



# Marine Data Management & Data Quality Control: Marine Datascape

Simona Simoncelli

**IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System**  
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-  
Mission 4 "Education and Research" - Component 2: "From research to business" - Investment  
3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures"



# Outline

- 🌐 Marine data landscape (datascape)
- 🌐 Data and information within the UN Ocean Decade
- 🌐 EU datascape and Blue Data Infrastructures
- 🌐 Ocean Observing
- 🌐 Data processing (analytics, workflows and virtual research environments)
- 🌐 Marine data value chain
- 🌐 Data life cycle and quality control
- 🌐 Data products
- 🌐 FAIR principles and FAIRness assessment

# Objectives

- 🌐 to present the marine datascape at international and European level and **demonstrate the value of data-driven science and data management to advance knowledge and to tackle emerging societal challenges**
- 🌐 to show the value of marine data when they are efficiently managed and integrated across data products and service organization
- 🌐 to present the need of a sustained Observing System and its network components
- 🌐 to demonstrate the importance of effective data management for their proper re-use
- 🌐 to introduce FAIR principles as key enablers of data-driven science and the importance of FAIRness assessment to fully unlock data potential



# UN Decade of Ocean Science for Sustainable Development 2021-2030

A screenshot of the oceandecade.org website homepage. The browser address bar shows "oceandecade.org". The navigation bar includes "OCEAN DECADE NETWORK", "GenOcean", and "ENGLISH". The main content area features a large image of dolphins leaping from the water. The text "THE OCEAN DECADE" is prominently displayed in white and cyan, with "THE" and "DECADE" in white and "OCEAN" in cyan. Below this, the tagline "The Science We Need for the Ocean We Want" is written in white. A vertical credit line on the right side reads "video © Universeum Ocean Science Lab".

oceandecade.org

OCEAN DECADE NETWORK GenOcean ENGLISH

2021 United Nations Decade of Ocean Science for Sustainable Development 2030

ABOUT US WHAT'S HAPPENING DECADE ACTIONS ORGANIZATION GET INVOLVED

# THE OCEAN DECADE

The Science We Need for the Ocean We Want

video © Universeum Ocean Science Lab



# Ocean Decade: 7 expected outcomes



## A clean ocean

where sources of pollution are identified and reduced or removed.



## A healthy and resilient ocean

where marine ecosystems are understood, protected, restored and managed.



## A productive ocean

supporting sustainable food supply and a sustainable ocean economy.



## A predicted ocean

where society understands and can respond to changing ocean conditions.



## A safe ocean

where life and livelihoods are protected from ocean-related hazards.



## An accessible ocean

with open and equitable access to data, information and technology and innovation.



## An inspiring and engaging ocean

where society understands and values the ocean in relation to human wellbeing and sustainable development.

→ **data and information are key enablers**

<https://www.oceandecade.org/>



## Challenge 1

### Understand and beat marine pollution

Understand and map land and sea-based sources of pollutants and contaminants and their potential impacts on human health and ocean ecosystems and develop solutions to remove or mitigate them.



## Challenge 2

### Protect and restore ecosystems and biodiversity

Understand the effects of multiple stressors on ocean ecosystems, and develop solutions to monitor, protect, manage and restore ecosystems and their biodiversity under changing environmental, social and climate conditions.



## Challenge 3

### Sustainably nourish the global population

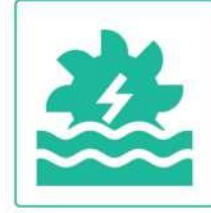
Generate knowledge, support innovation and develop solutions to optimize the role of the ocean in sustainably nourishing the world's population under changing environmental, social and climate conditions.



## Challenge 4

### Develop a sustainable, resilient and equitable ocean economy

Generate knowledge, support innovation and multi-sectoral partnerships and develop solutions for equitable, resilient and sustainable development of the ocean economy under changing environmental, social and climate conditions.



## Challenge 5

### Unlock ocean-based solutions to climate change

Enhance understanding of the ocean-climate nexus and generate knowledge and solutions to mitigate, adapt and build resilience to the effects of climate change across all geographies and at all scales, and to improve services including predictions for the ocean, climate and weather.



## Challenge 6

### Increase community resilience to ocean and coastal risks

Enhance multi-hazard early warning services for all geophysical, ecological, biological, weather, climate and anthropogenic related ocean and coastal hazards, and mainstream community preparedness and resilience.



## Challenge 7

### Sustainably expand the Global Ocean Observing System

Ensure a sustainable and sustained ocean observing system across all ocean basins that delivers accessible, timely and actionable data and information to all users.



## Challenge 8

### Create a digital representation of the ocean

Through multi-stakeholder collaboration, develop a comprehensive digital representation of the ocean, including a dynamic ocean map, which provides free and open access for exploring, discovering, and visualizing past, current, and future ocean conditions in a manner relevant to diverse stakeholders.



## Challenge 9

### Skills, knowledge, technology and participation for all

Ensure comprehensive capacity development and equitable access to data, information, knowledge and technology across all aspects of ocean science and for all stakeholders.

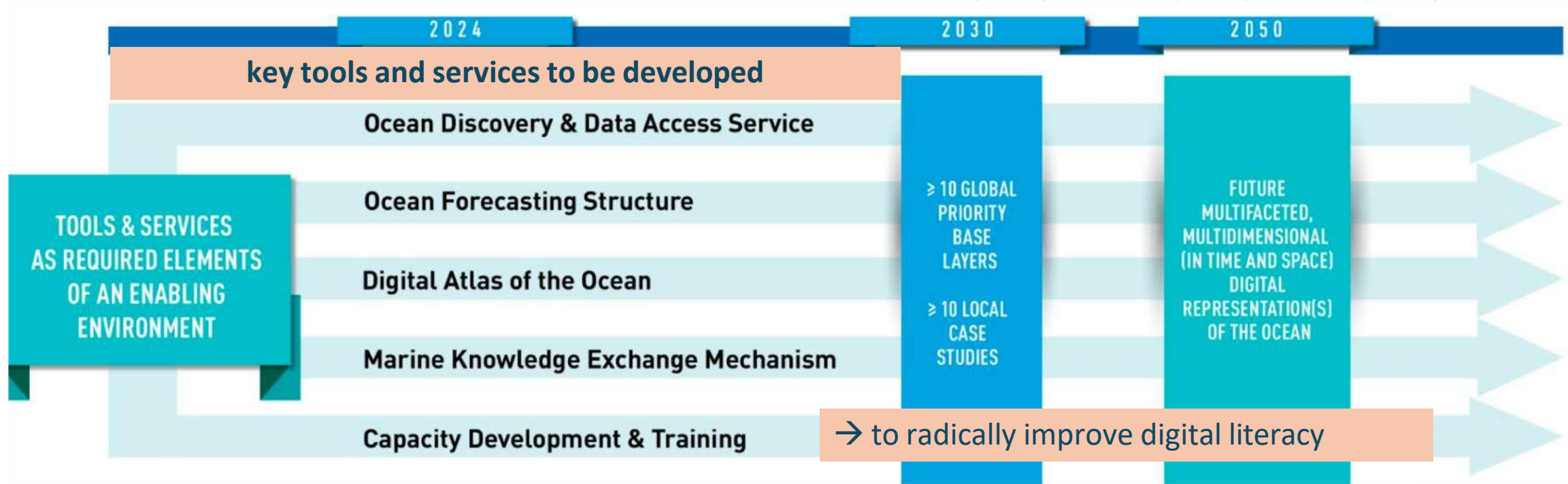


## Challenge 10

### Restore society's relationship with the ocean

Meaningful society-ocean connections are strengthened, driving increased motivation, capability and opportunity for people, across all sectors of society, to make decisions, act and behave in ways that ensure a healthy ocean.

# Digital Representation of the Ocean (challenge 8)



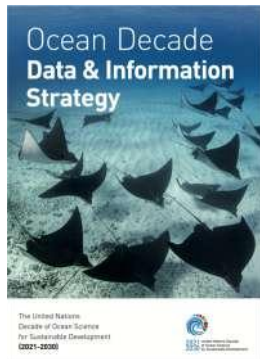
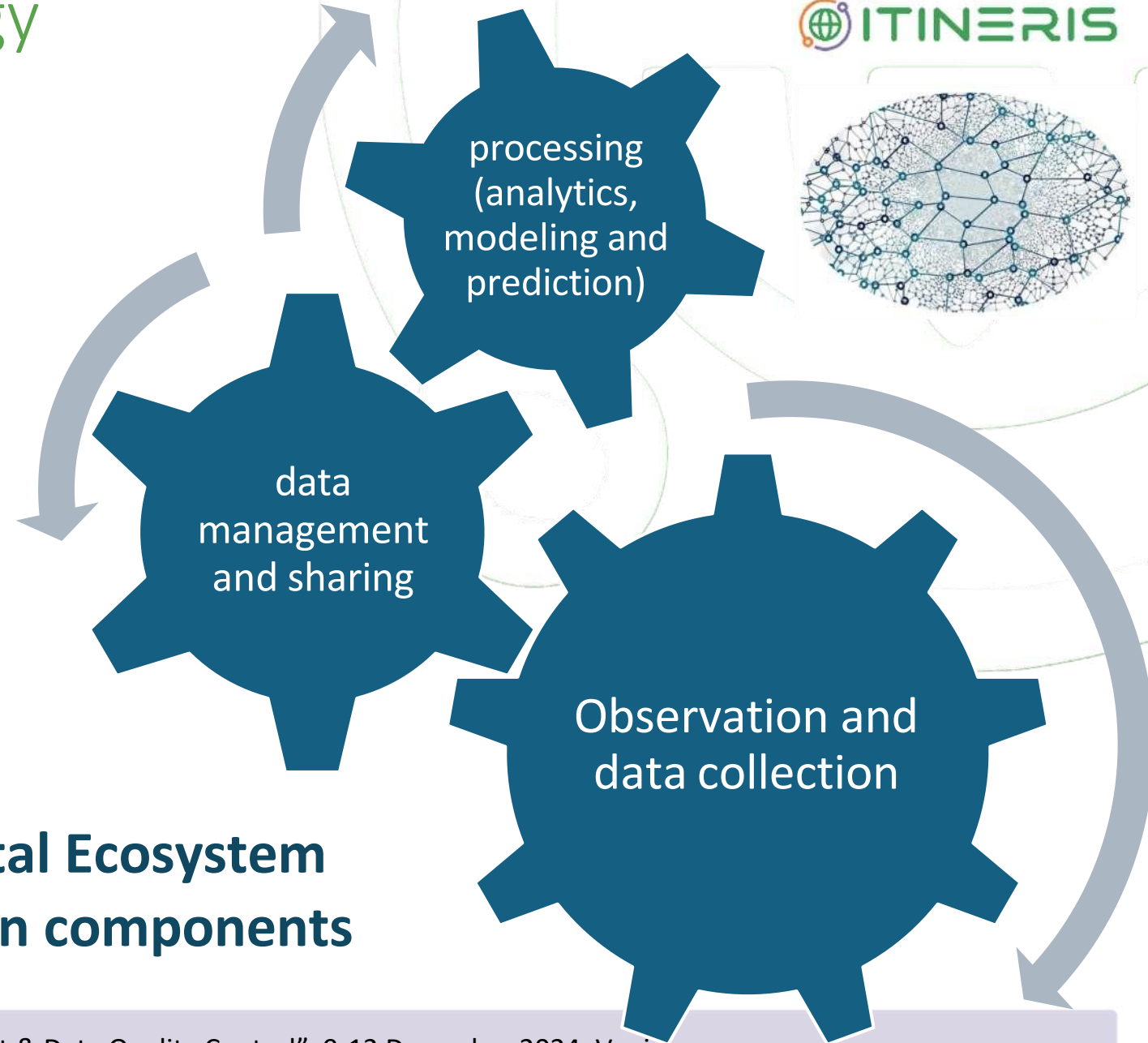
Need to put in place a transparent process for requesting, co-designing/developing, and sharing digital products and services for use by all: **Data Coordination Group (DCG)** coordinating the co-design and co-development process of these layers and case studies  
→ promote recommended practices, data standards, shared architectures and tools to foster ocean data sharing, application, and ocean prediction, leveraging innovations (artificial intelligence and digital twinning)

