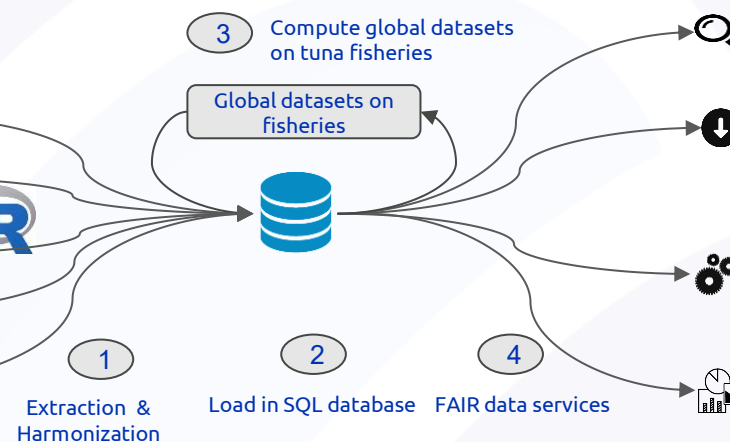
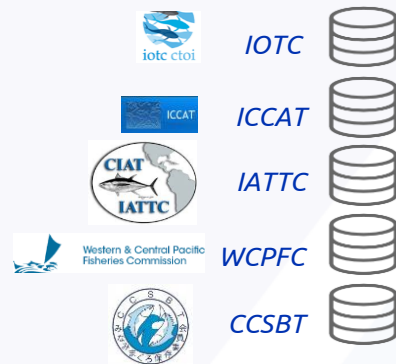
Data curation &  
validation



new!  
Workflows  
merging

Fisheries Atlas  
(data level)

(eg Tuna) RFMOs



**Discover** available data  
*What datasets exist?  
How they were built?*

**Access** the data and  
code  
*Different protocols and  
formats?*

**Process** the data  
*How to customize a fisheries  
atlas?*

**Visualize** the data  
*How to easily create maps, plots?*



# Blue-Cloud 2026 EOVS Workbenches

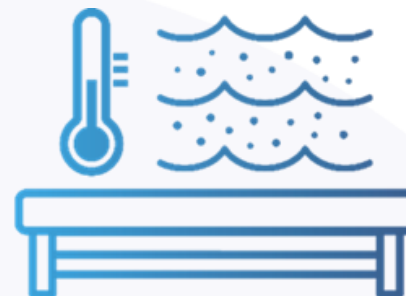
<https://blue-cloud.org/workbenches-essential-ocean-variables-eovs>



Cloud-based, analytical pipelines  
Leveraging EOSC Node Generic  
Capabilities

To produce highly qualified  
datasets for chosen  
Essential Ocean Variables &  
Essential Biodiversity Variables

From various major ocean  
research infrastructure data  
collections



**physical** workbench  
for temperature,  
salinity



**chemical** workbench,  
linked to  
eutrophication:  
nutrients, chlorophyll,  
oxygen



**ecosystem**  
workbench for  
plankton biomass and  
diversity

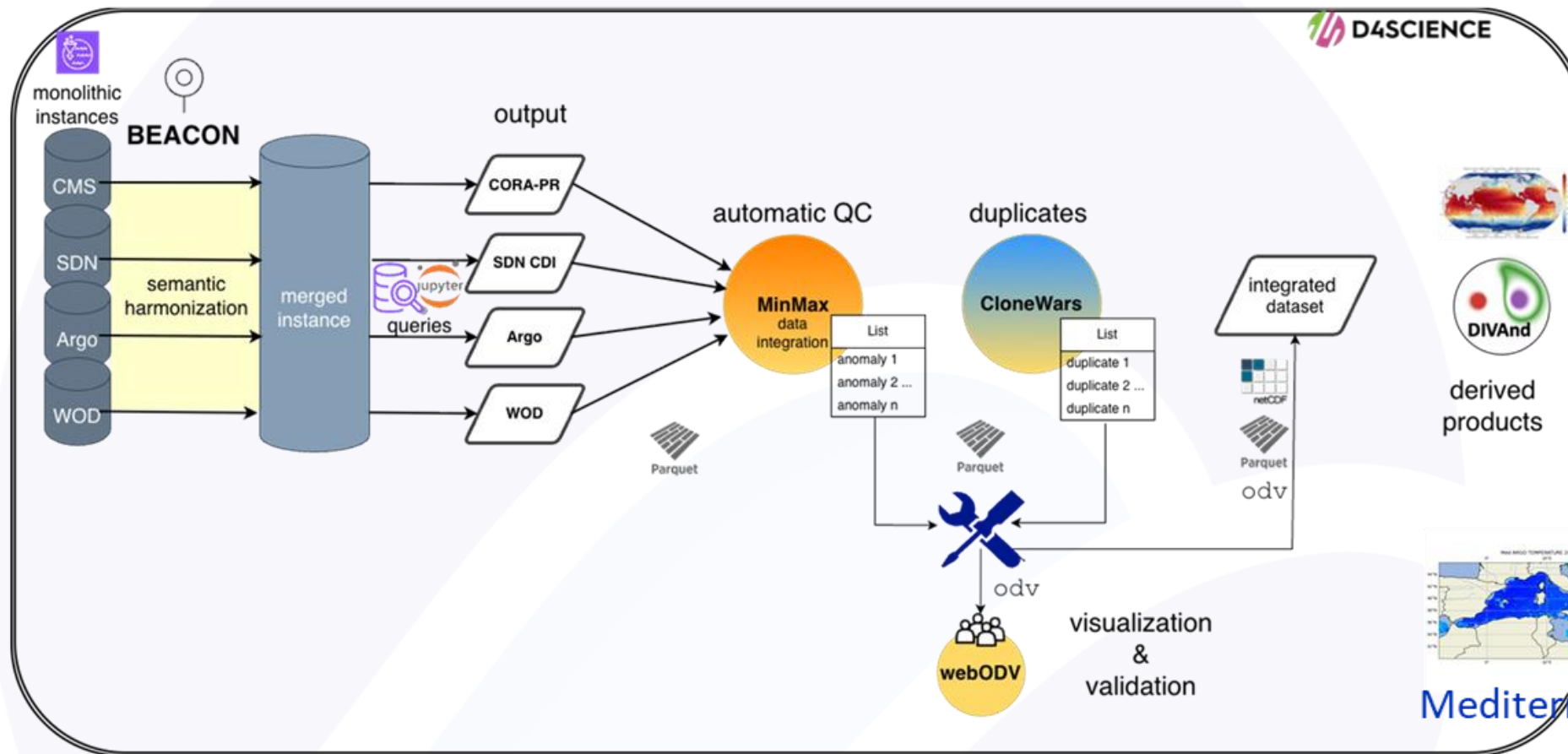




## Physical Workbench for Temperature & Salinity EOV

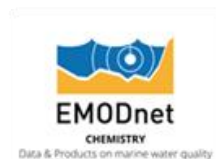


BDI

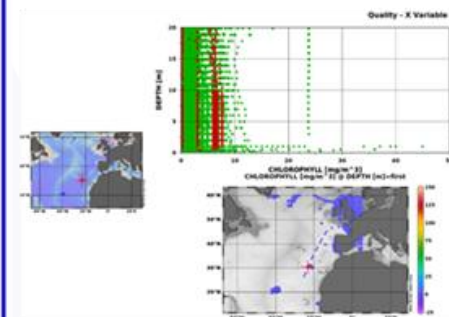
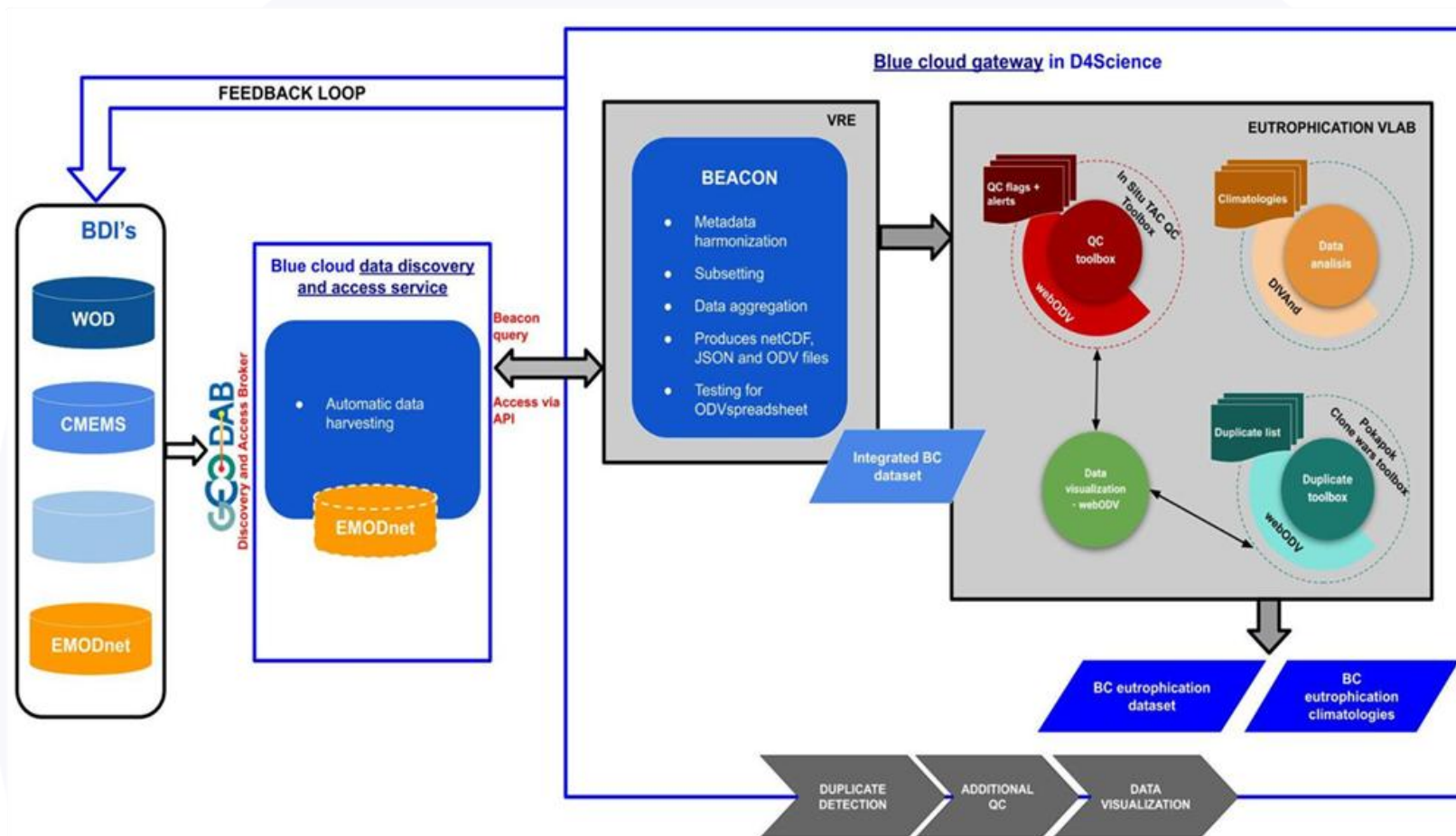