# Data & Information Strategy

Data Strategy Implementation  
Group (DSIG) →

how data systems can co-create a distributed, robust and collaborative **digital ecosystem** to be actively used for decision making and to support sustainable ocean management

## Digital Ecosystem 3 main components



## Strategic Objectives



**Develop an ocean digital ecosystem** that encourages the sharing and equitable access of multidisciplinary data, information and knowledge by all.



**Improve data and information discovery and usability** across the ocean digital ecosystem.



**Build trust in data and information** shared across the ocean digital ecosystem.



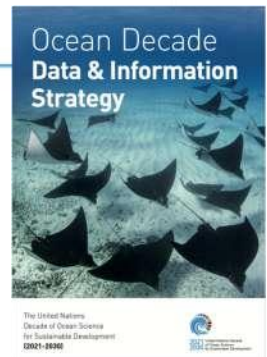
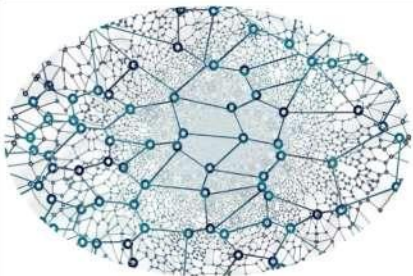
**Prioritise digital solutions that support decisions** for sustainable ocean management.



**Expand, empower, and mobilise global communities** to advance and maintain the ocean digital ecosystem.

## Enablers

logical Innovation // Partnerships // Durable Resourcing // Policy & Regulatory Frameworks



- 🌐 **Open science** and **FAIR** (Findable, Accessible, Interoperable and Reusable) principles drive the development of a digital ocean ecosystem capable of supporting the preservation and sharing of a wide array of data, knowledge and information products
- 🌐 **Ocean Best Practices** and a **FAIR digital asset** facilitate technological advances and community approaches for all methods to better understand and sustain our oceans
- 🌐 **Virtual Research Environments** will provide to the new generation of scientists an easy access to big amount of data, analytics and HPC resources to create innovative applications and tackle the emerging societal challenges (DTO - Digital Twin of the Ocean)

# (1) Ocean observation and data collection

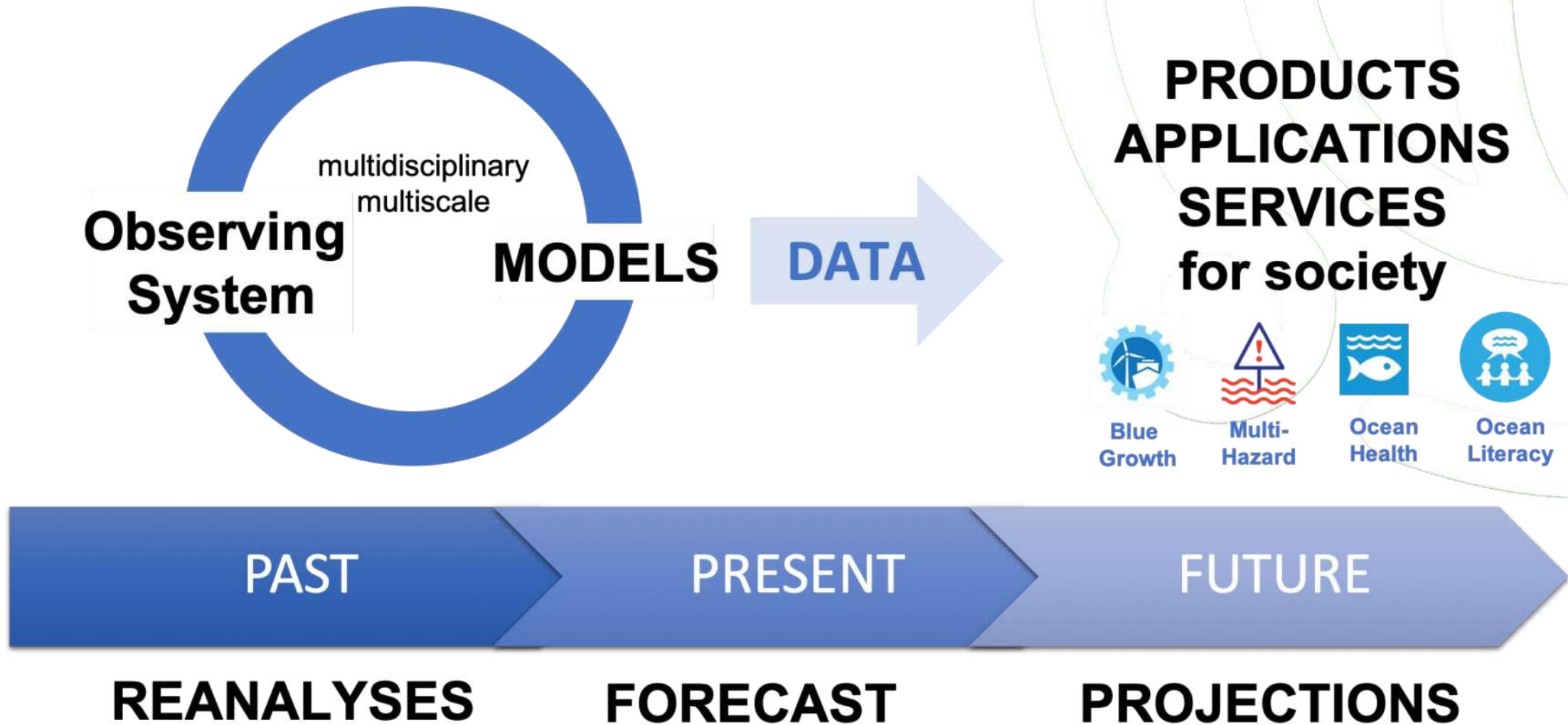
An **healthy, resilient and productive ocean** can be restored and preserved through constant and **systematic observation and monitoring of the marine environment** which allows to:

- 🌐 assess past and present ocean state, anticipate future scenarios
- 🌐 understand the ocean complex dynamics and evolution
- 🌐 mitigate the impact of human activity in a changing environment/climate
- 🌐 exploit the ocean natural resources sustainably

Ocean knowledge and information rely on a sustained **integrated multidisciplinary and multiscale ocean observing and monitoring system**

- systematic assessment of its adequacy to the emerging societal needs through the engagement of multiple stakeholders (institutional, academia, research, private sector, industry)
- **gaps identification for observing system optimization**

# Integrated Observing and Monitoring System



Simoncelli et al. 2022 <https://doi.org/10.1016/B978-0-12-823427-3.00001-3>

## (2) Data management and sharing

The management of the deriving marine data represents a societal priority

- data sharing approach and **open data** policies maximize users' uptake, promoting knowledge and innovation, reducing costs, avoiding duplication of efforts
- adoption of **common vocabulary, formats and standards** enable a rapid and efficient access to the data
- **FAIR principles** (Findable Accessible Interoperable and Reusable - *Wilkinson et al., 2016*) and **linked data approach** are the foundation of this revolution to reduce uncertainties but preserving transparency and traceability

# Global Ocean Observing System (GOOS)

- 🌐 led by the Intergovernmental Oceanographic Commission (IOC) of UNESCO, and co-sponsored by the World Meteorological Organization (WMO), the United Nations Environment Programme (UNEP) and the International Science Council (ISC) to support a community of international, regional and national ocean observing programmes, governments, UN agencies, research organizations and individual scientists
- 🌐 **goal: implementation of an integrated and sustained ocean observing system to deliver maximum impact for our user base and society**
- 🌐 with **systems approach** designed to be flexible and adapt to evolving scientific, technological and societal demands
- 🌐 fit-for-purpose: efficient data flow on **Essential Ocean Variables (EOVs)**, information services and applications that benefit the end users