## Eutrophication Workbench for nutrients, oxygen and chlorophyll EOVS



BDI



North East Atlantic



## Ecosystem Workbench for plankton diversity & biomass EOV & EBV

ETH zürich



OBIS OCEAN BIODIVERSITY  
INFORMATION SYSTEM

GBIF Global Biodiversity  
Information Facility

AtlantECO Atlantic Ecosystems Assessment, Forecasting & Sustainability

Mgnify

EcoTaxa<sup>2.0</sup>

BDI

Plankton data

urOBIS  
European Ocean Biodiversity  
Information System

GBIF

AtlantECO

EcoTaxa<sup>2.0</sup>  
Mgnify

Global Environmental data

Temperature, O<sub>2</sub>, Chlorophyl,  
Alkalinity, EKE, HCO<sub>3</sub>, PAR, DIC, ...

Copernicus  
Marine Service



Input

Ecosystem Workbench

GEPHALOPOD

Species distribution modelling pipeline

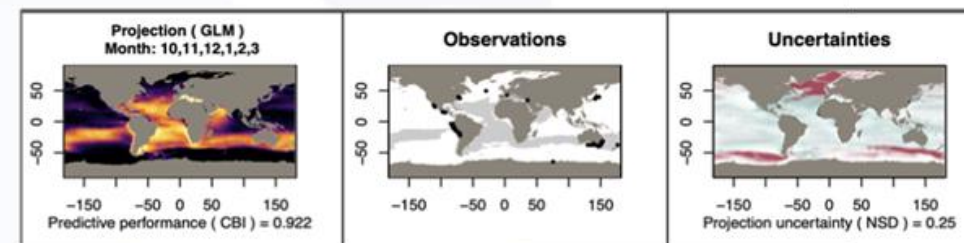
- Comprehensive suite of methods:  
(GLM, GAM, SVM, MLP, RF, BRT, MBTR)
- Quality checks, uncertainty estimation

Input

Output



Global plankton EOVs and EBVs  
(NetCDF files & Summary PDF)





eosc | Blue-Cloud2026

# Any questions?

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union

eosc | Blue-Cloud2026

# Eutrophication: chlorophyll, nutrients, oxygen

Nydia Catalina Reyes Suarez (OGS)

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union 53



## Essential Ocean Variables Workbench: Eutrophication

Partners: OGS (lead), Ifremer, HCMR, SMHI & Pokapok.

*Blue cloud 2026 General Assembly  
5-7 Nov 2025*



Funded by  
the European Union

eosc | Blue-Cloud2026

# Marine Environmental Indicators

Francesco Palermo (CMCC)

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union **55**



eosc | Blue-Cloud2026

# Physics: temperature & salinity

Simona Simoncelli (INGV)

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union 56



# EOV Workbench for Physics

*Simona Simoncelli*



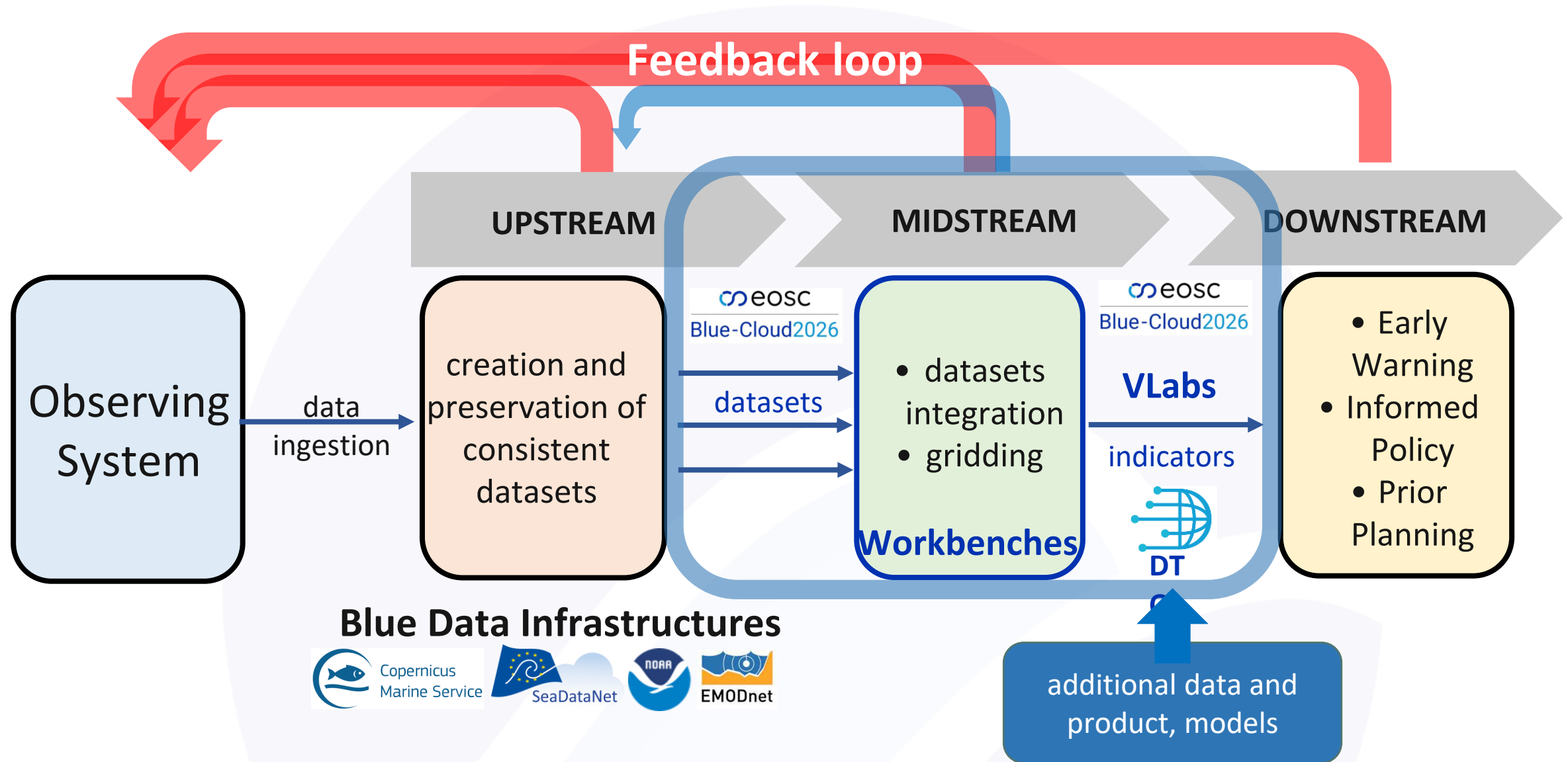
[simona.simoncelli@ingv.it](mailto:simona.simoncelli@ingv.it)  
[ORCID 0000-0003-1283-2798](https://orcid.org/0000-0003-1283-2798)

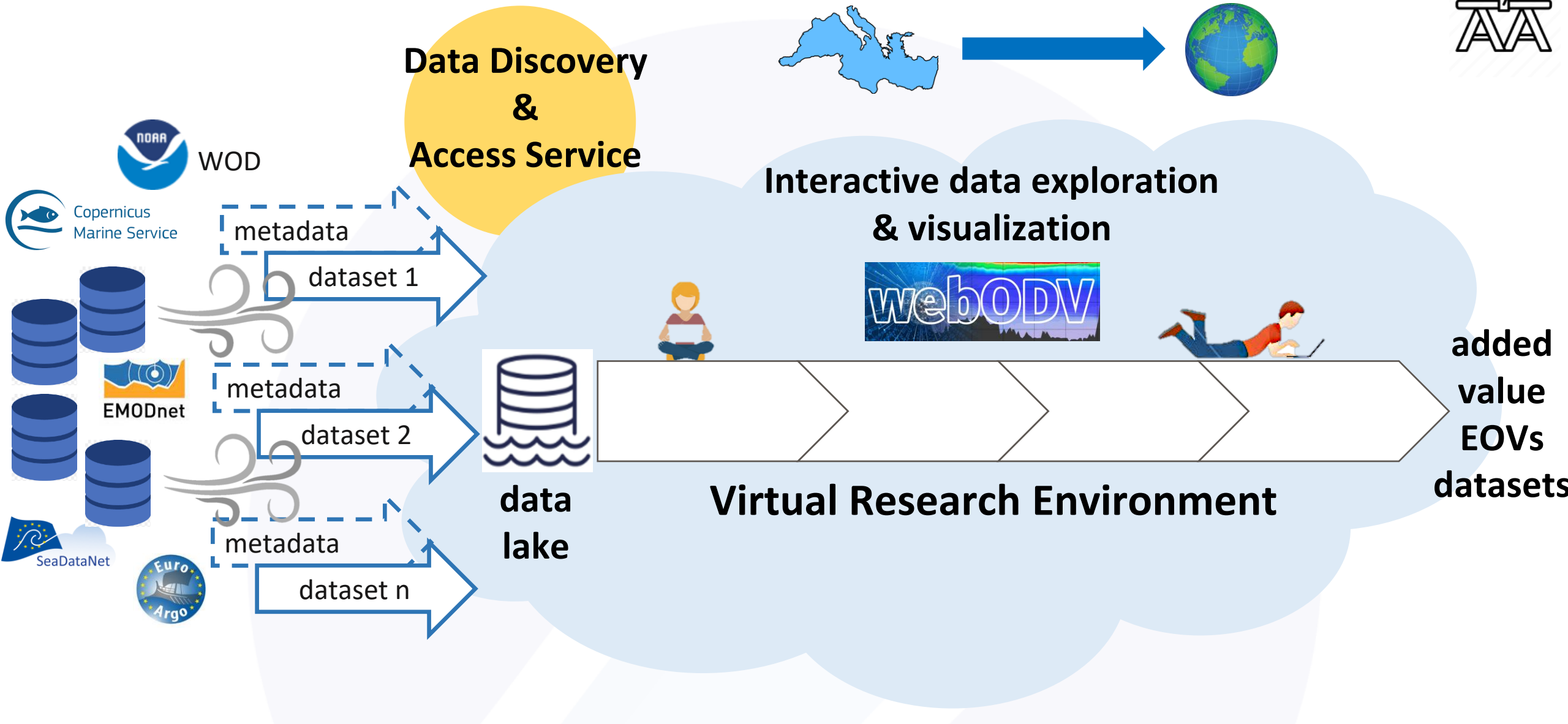
and the **PWB team**



Funded by  
the European Union

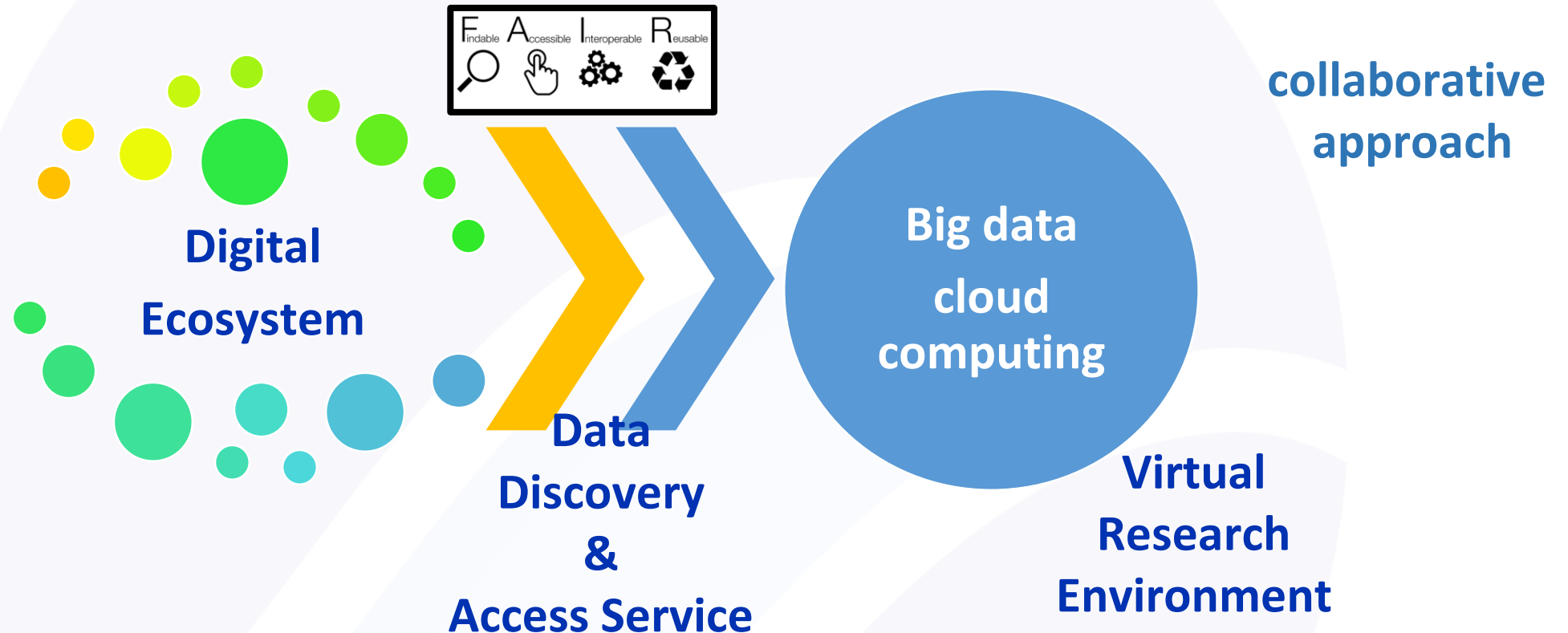
# Speeding up the data value chain



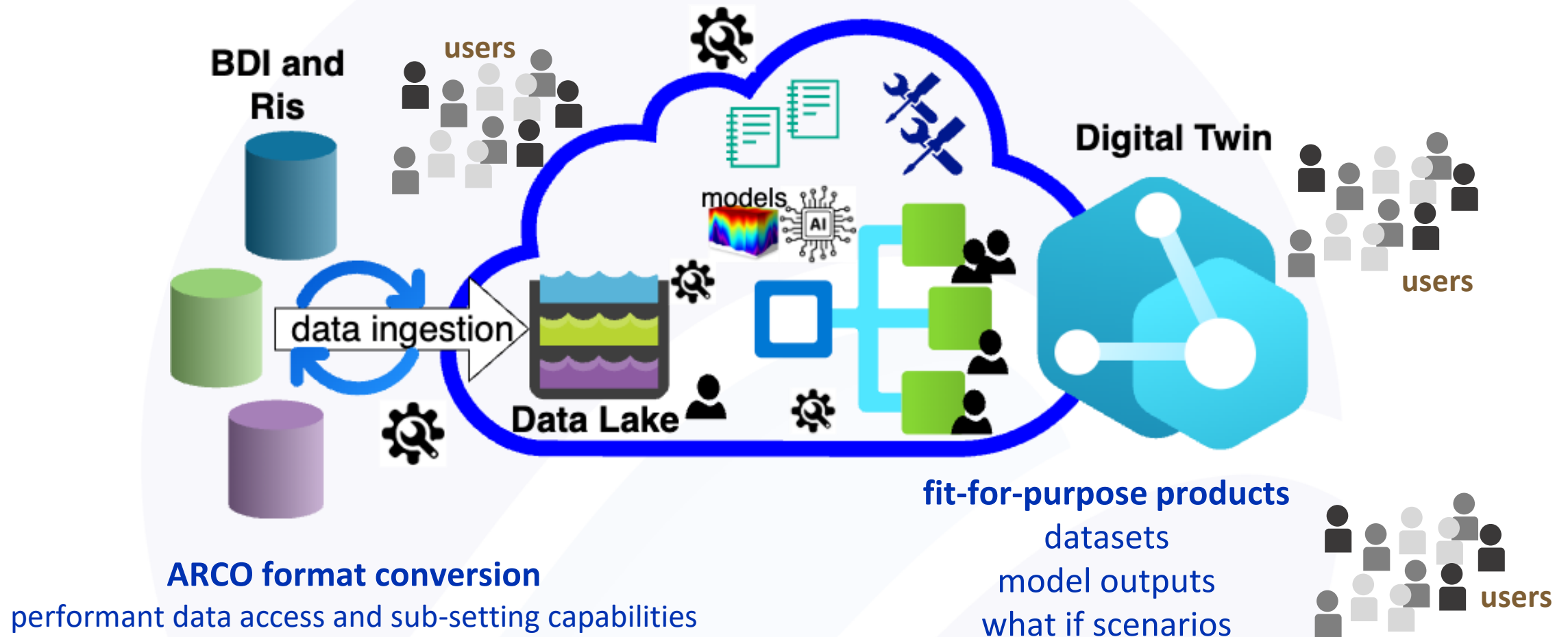


Deployment and orchestration of analytical tools on a Virtual Research Environment

- **automation** increases efficiency and replicability, optimizing the the results
- **FAIR principles** and **advanced services** for data access, web apps for software development and interactive computing