<https://goosocean.org>

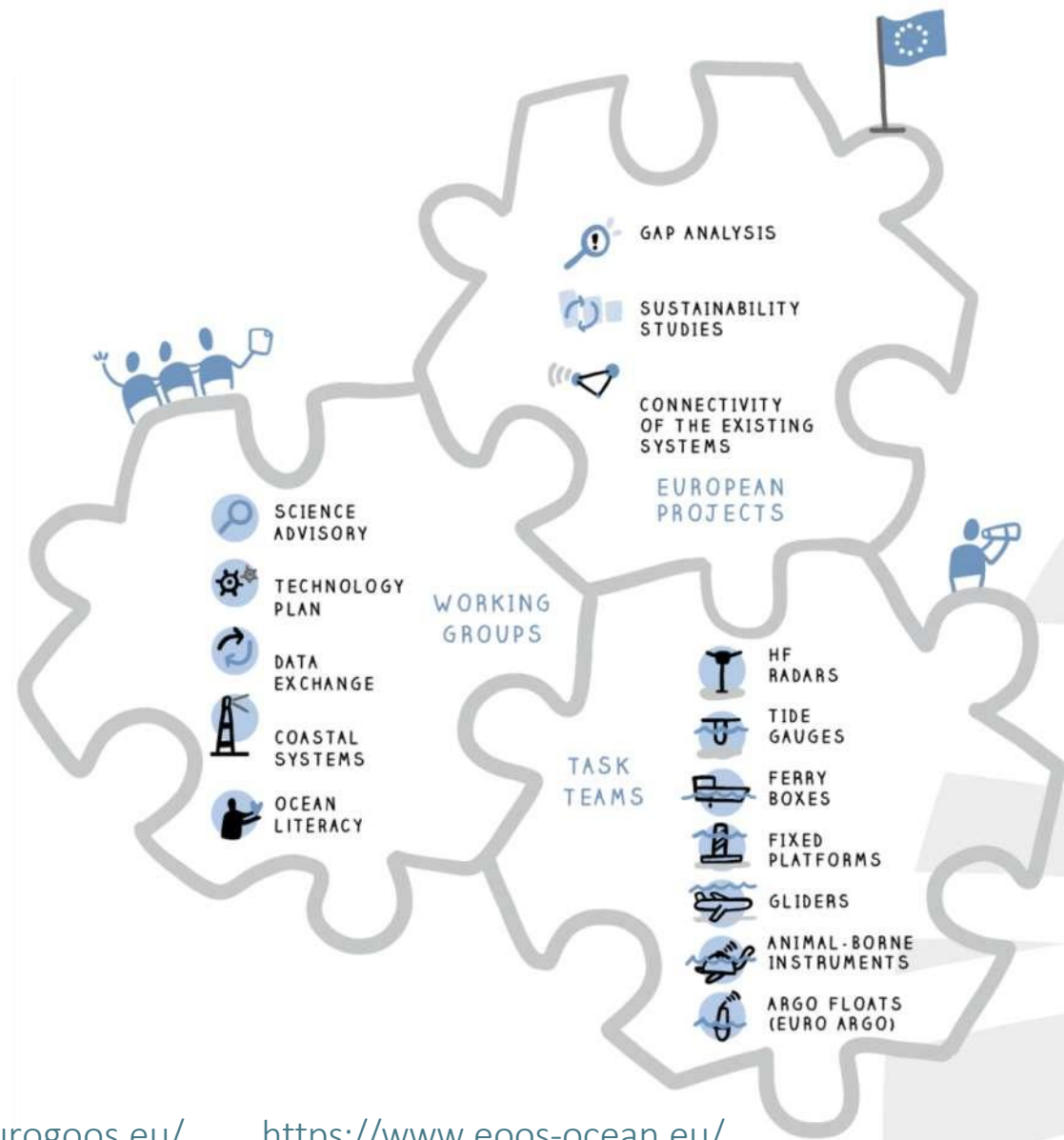


- 🌐 There are currently ~ 8000 in situ ocean observing platforms monitored by the GOOS operational centre OceanOPS
- 🌐 systems of buoys, moorings, profiling floats, ships (merchant, academic, fishing), gliders and other uncrewed systems (e.g., sail drones), drones, planes, land-based stations (e.g., global sea level), satellites, HF radars, deep sea cables, and animal borne sensors measuring mostly physical data and some biogeochemical data
- 🌐 14 global **ocean observing networks** that vary in stage of implementation
- 🌐 GOOS is organised in thirteen GOOS Regional Alliances (GRAs)



# EuroGOOS

- 🌐 European component of the GOOS
- 🌐 48 members
- 🌐 5 Regional Systems
- 🌐 Working Groups → Data Management WG
- 🌐 **networks of observing platforms** (Task Teams)
- 🌐 cooperation to unlock quality marine data and deliver common strategies, priorities and standards towards an integrated, sustained and fit-for-purpose European Ocean Observing System (EOOS)
- 🌐 EOOS framework is driven by the European community to better coordinate Europe's ocean observing capacity



<https://eurogoos.eu/>

<https://www.eoos-ocean.eu/>

# EU *in situ* marine data landscape

[https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/PDF/EUInSituMarine\\_EMODnet\\_CMEMS\\_FINAL.pdf](https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/PDF/EUInSituMarine_EMODnet_CMEMS_FINAL.pdf)

- 🌐 EU is strongly committed to the development of a coordinated system of global *in situ* ocean observations and has mature capabilities surrounding *in situ* ocean observations and marine monitoring
- 🌐 two cornerstone data services of the EC: **EMODnet** and the **Copernicus Marine Service**
- 🌐 **EMODnet** is a key marine knowledge initiative of **EC DG MARE**, with a core mandate for *in situ* marine data services delivered at a multi-disciplinary pan-European scale
- 🌐 **Copernicus Marine Service**, led by **EC DG DEFIS**, provides the EU with a world-leading capacity for monitoring the ocean worldwide, through satellites, targeted *in situ* observations, and numerical models
- 🌐 complementarity of their offers, that are developed and evolve in close interaction and synergy (high-level agreement between DG DEFIS and DG MARE and Memorandum of Understanding in 2016)
- 🌐 **Marine In Situ Coordination Inter-service Working Group (2021)**: experts from EMODnet, CMS and the EuroGOOS operational oceanography networks → to optimise data flows, coordination, and synergies of *in situ* marine data
- 🌐 working together to provide the **backbone data infrastructure and the core data streams into a common data lake underpinning the European Digital Twin Ocean (DTO)**

# EU *in situ* marine datascape

[https://cinea.ec.europa.eu/featured-projects/emodnet-catalyst-digital-twin-ocean-and-blue-innovation\\_en](https://cinea.ec.europa.eu/featured-projects/emodnet-catalyst-digital-twin-ocean-and-blue-innovation_en)



## upstream observation infrastructures

upstream global, pan EU and coastal *in situ* obs networks (GOOS, EuroGOOS, **EU Marine Research Infrastructures**), and associated data assembly and processing infrastructures (National Oceanographic Data Centres, SeaDataNet)  
→ need to be consolidated and sustained at the EU level as part of the EOOS (European Ocean Observing System) initiative

## data aggregator services

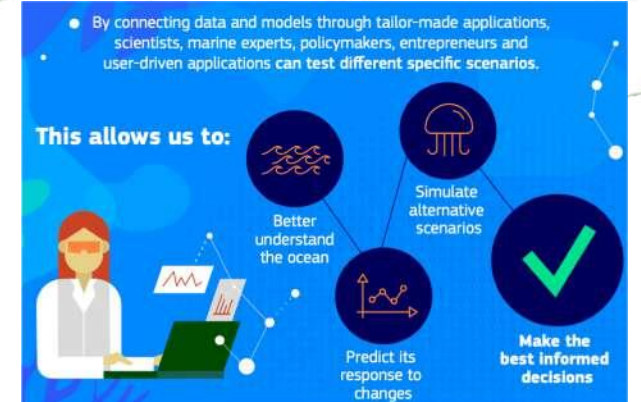


## European Digital Twin Ocean

An interactive replica of the ocean for better decision-making

### What is it?

A digital space providing access to vast amounts of data, models, artificial intelligence and other tools, which will allow the replication of the properties and behaviours of marine systems, including ocean currents and waves, marine life and human activities, and their interactions, in and near the sea.



# EMODnet: *in situ* marine data service

<https://emodnet.ec.europa.eu/en>

EMODnet serves users in  
policy, research, industry, and society,  
the EU Digital Twin Ocean and  
global ocean data initiatives

FAIR data layers and added value  
data products with metadata,  
standardised, harmonized and  
validated according to  
international standards and  
openly and freely available



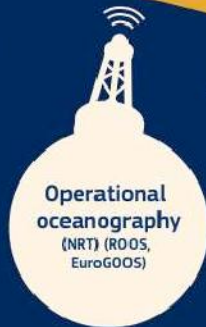
EC marine data service developing  
pan-European data products\*  
and aggregated datasets



Marine *in situ* data pipeline to EMODnet:  
National Oceanographic Data Centres, regional  
and European data infrastructures and services



*in situ* ocean observations (collected in  
water/air/seafloor) are the foundation  
of the EMODnet marine data pipeline



\*All EMODnet data products are published under an open data licence

# One central map viewer

to visualise all EMODnet data

<https://emodnet.ec.europa.eu/geoviewer/#/>

120  
partners

# 1 OCEAN 1 EMODnet

One single portal

# One central metadata catalogue

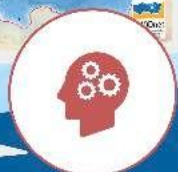
to enhance data search and discovery

+100  
use cases

Discover, visualise and  
download marine data and products  
across 7 thematics and hundreds of parameters



BATHYMETRY



HUMAN ACTIVITIES



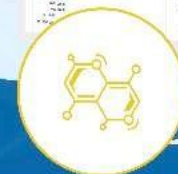
PHYSICS



GEOLOGY



SEABED HABITATS



CHEMISTRY



BIOLOGY

## EMODNET.EC.EUROPA.EU

YOUR GATEWAY TO *IN SITU* MARINE DATA IN EUROPE AND BEYOND



The European Marine Observation and Data Network (EMODnet) is financed by the European Union under Regulation (EU) 2021/1139 of the European Parliament and of the Council of 7 July 2021 establishing the European Maritime, Fisheries and Aquaculture Fund.



EMODnet

European Marine Observation and Data Network



# Copernicus Marine Service

Providing free and open marine data and services to enable marine policy implementation, support Blue growth and scientific innovation.

Access Data >

## DATA

### OCEAN PRODUCTS

A robust ocean data catalogue, to download or visualise data including hindcasts, nowcasts and forecasts.

## EXPERTISE

### OCEAN STATE REPORT

Extensive annual analysis on the state of the ocean over nearly 20 years and severe/notable annual events.