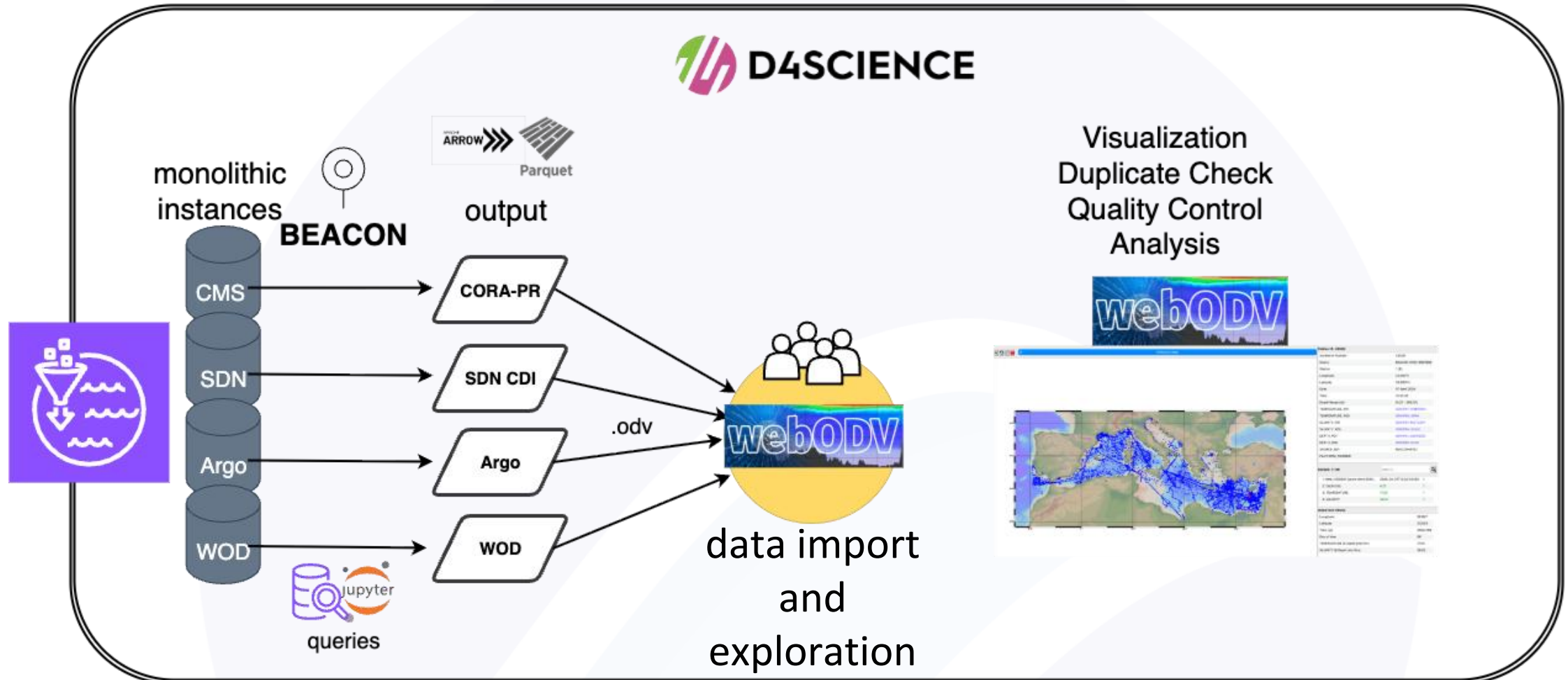


Innovative applications can be built thanks to cloud-based environments implementing operational workflows and exploiting data lake solutions



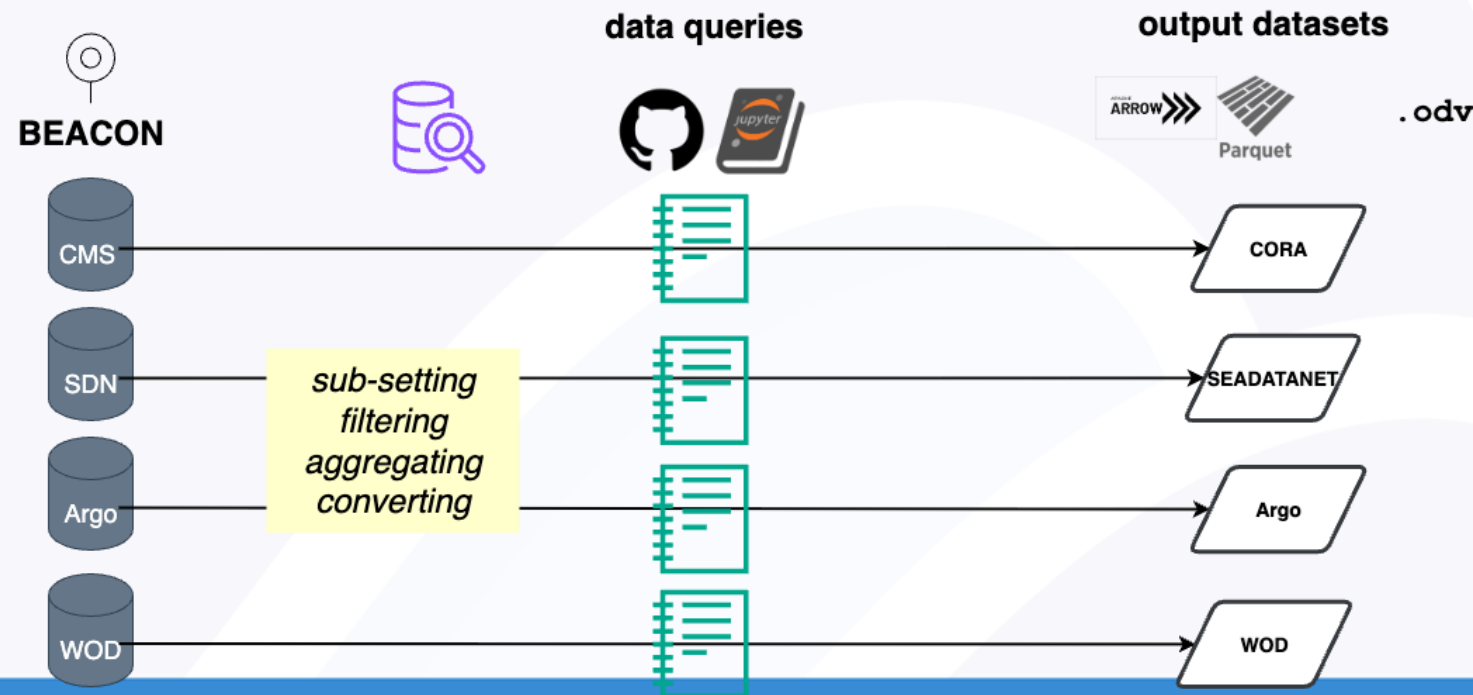


Beacon monolithic instances have been tested and Jupyter notebooks to query them have been set up



## Retrieval of ready to use datasets for scientific applications

- notebooks for the rapid/easy access to 4 data sources: Argo, World Ocean Database, SeaDataNet, Copernicus CORA
- blending the data experts knowledge with code optimization by the software developers
- the notebooks are managed through github and soon published

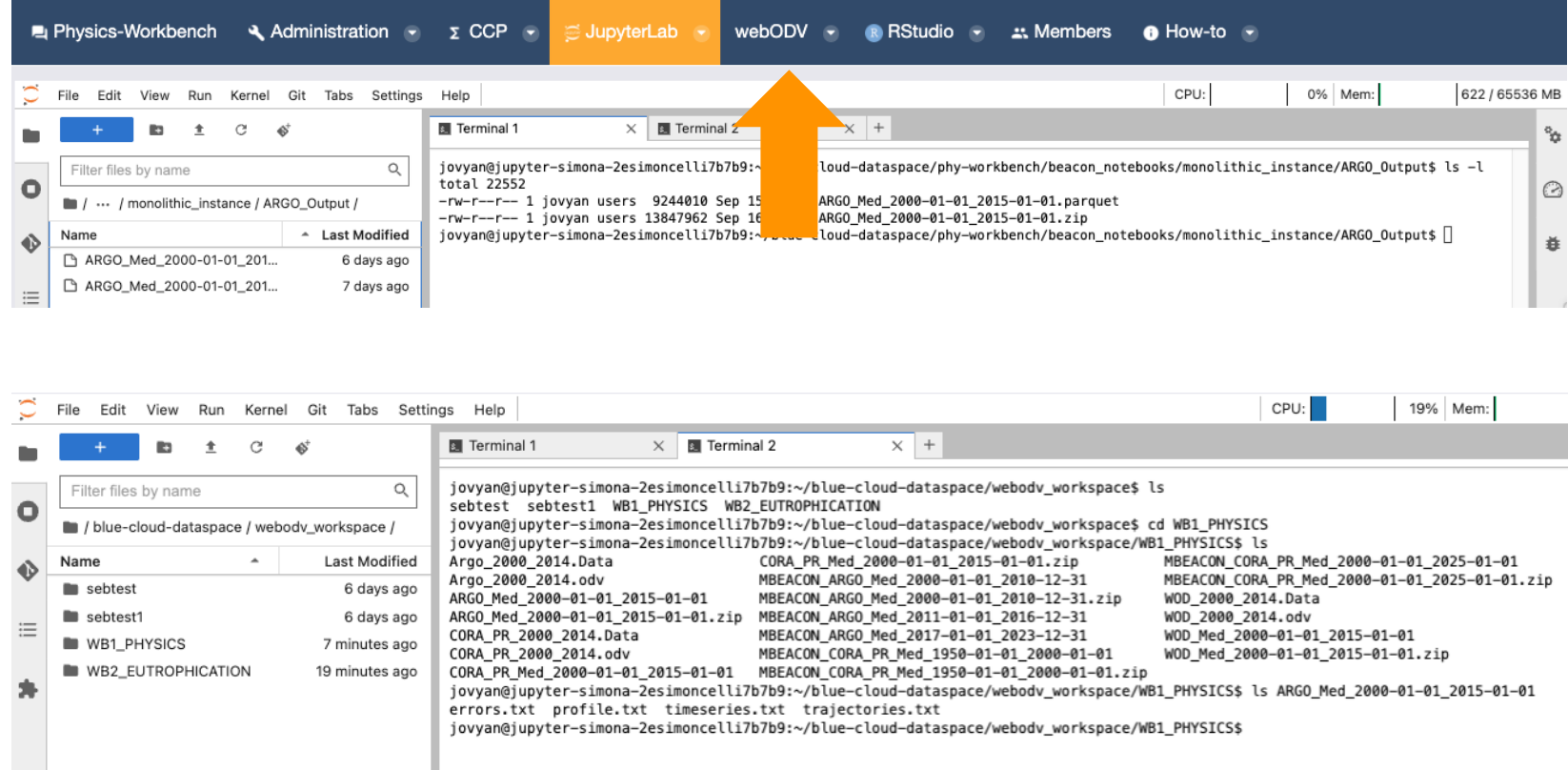


# EXAMPLE

Retrieve Argo T and S values in the Mediterranean Sea between 0-150m in the time period 2000-2014, filtering only GOOD data and obtain an output.zip (odv spreadsheet format) file

- diagnostics available in the notebook
- copy output.zip in webODV\_workspace
- unzip
- launch webODV and import the profile.txt file

.parquet, .nc, .arrow, .pandas, .geopandas, .zarr, .geoparquet, .odv, .csv formats are also available options to speed up you workflow!



Uncompress

```
$ls ARGO_Med_2000-01-01_2015-01-01*
```

**ARGO\_Med\_2000-01-01\_2015-01-01.zip**

```
ARGO_Med_2000-01-01_2015-01-01:
```

```
errors.txt
```

**profile.txt**

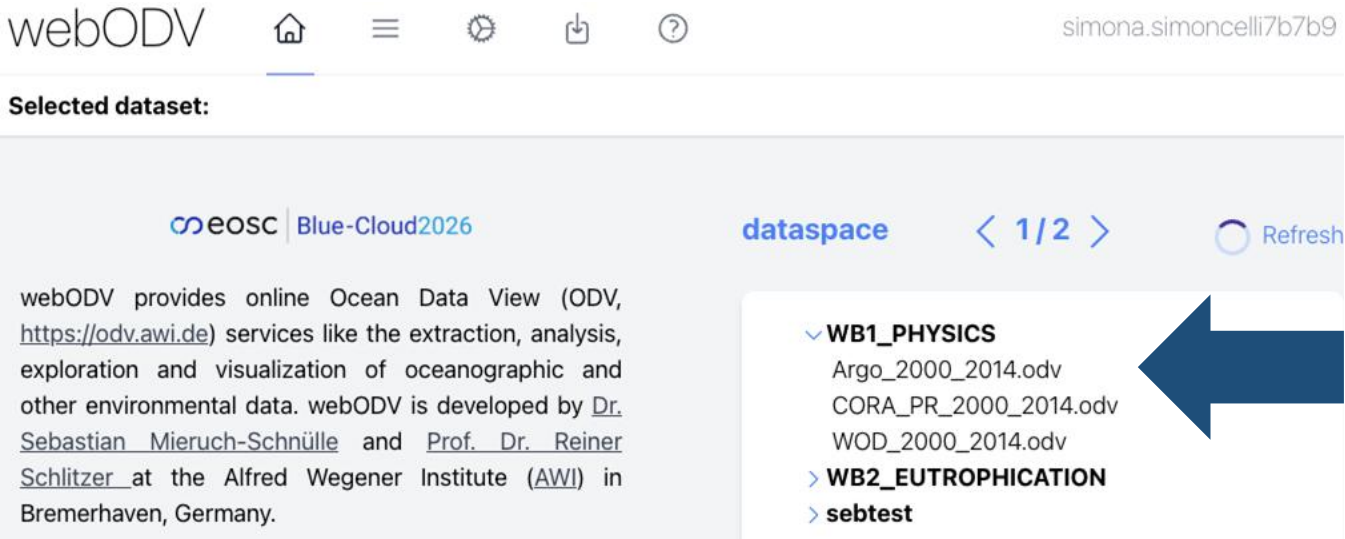
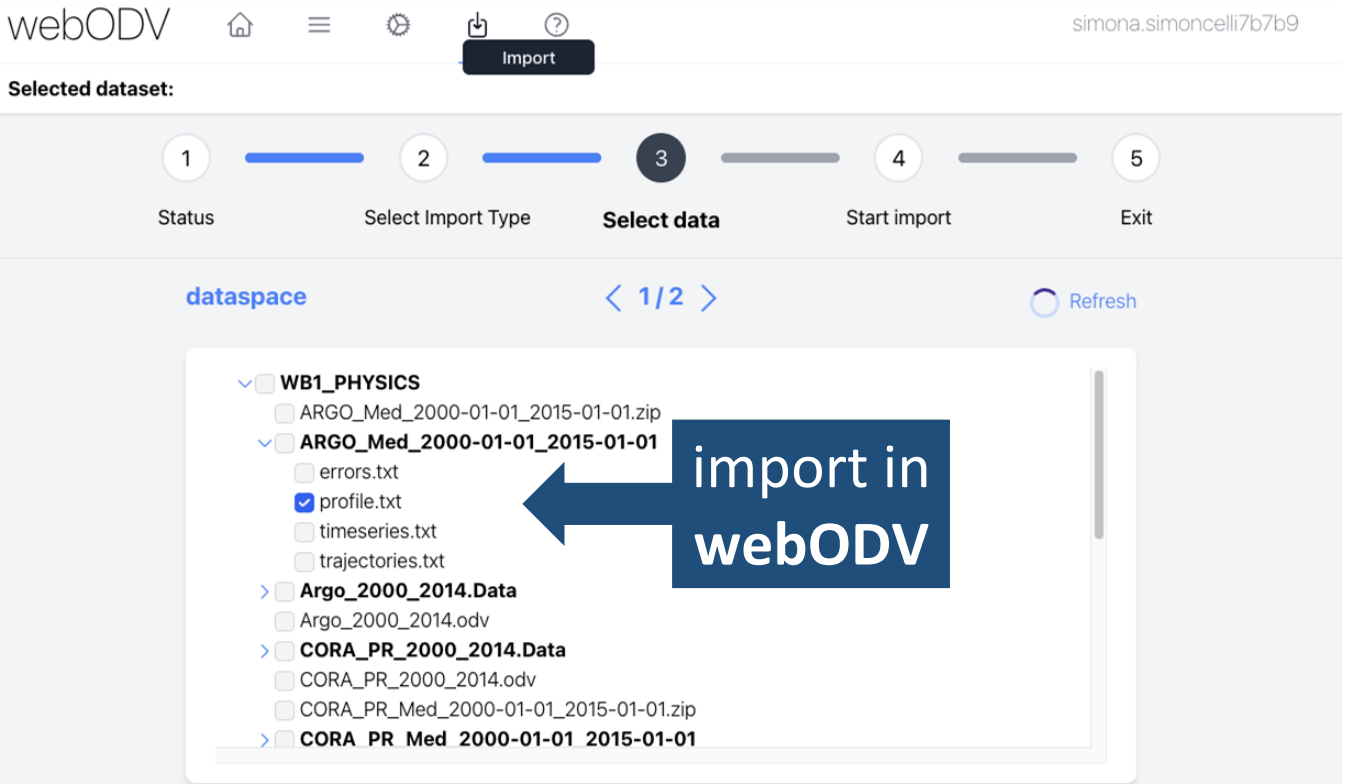
```
timeseries.txt
```

```
trajectories.txt
```

import in  
**webODV**

# EXAMPLE

Retrieve Argo Temperature and Salinity values in the Mediterranean Sea between 0-150m in the time period 2000-2014, filtering only GOOD data and obtain an output.zip (odv spreadsheet format) file





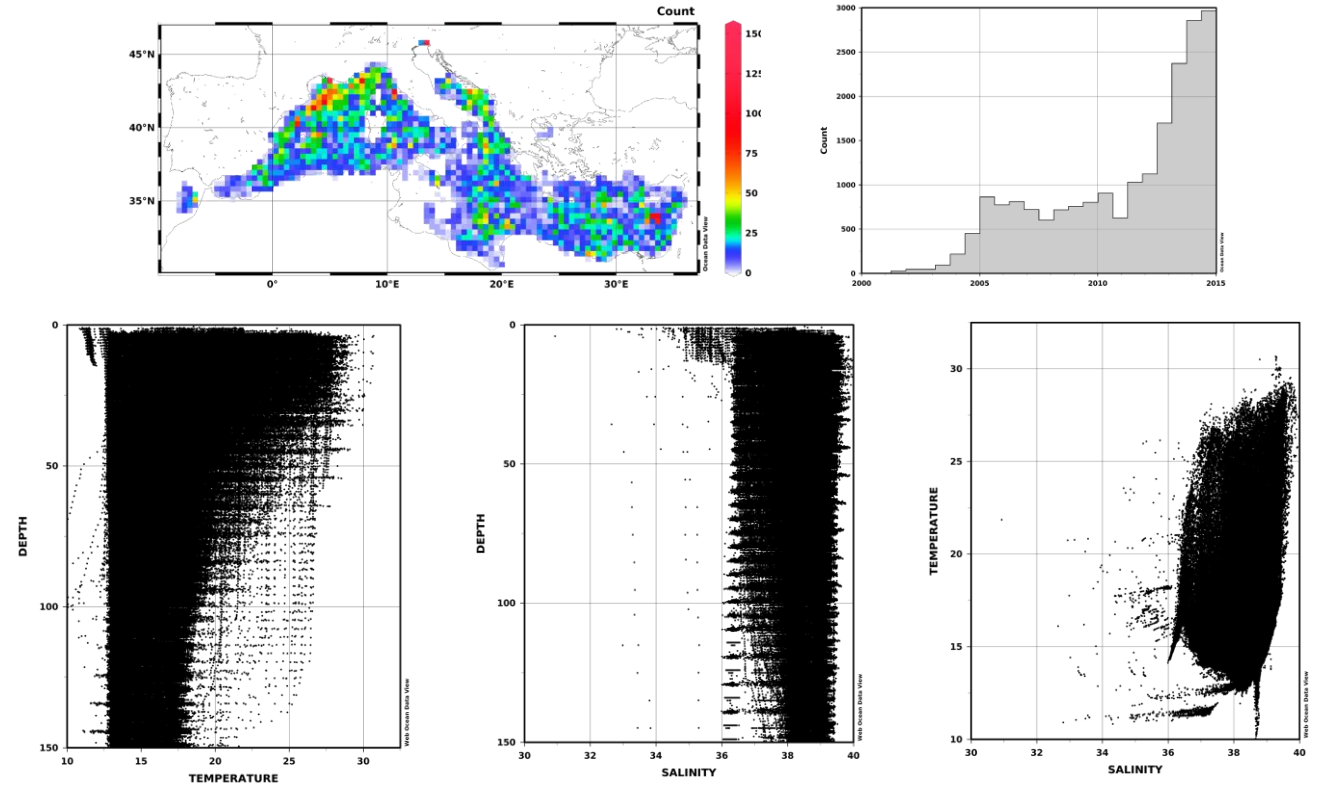
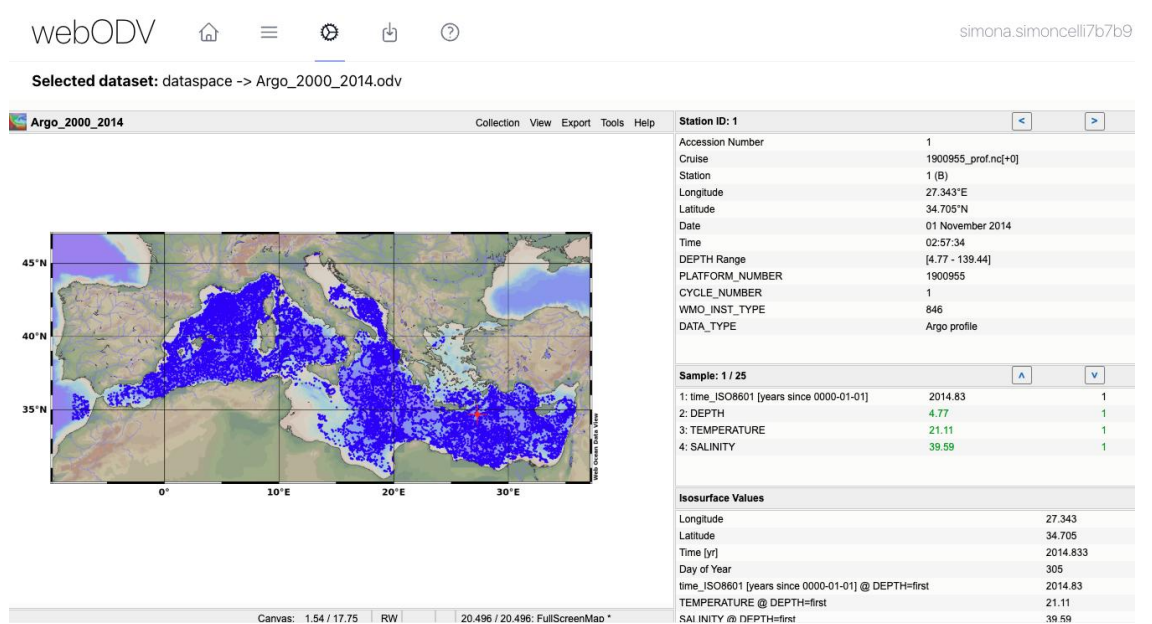
# EXAMPLE