## TRENDS

### OCEAN CLIMATE TRENDS

Monitoring the health of the ocean.  
[Ocean Monitoring Indicators](#)  
[Ocean Climate Portal](#)

## EXPLORATION

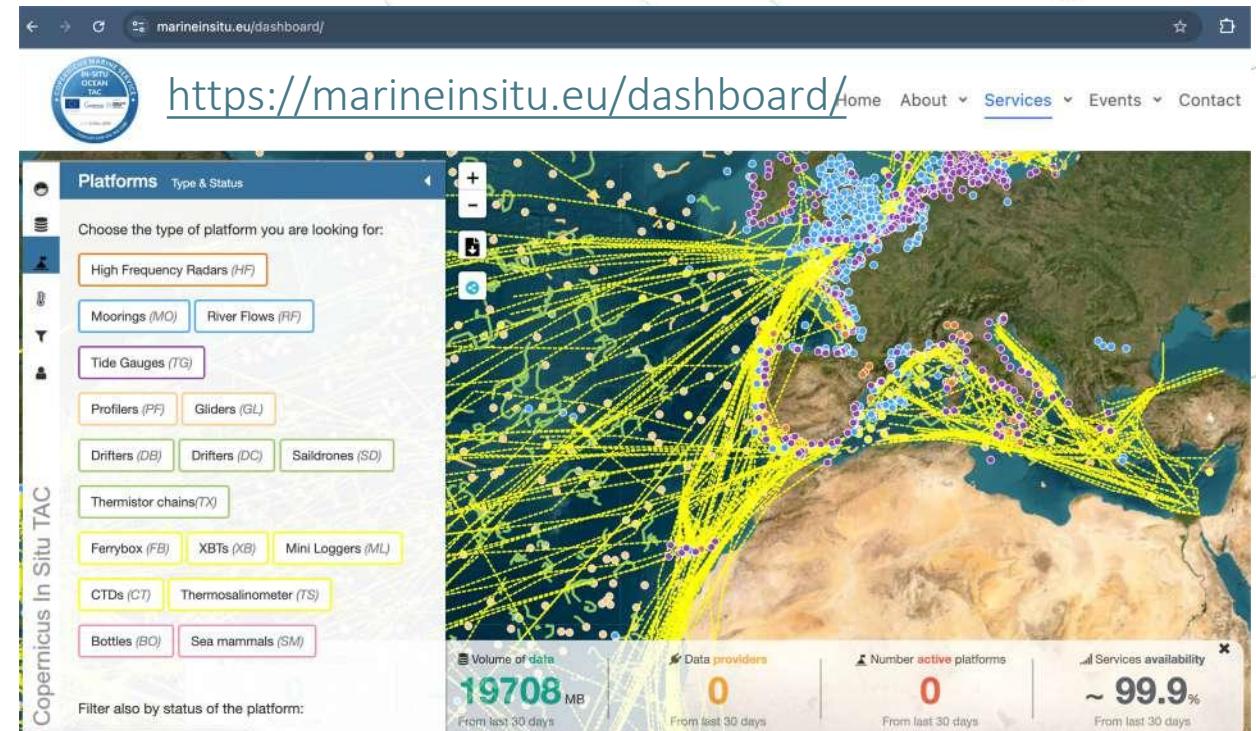
### OCEAN VISUALISATION

Dive into our 4D digital oceans through our 3 visualisation tools for beginner, intermediate and advanced users

*In situ* observations play a key role in constraining global and regional ocean analysis and forecasting systems, are collected, quality-controlled and delivered by a dedicated service-unit → **In Situ TAC**

- based on a distributed model of production, and developed in coherence with other components of the pan European data management landscape
- more than 350 data-providers contributing to a **targeted core set of parameters (temperature, salinity, currents, sea level, waves, chlorophyll-a, oxygen, nutrients, and carbon) delivered in near-real time and in delayed mode**, and directly useable for: initialization, forcing, assimilation and validation of ocean numerical models, validation of satellite products
- provides thorough quality control and assurance based on internationally agreed standards and best practices

## INSTAC Dashboard



## EU DTO: EDITO



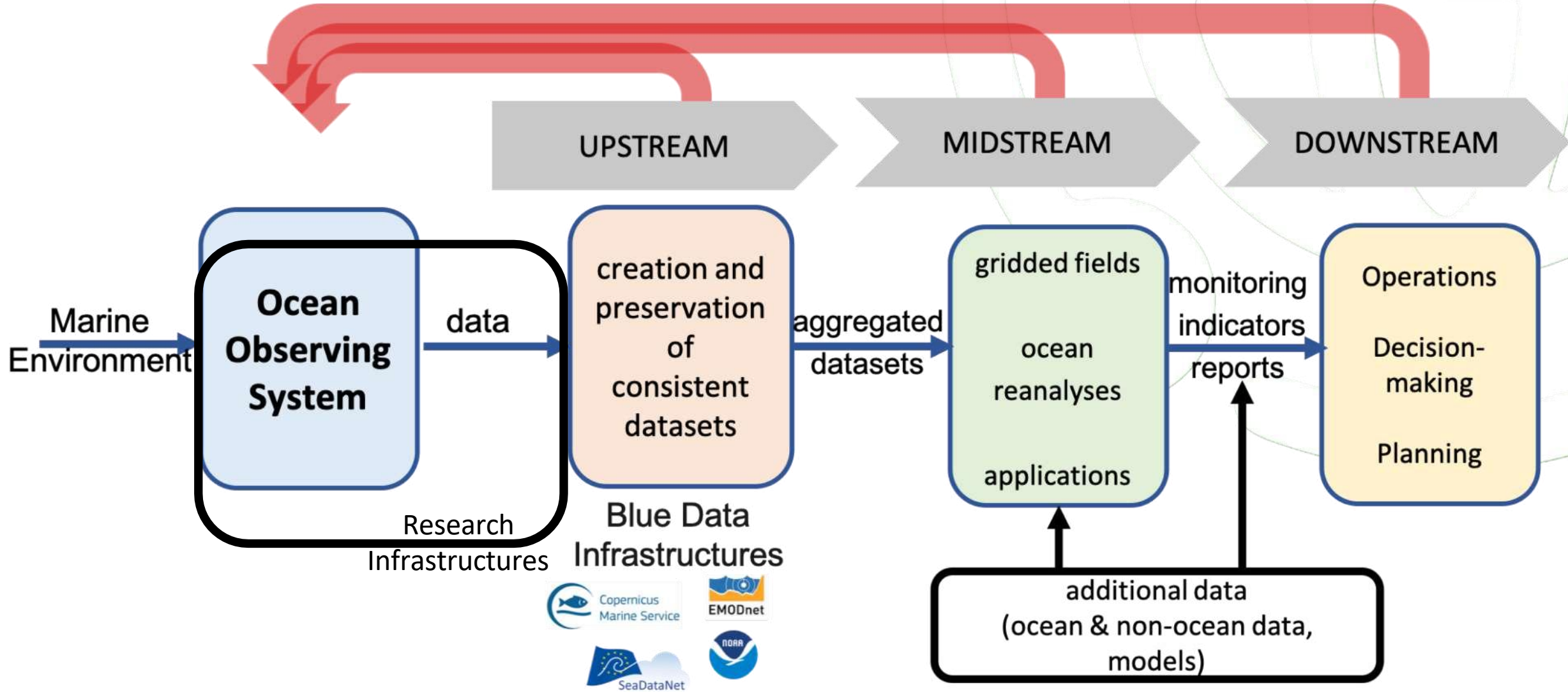
<https://www.edito.eu/>

<https://events.edito.eu/2024-digital-ocean-forum/content/eu-dto-platform>

- 🌐 announced in Feb 2022: EC has been investing (*Mission Restore our Ocean and Waters work programme*) ~€15M annually since 2021 to develop it
- 🌐 **It will transform data into knowledge:** consistent high-resolution, multi-dimensional descriptions of the ocean, its physical, chemical, biological, socio-ecological and economical dimensions, with forecasting periods ranging from seasons to multi-decades → **what if scenarios**
- 🌐 **EDITO** is an open infrastructure, conceived to facilitate the development of digital twin applications
- 🌐 co-designed and co-developed with all interested parties, who may also contribute with data, models, and intermediate or final services
- 🌐 The first prototype of the core infrastructure has been unveiled at the Digital Ocean Forum 2024 [https://maritime-forum.ec.europa.eu/events/digital-ocean-forum-2024-2024-06-13\\_en](https://maritime-forum.ec.europa.eu/events/digital-ocean-forum-2024-2024-06-13_en)



# Marine Data Value Chain



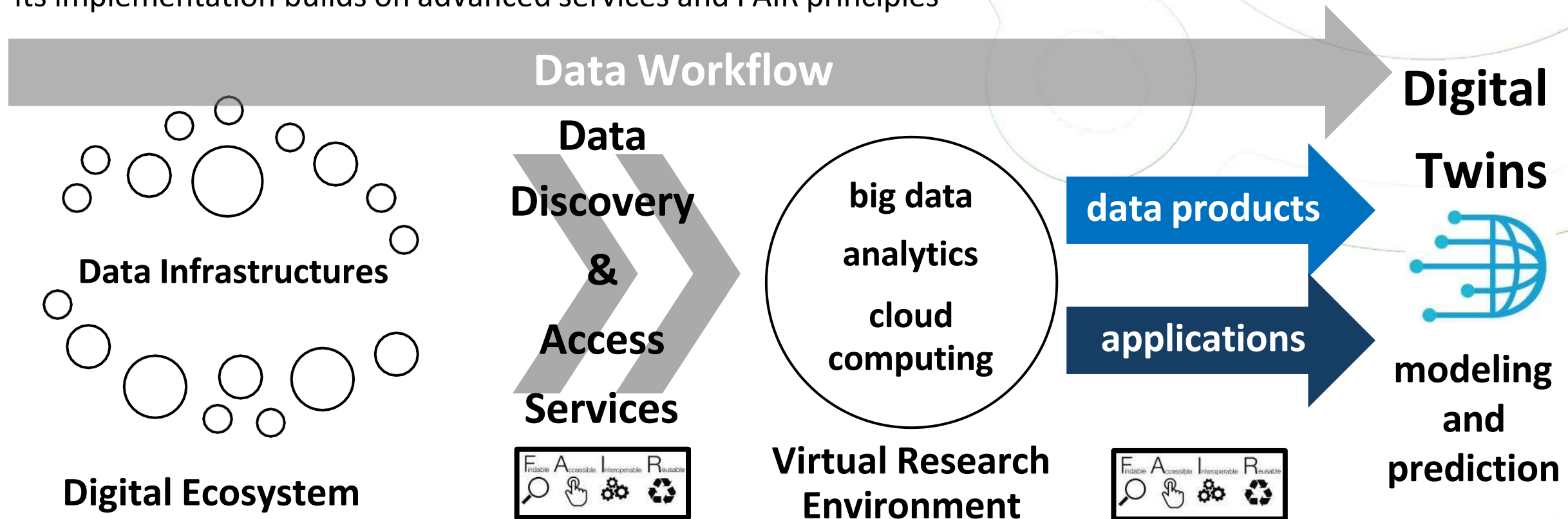
Simoncelli et al. 2022 <https://doi.org/10.1016/B978-0-12-823427-3.00001-3>