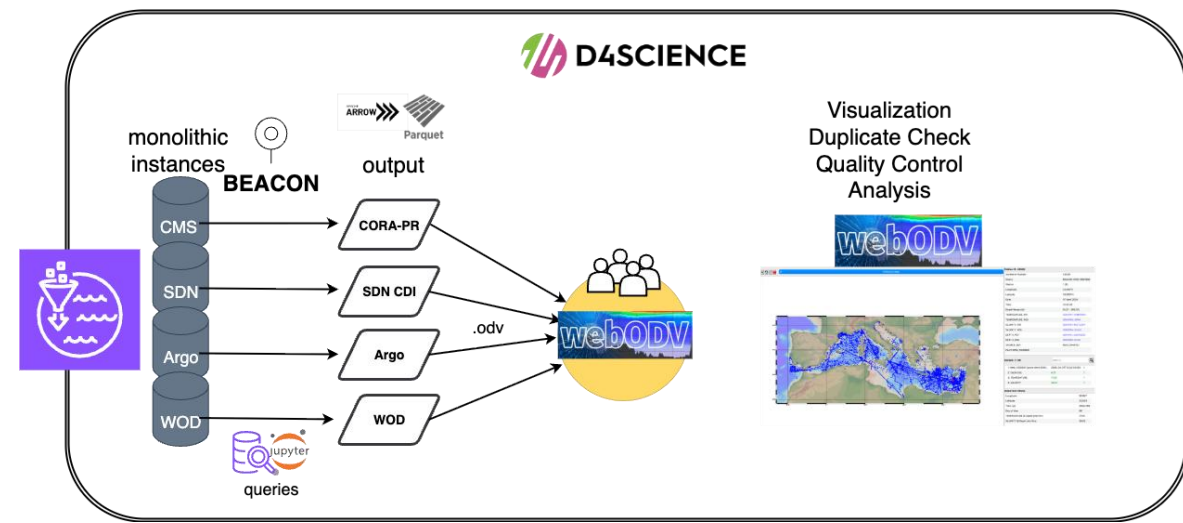
Retrieve Argo Temperature and Salinity values in the Mediterranean Sea between 0-150m in the time period 2000-2014, filtering only GOOD data and obtain an output.zip (odv spreadsheet format) file

diagnostics obtained in webODV

- distribution map
- time histograms
- scatter plots
- TS diagrams

webODV for data QC and duplicates detection

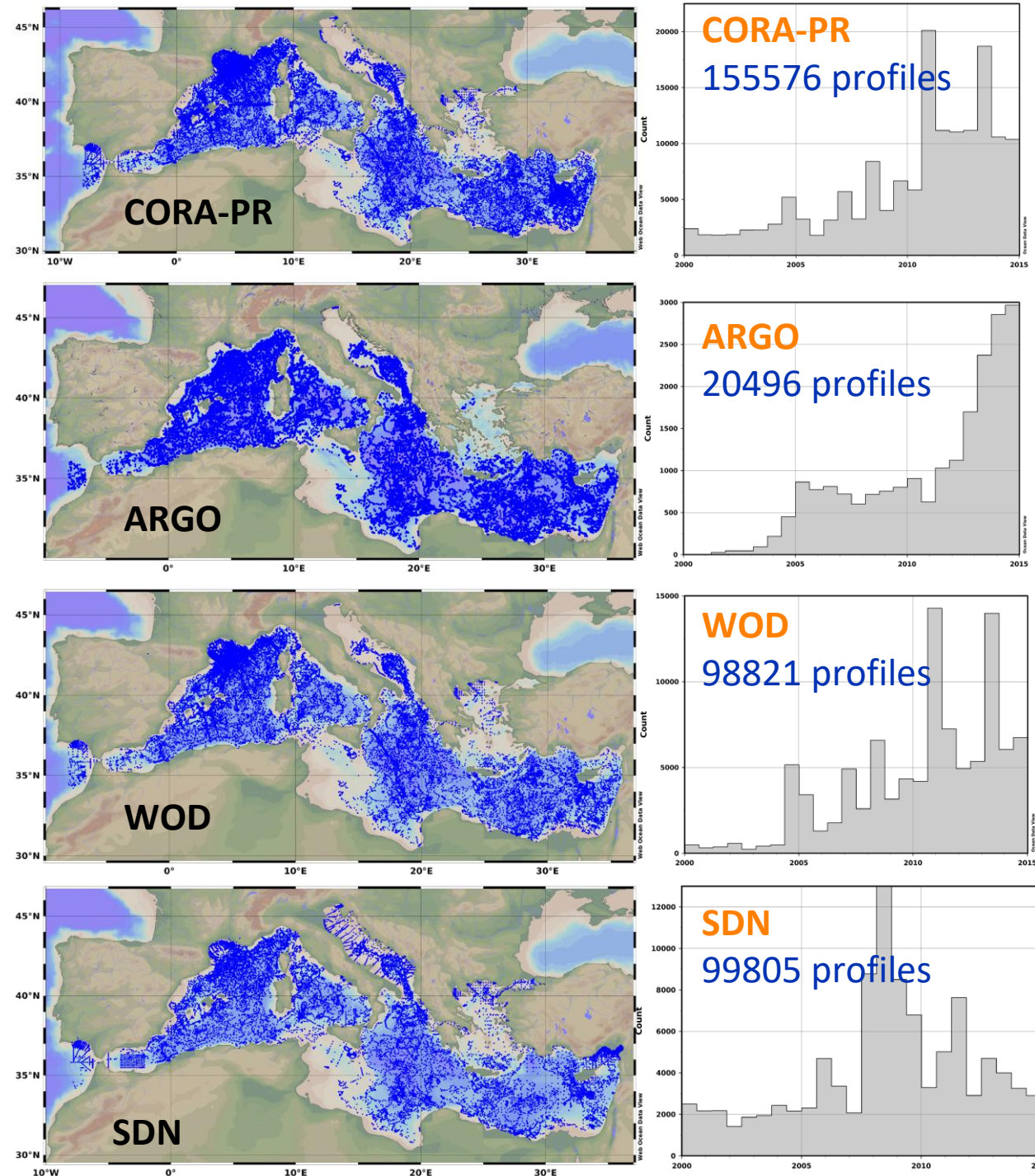




❓ You can query **4 Beacon Monoliths** to sub-set T and/or S in a specific domain, layer and time period, filtering GOOD data and obtain an output in many formats

Is there a dataset fitting your needs in terms of coverage, quality,...?

**Data integration is the next step!**



## Essential metadata retrieval:

- BDI ID
- platform
- instrument
- data provider

## Sub-setting capabilities:

- time period
- box domain and polygon selection
- depth layer
- variable

## Pressure to depth conversion

## Aggregation of parameters and variables (i.e. Argo ADJUSTED variables)

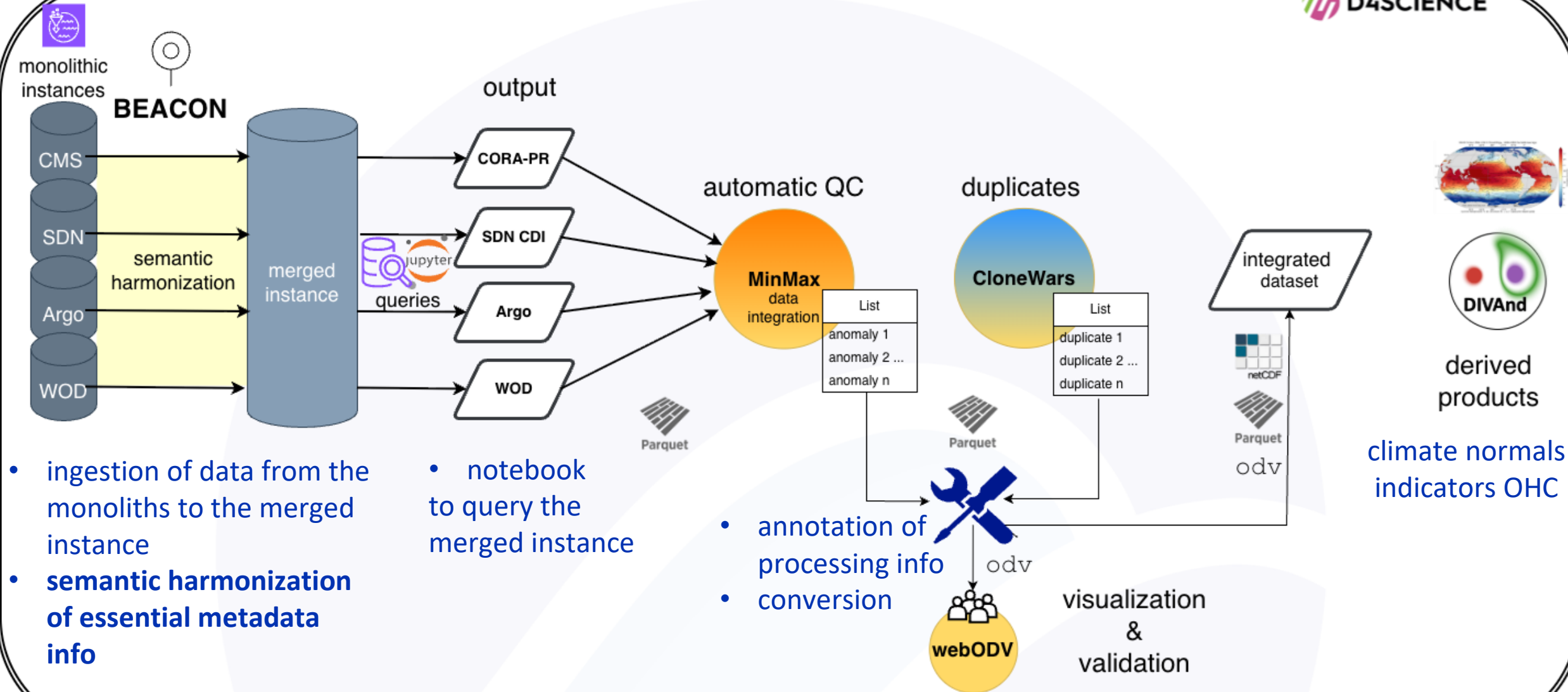
## Filtering capabilities:

- datasets (platforms, instrument, data provider)
- quality flags (i.e. GOOD data)

**Multi output:** pandas, geopandas, netcdf, zarr, parquet, geoparquet, odv, arrow, csv

- ✓ Consistency of the monoliths with the data source
- ✓ Correctness of the ingestion and harmonization process from monoliths to merged instance
- ✓ Set up of the merged beacon query and refinement of harmonization









Go to ▾

20

Simona Simoncelli ▾

Physics-Workbench

Administration ▾

CCP ▾

JupyterLab ▾

webODV ▾

RStudio ▾

Members

How-to ▾

File Edit View Run Kernel Git Tabs Settings Help

CPU:

0%

Mem:

213 / 65536 MB

Filter files by name

/ ... / beacon\_notebooks / merged\_instance /

Name	Last Modified
nopf_output_odv	2 minutes ago
nopf_output_parquet	2 minutes ago
full_output_parquet	13 days ago
full_output_odv	13 days ago
old_nb_202508	15 days ago
merged_instance_query.ipynb	12 hours ago
monolith_wod_4testing.ipynb	6 days ago
monolith_seadatanet_4testing.ipynb	12 days ago
monolith_argo_4testing.ipynb	12 days ago
monolith_cora_pr_4testing.ipynb	12 days ago

Terminal 1

merged\_instance\_query.ipynb

Markdown ▾

Python [conda env:base] \*

## Build a query

```
[5]: # Input parameters
regionname = "Med"
BDI = "BEACON_SEADATANET"
bigram = "PF"
dataset = "profiling float" # "mooring" # "profiling float"
mindate = "2000-01-01T00:00:00"
maxdate = "2010-12-31T23:59:59"
minlon = -8
maxlon = 36.5
minlat = 30
maxlat = 46
region = [
    (-9.50, 30.00),
    (-9.50, 39.00),
    (-0.80, 43.20),
    (16.00, 46.50),
    (36.50, 38.00),
    (36.50, 30.00),
    (-9.50, 30.00)
]

format = "pandas" # Options are: pandas, geopandas, netcdf, zarr, parquet, geoparquet, odv, arrow, c
output_file_name = f"{BDI}_{regionname}_2000_2010_nopf"
df = None
gdf = None
```

## Filters for optimized integration

[illegible]

# eosc | Blue-Cloud2026



[blue-cloud.org](https://blue-cloud.org)



[@bluecloudeu](https://twitter.com/bluecloudeu)



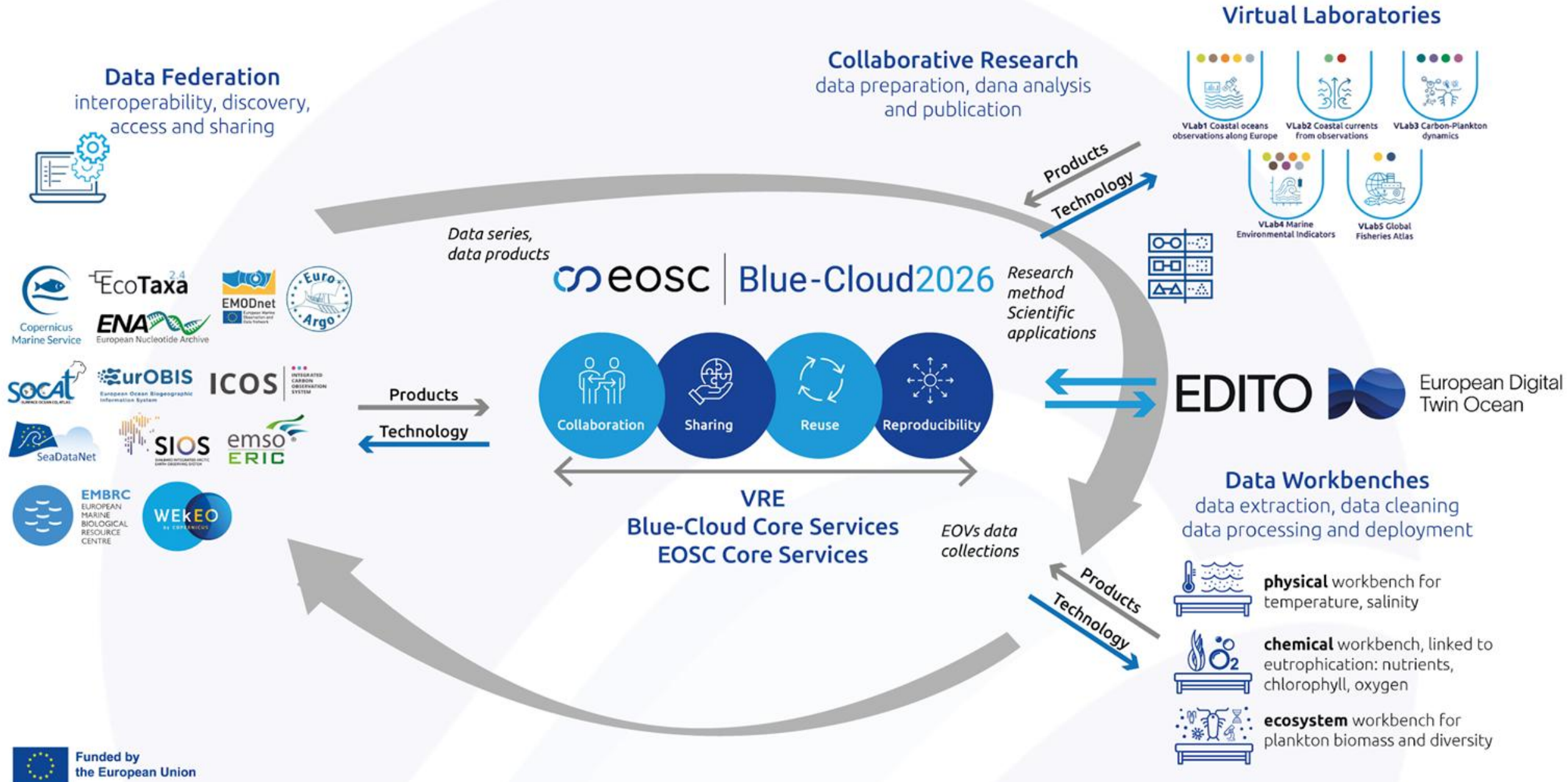
[blue-cloud org](https://www.linkedin.com/company/blue-cloud-org)