# (3) Processing (analytics, modeling and prediction)

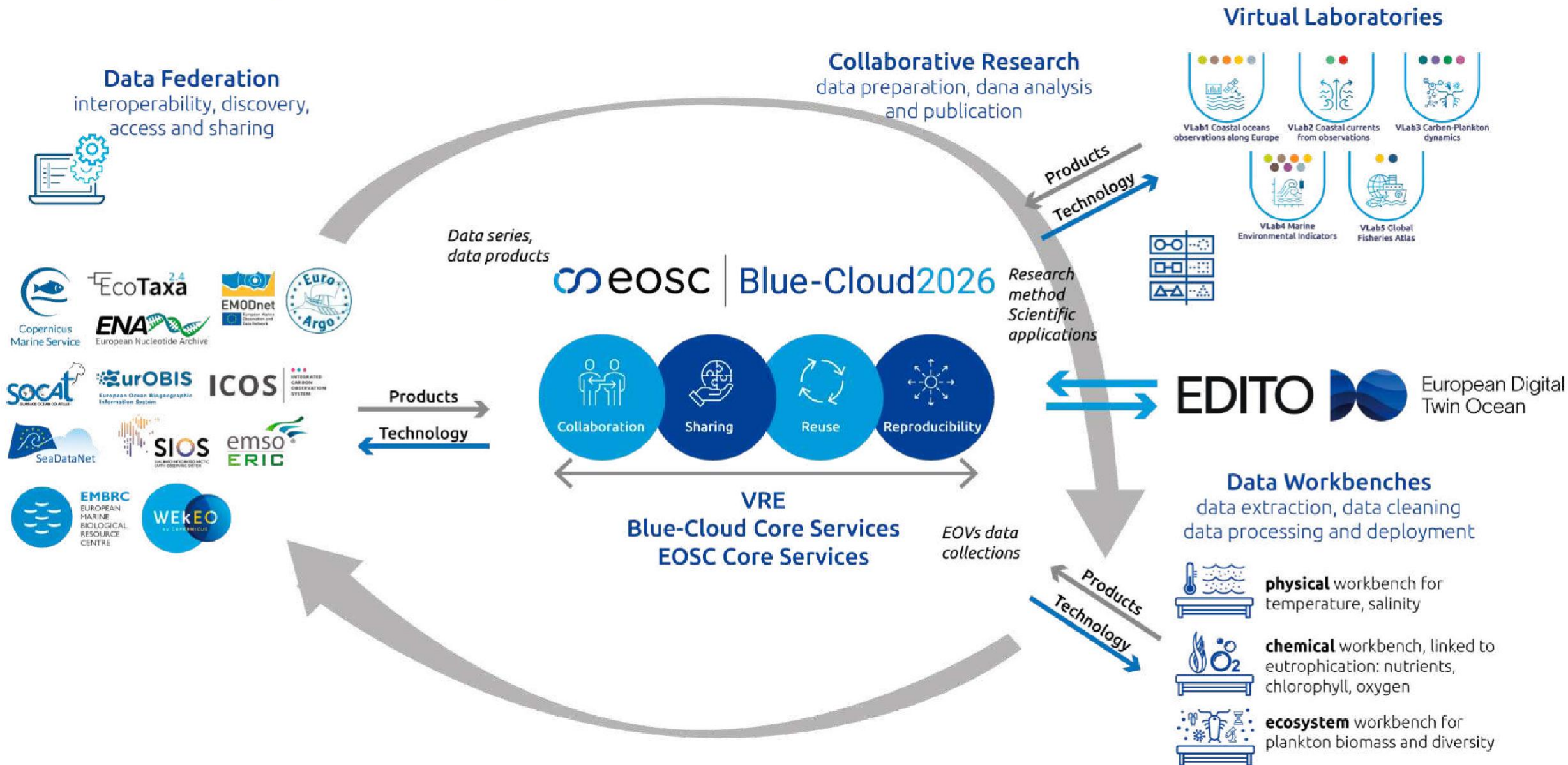
Workflow → a series of activities/processes that are necessary to complete a task

→ its **automation and management process** increase efficiency, optimize the results, allows its replicability and monitoring

→ its implementation builds on advanced services and FAIR principles

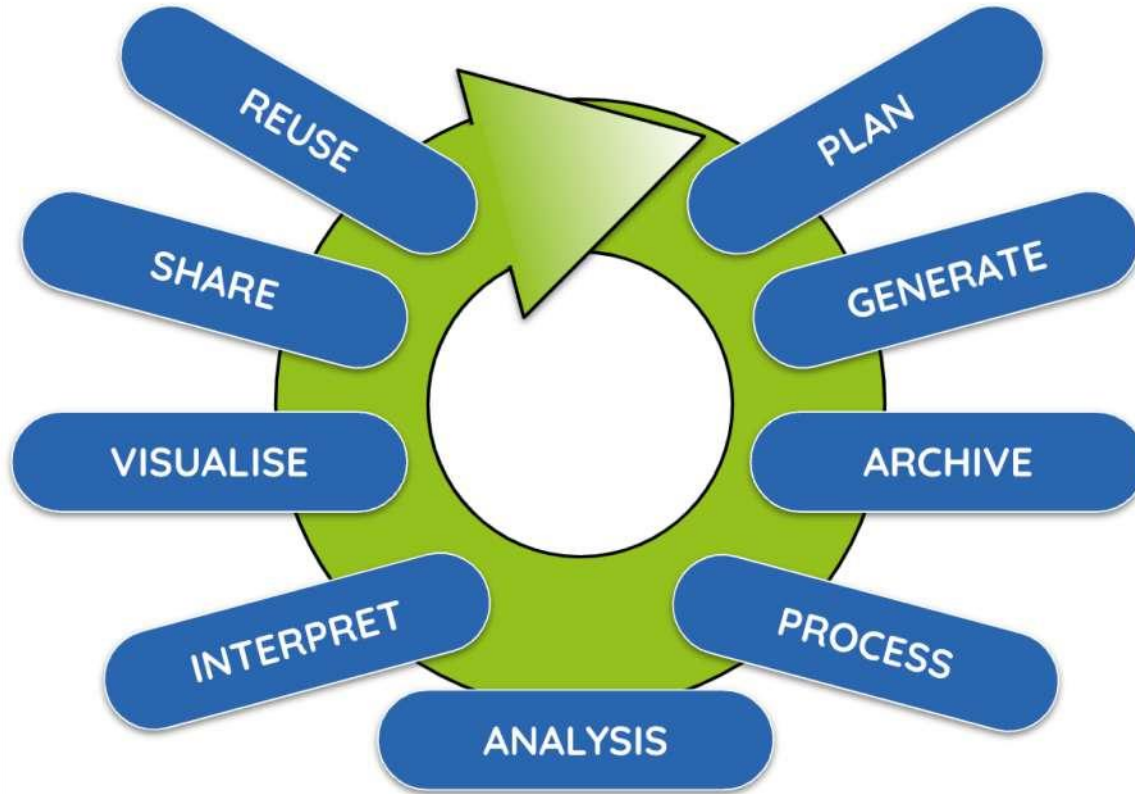


# Blue Cloud in a nutshell



Training current future RI research staff: "Marine Data Management & Data Quality Control", 9-12 December 2024, Venice

# Research Data Life Cycle



adapted from Locati & Cacciola (2024)

Research Data Management

<https://istituto.ingv.it/ufficio-gestione-dati>



## Research Data Management:

Organization, storage, preservation, and sharing of data in a research environment throughout the **entire data life cycle and beyond**.

RDM ensures the efficiency, transparency, and reproducibility of research and include **practices and policies that aim to improve the quality and value of research data**.

- 🌐 Traditional view of data (collecting, processing, analyzing, and publishing of results) substituted by a **life cycle approach that highlights the importance of finding, storing, and sharing data**
- 🌐 **Prior planning** has become mandatory to document data creation, content, context, but also to fulfill data quality requirements
- 🌐 Data quality requires predefined Quality Assurance (QA) strategies based on the selection of internationally validated methodologies for sampling and analysis, the mandatory use of reference materials

## The data providers (PI)

- 🌐 are responsible of the adequacy of the sampling strategy to the scope
- 🌐 must follow specific QA procedures and protocols applied before and during the dataset creation
- 🌐 have responsibilities in terms of documentation, calibration/intercalibration exercise, sampling strategy, admissible ranges of data, algorithms used, corrections, and flags

# QA/QC protocols and Standard Operating Procedures



**Best practices** have been developed from expert groups to have agreed and broadly adopted methods across ocean research, operations and applications

## Best Practice:

- is a methodology that has repeatedly produced superior results relative to other methodologies with the same objective
- is method that has been adopted and employed by multiple organizations

<https://www.oceanbestpractices.org/>

A screenshot of the Ocean Best Practices System website. At the top left is the "Ocean best practices" logo. To the right is a search bar with the word "Search" and a magnifying glass icon. Below the logo and search bar is a navigation menu with links: "ABOUT US", "NEWS AND EVENTS", "REPOSITORY", "COMMUNITY AND DEVELOPMENT", "OUR WORK", and "RESOURCES". The main content area features a large banner with a sunset over water. The banner text reads: "OCEAN BEST PRACTICES SYSTEM" in large white letters, followed by "Providing technological advances and community approaches for all ocean methods to better understand and sustain our oceans" in smaller white text. Below this is a link: "OBPS WORKSHOP VIII, 14-18 OCT 2024 HERE : RECORDINGS AVAILABLE HERE". At the bottom of the banner are three white boxes with blue icons and text: 1. A magnifying glass icon over a gear, with the text "SEARCH FOR PRACTICES". 2. An icon of four people silhouettes around a central person, with the text "SUBMIT A PRACTICE". 3. An icon of a database cylinder, with the text "EXPLORE OUR PROGRAMMES".