Funded by  
the European Union



# Blue-Cloud open science platform

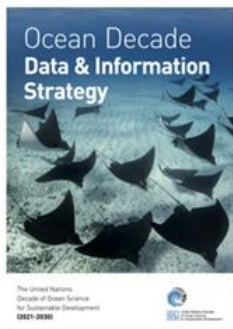
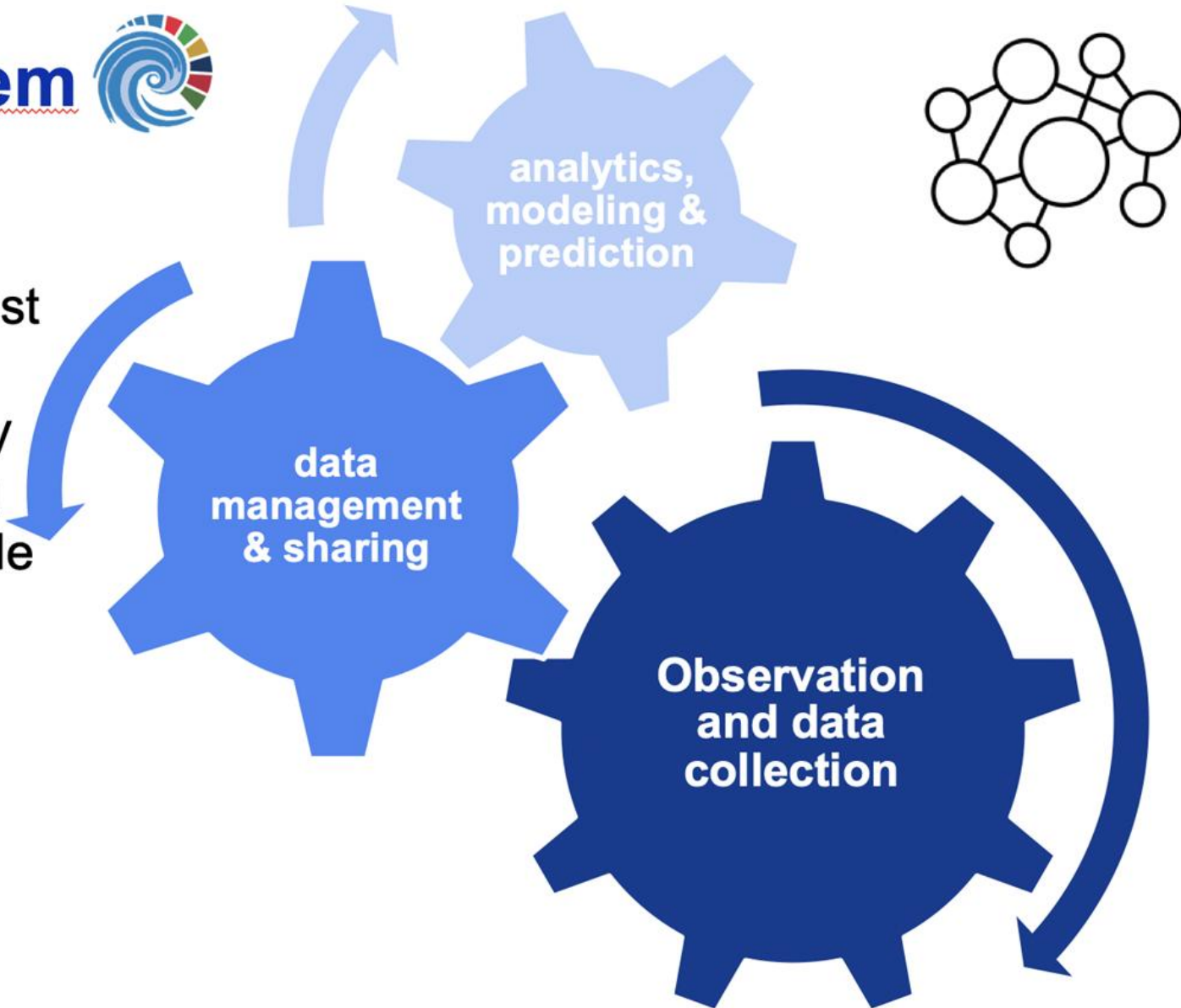




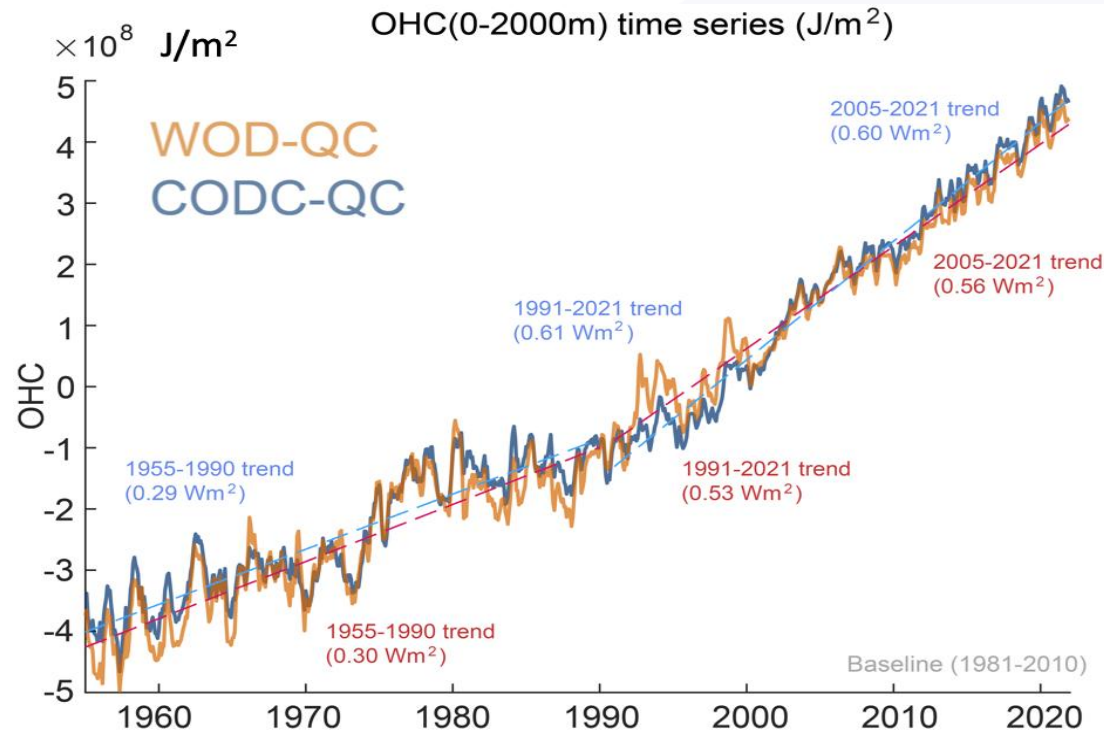
# Digital Ecosystem



Data systems (BDIs) co-create a distributed, robust and collaborative **digital ecosystem** to be actively used for decision making and to support sustainable ocean management

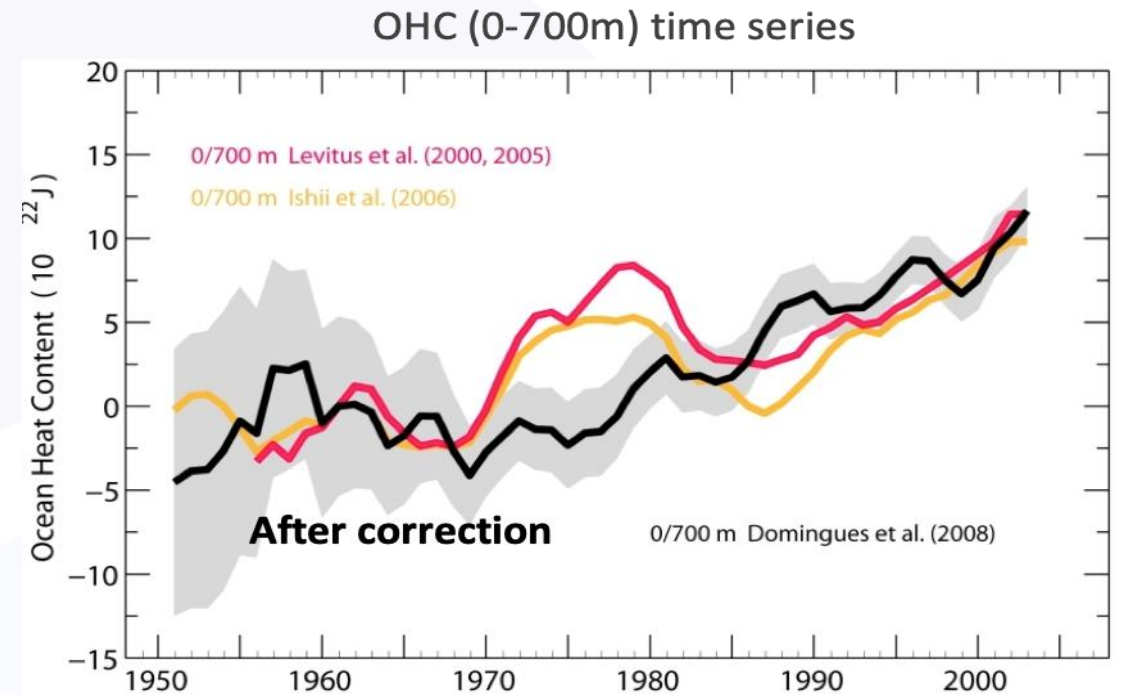


## Ocean Heat Content indicator



### Impact of QC on OHC 0-2000 m

~8% trend difference from 2005-2021



### Impact of instrumental bias on OHC 0-700 m

~50% trend difference from 1970-2000

eosc | Blue-Cloud2026

# Integration of coastal ocean observations along Europe

Joao Paulo do Nascimento Vitorino <sup>(IH)</sup>

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union

eosc | Blue-Cloud2026

# webODV

Sebastian Mieruch (AWI)

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union



eosc | Blue-Cloud2026

# The Global Fisheries Atlas

Julien Barde (IRD)

---

**5-6 November** 2025, Brussels [Belgium]



Funded by  
the European Union

# eosc | Blue-Cloud2026



[blue-cloud.org](https://blue-cloud.org)



[@bluecloudeu](https://twitter.com/bluecloudeu)



[blue-cloud org](https://www.linkedin.com/company/blue-cloud-org)



Funded by  
the European Union