# Metadata

- 🌐 Records of monitored parameters must include a minimum set of information mapped through metadata
- 🌐 overview of the sensors and the methodologies (platform, instrument type, sensor's accuracy, calibration info)
- 🌐 measurement position, date and time
- 🌐 units
- 🌐 quality information (quality flags)

# Standard and Formats

- 🌐 controlled vocabularies
- 🌐 ISO 19115 metadata standards
- 🌐 Data Transport Formats
- 🌐 common QC protocols and flag scales

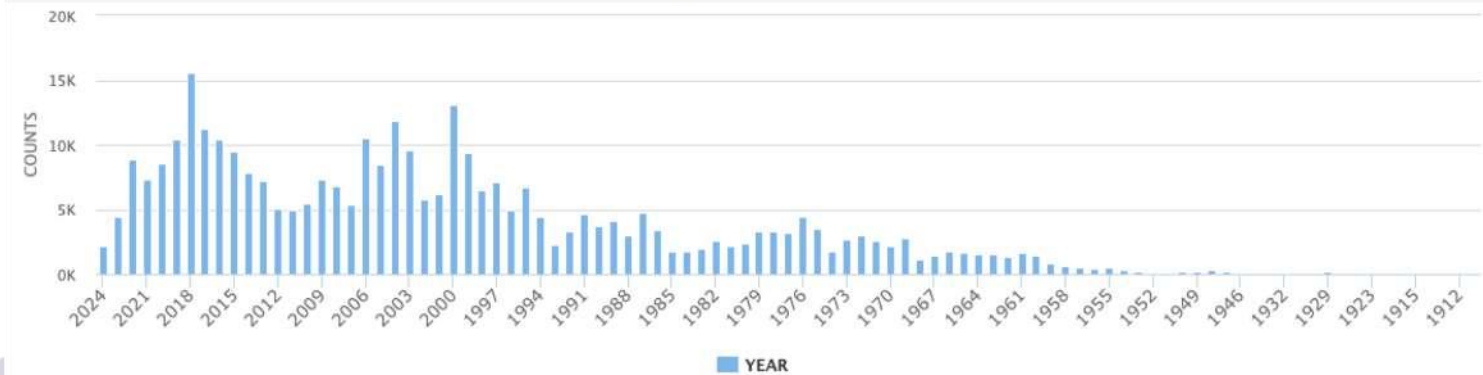
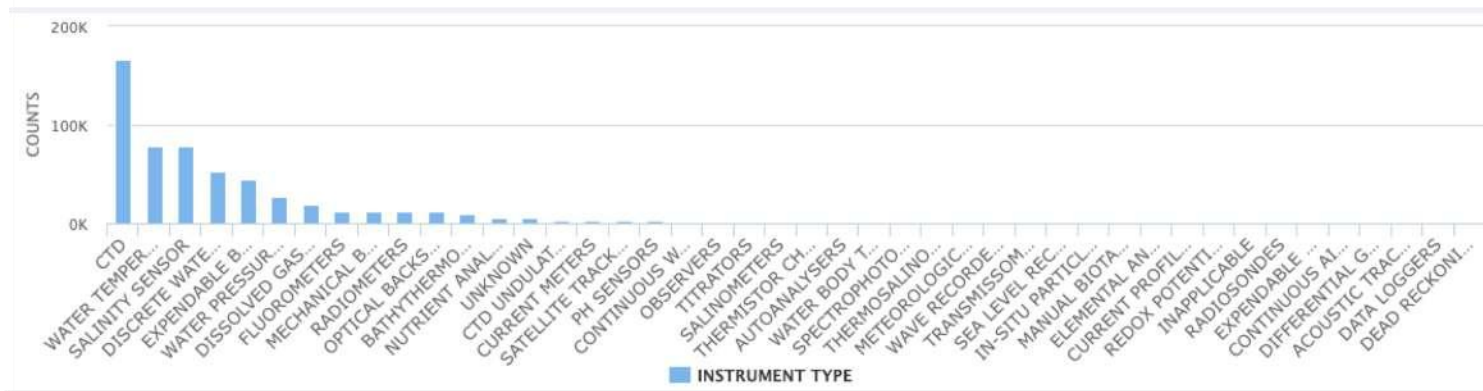
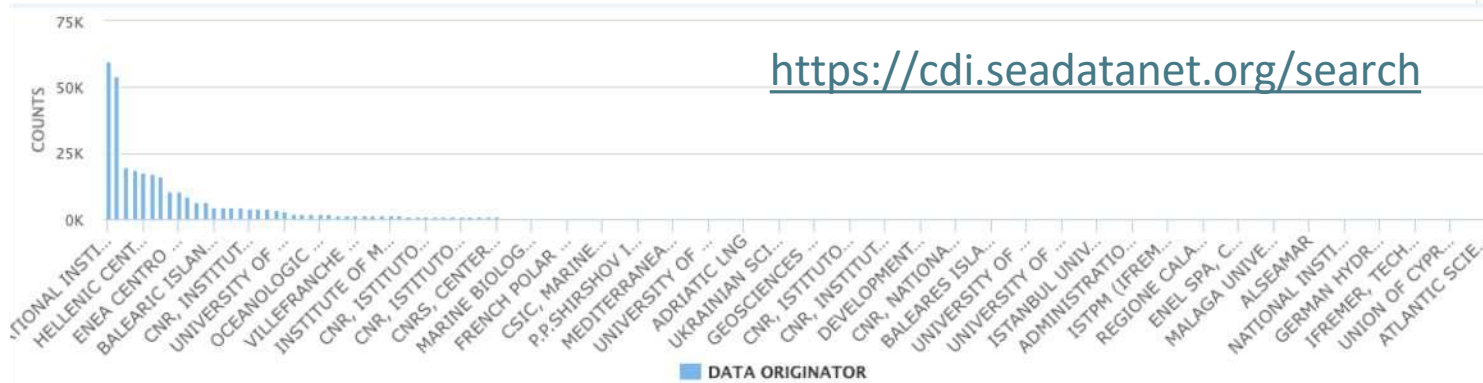
# The Power of Metadata

Metadata analysis can be used to monitor the **data sharing monitoring landscape**: data originator/distributor, instrument type, platform type, temporal distribution, data access restrictions

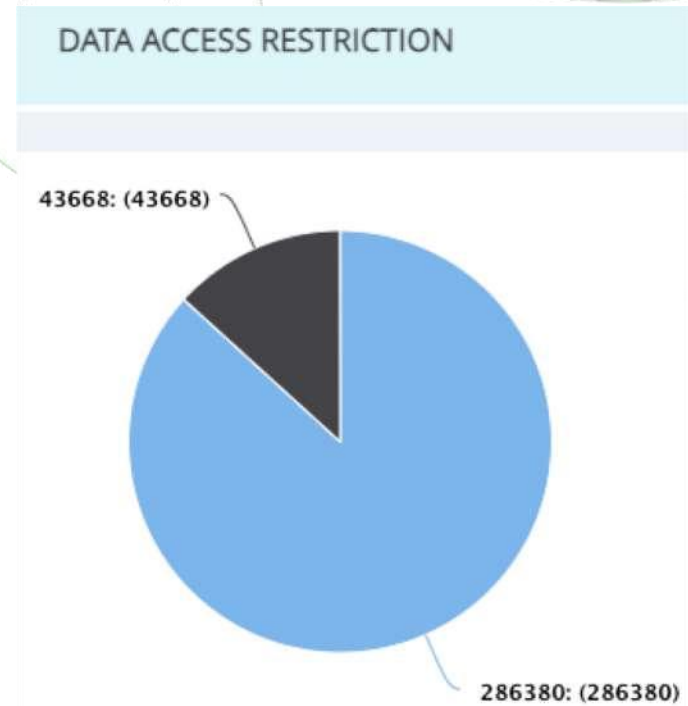
<https://cdi.seadatanet.org/search>

The screenshot displays the SeaDataNet search interface. At the top, there are navigation tabs: 'NEW SEARCH', 'REFINE SEARCH', and 'SEARCH RESULTS'. Below these, a 'SUMMARY OF YOUR SEARCH RESULTS' section lists various filters with corresponding bar charts and icons for grid and list views. The filters include: POINT OF CONTACT, DATA ORIGINATOR, DATA CUSTODIAN, DATA DISTRIBUTOR, PARAMETER GROUP (P03), DISCOVERY PARAMETER (P02), INSTRUMENT TYPE, PLATFORM TYPE, YEAR, and DATA ACCESS RESTRICTION. To the right of the filters is a map of Europe and the Mediterranean region, overlaid with red and orange data distribution patterns. A red box highlights a specific area in the Balkans. The interface also includes a 'Hello Simona SIMONCELLI My CDI' user profile, a 'DATASET BASKET' with a count of 0, and a vertical sidebar on the right with various navigation icons.

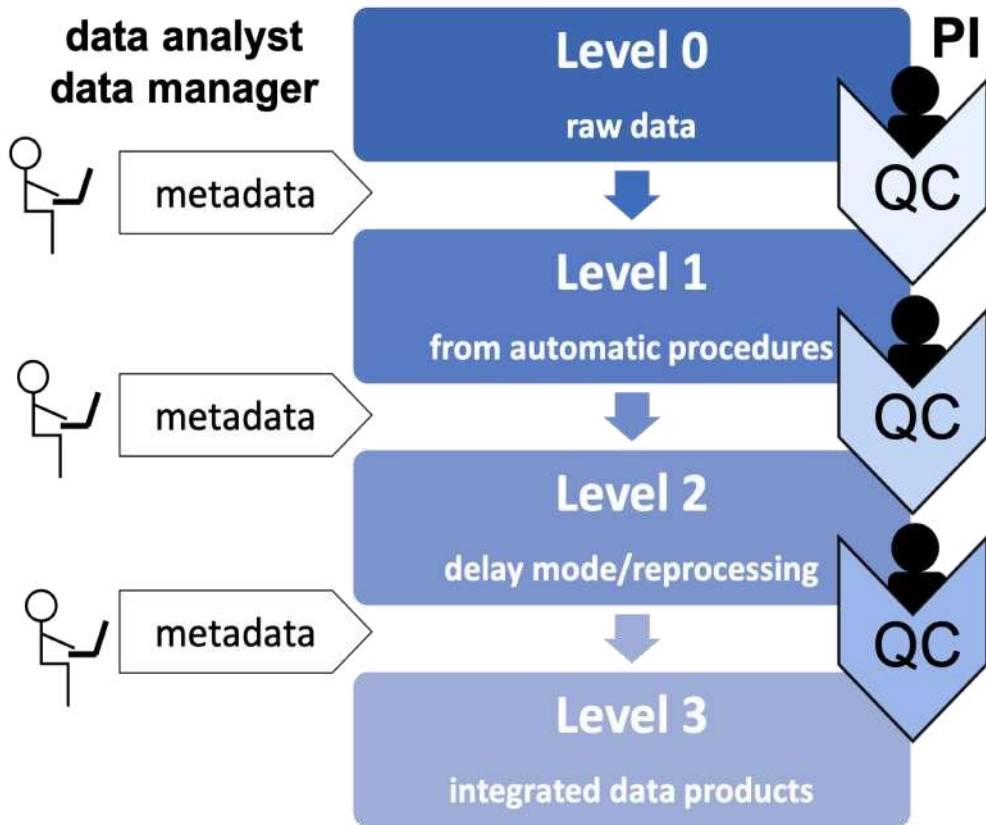
# The Power of Metadata



**Metadata analysis** from SeaDataNet data access portal on the specific data query results (i.e selected space-time domain, parameter)



# Data Processing Levels (DPL)



## Data Latency

Real Time

Near Real Time

Delay Mode



## Purpose

- ocean prediction
- early warning

- ocean state assessment
- climate studies

This could generate issues related to different data versions and duplicates within data infrastructures

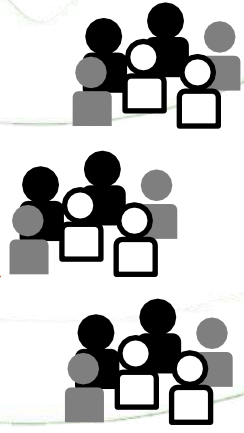
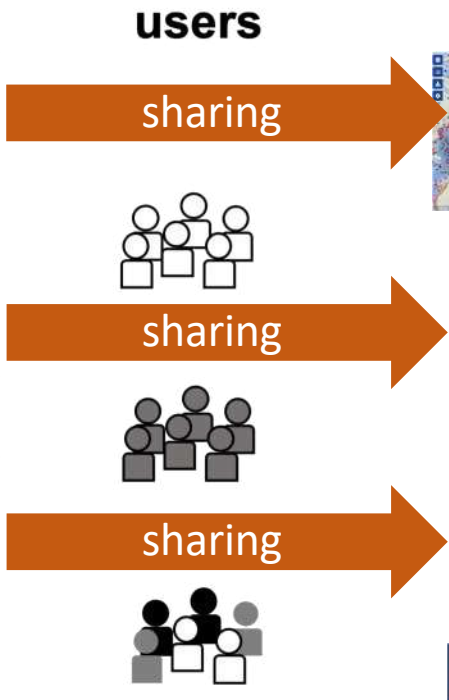
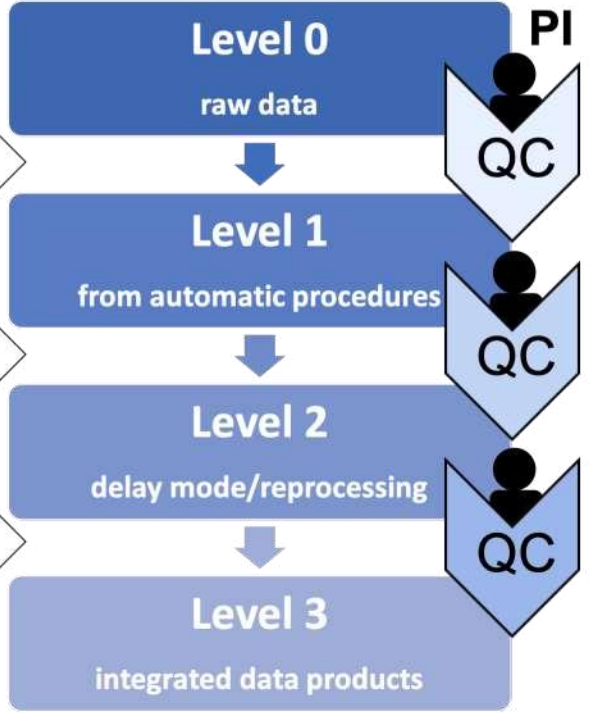
# Dataflow

## Data Provider

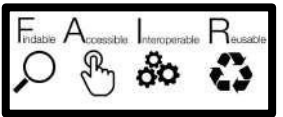
## Blue Data Infrastructures

## end users

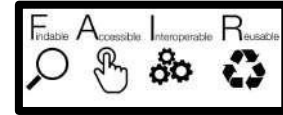
data analyst  
data manager



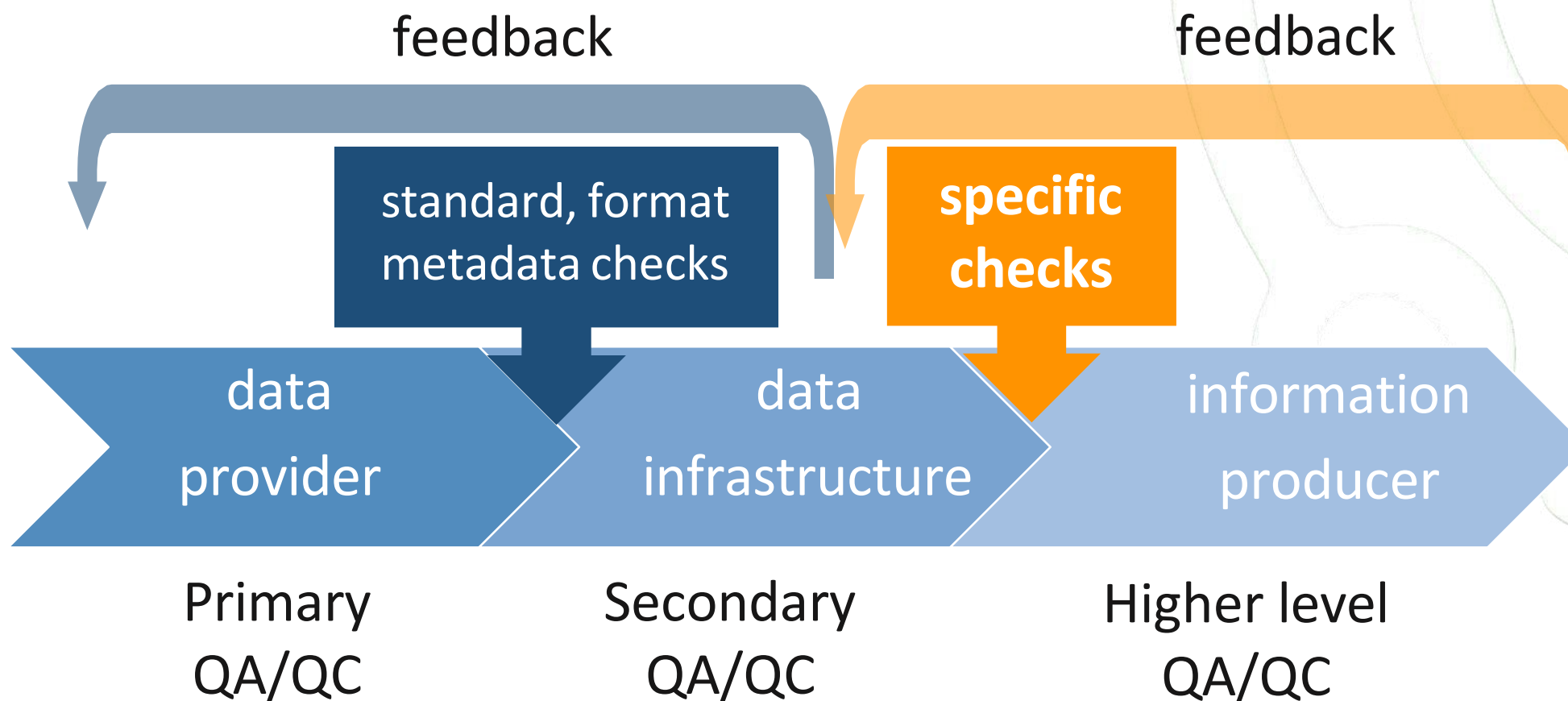
DOI  
dataset paper +  
scientific paper



data integration &  
metadata harmonization



# Quality Control

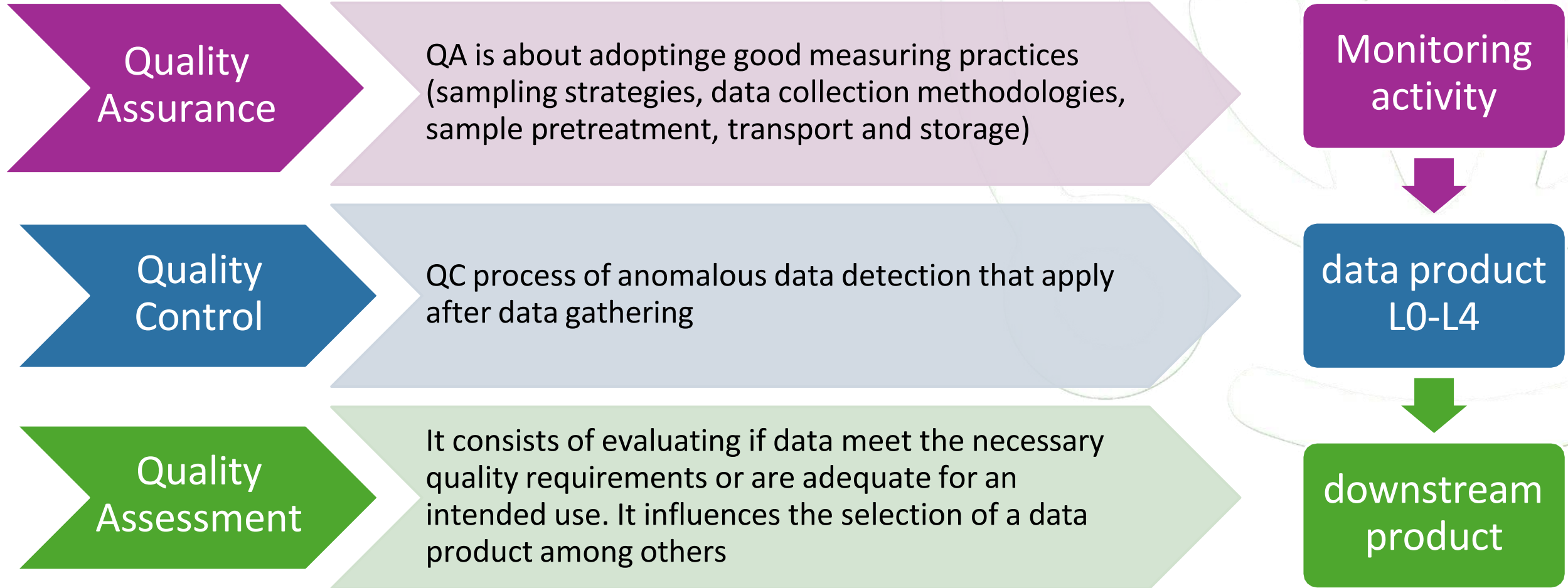


- **FAIRness**
- **integrity**
- **consistency**
- **completeness**

- several actors apply QA/QC procedures along with the data value chain
- **provenance** and **lineage information** are key elements to preserve

Simoncelli et al. (2022) <https://doi.org/10.1016/B978-0-12-823427-3.00001-3>

# about Quality



**Data accuracy and uncertainty** are key quality elements of data reliability

**Data uncertainty determination and its propagation along with the data value chain** is still a present challenge

It is very important to annotate data uncertainty in the metadata

# Quality Control

- 🌐 QC is vital for data reuse, and without it data from different sources cannot be combined to gain value
- 🌐 Scientific, analytical and statistical evaluations must determine if data present adequate quality to support the intended data usage, resulting in labeling each numerical value with a Quality Flag (QF) and avoid modifying the original data record following a harmonized scheme of QFs
- 🌐 QFs ensure that the quality of the data is apparent to the user, who holds sufficient information to decide the suitability for a specific task applying the proper data filtering

QC practices include:

- data integrity checks (e.g., format)
- data value checks (range checks, spikes and outliers checks, neighbor checks, climatology checks)
- QC highly depend on the data thematic, sensor type and the available amount of time for the analysis (RT vs DM)

Automatic QC (RT procedures) → algorithms

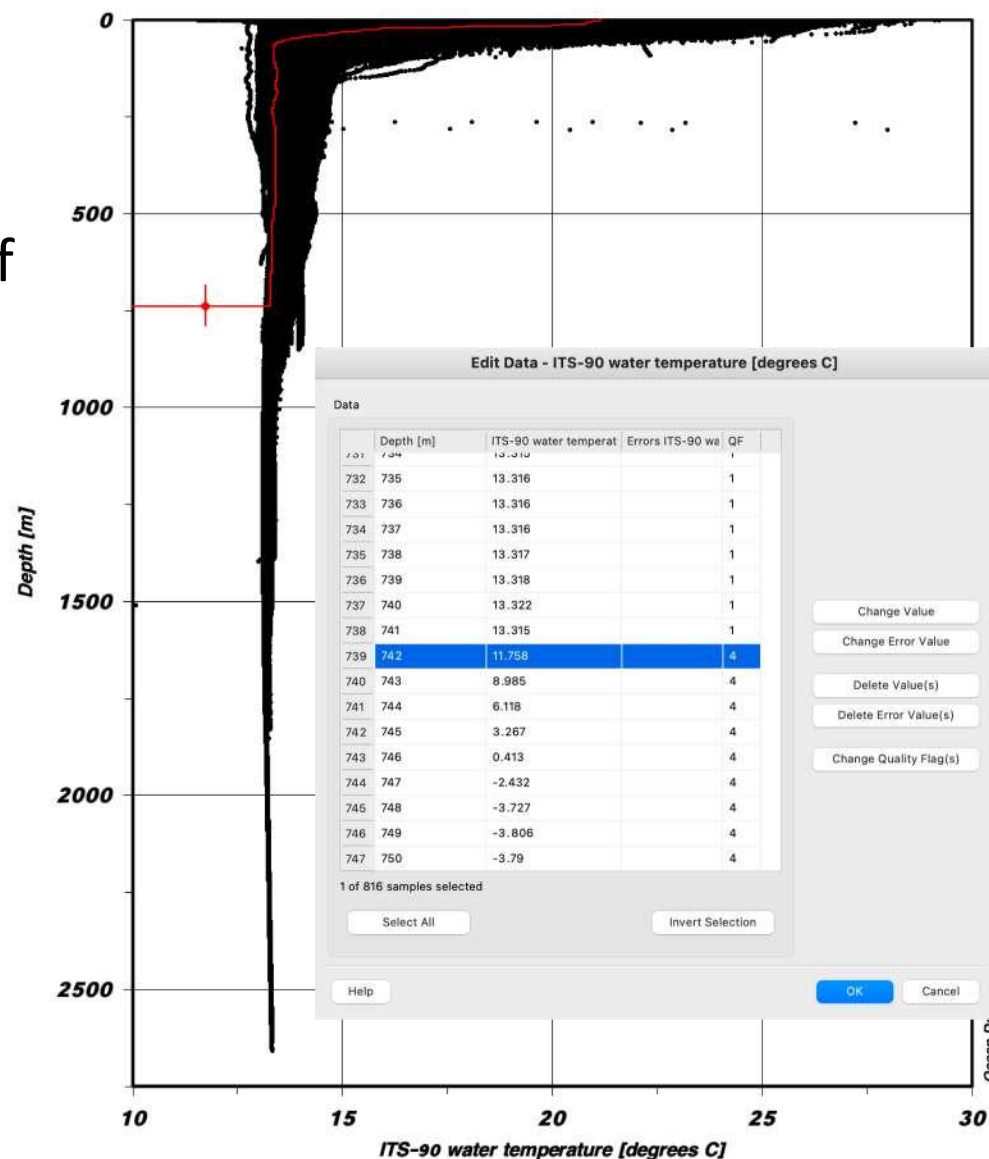
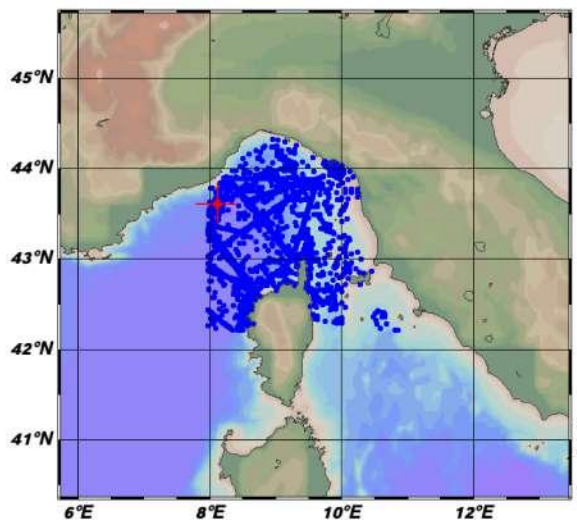
Visual QC (DM procedures) → visual tools (i.e. OceanDataView) and graphical interfaces

# Quality Flags

- Labels associated to each measurement that the user can use to filter data according to the needs
- SeaDataNet harmonised scheme of QC Flags to be used to label individual data values
- QC Flag scale is available in the SeaDataNet Common Vocabularies as list L20

Key	Entry Term	Abbreviated term	Term definition
0	no quality control	none	No quality control procedures have been applied to the data value. This is the initial status for all data values entering the working archive.
1	good value	good	Good quality data value that is not part of any identified malfunction and has been verified as consistent with real phenomena during the quality control process.
2	probably good value	probably_good	Data value that is probably consistent with real phenomena but this is unconfirmed or data value forming part of a malfunction that is considered too small to affect the overall quality of the data object of which it is a part.
3	probably bad value	probably_bad	Data value recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
4	bad value	bad	An obviously erroneous data value.
5	changed value	changed	Data value adjusted during quality control. Best practice strongly recommends that the value before the change be preserved in the data or its accompanying metadata.
6	value below detection	BD	The level of the measured phenomenon was too small to be quantified by the technique employed to measure it. The accompanying value is the detection limit for the technique or zero if that value is unknown.
7	value in excess	excess	The level of the measured phenomenon was too large to be quantified by the technique employed to measure it. The accompanying value is the measurement limit for the technique.
8	interpolated value	interpolated	This value has been derived by interpolation from other values in the data object.
9	missing value	missing	The data value is missing. Any accompanying value will be a magic number representing absent data.
A	value phenomenon uncertain	ID_uncertain	There is uncertainty in the description of the measured phenomenon associated with the value such as chemical species or biological entity.

**Ocean Data View (ODV)** software package for the interactive exploration, analysis and visualization of oceanographic and other geo-referenced profile, time-series, trajectory or sequence data



Edit Data - ITS-90 water temperature [degrees C]

Depth [m]	ITS-90 water temperat	Errors ITS-90 wa	QF
732	735	13.316	1
733	736	13.316	1
734	737	13.316	1
735	738	13.317	1
736	739	13.318	1
737	740	13.322	1
738	741	13.315	1
739	742	11.758	4
740	743	8.985	4
741	744	6.118	4
742	745	3.267	4
743	746	0.413	4
744	747	-2.432	4
745	748	-3.727	4
746	749	-3.806	4
747	750	-3.79	4

1 of 816 samples selected

Select All Invert Selection

Change Value Change Error Value Delete Value(s) Delete Error Value(s) Change Quality Flag(s)

OK Cancel

Station ID: 644562

Accession Num...	636844
Cruise	MOON2013
Station	2103 (C)
Longitude	8.133°E
Latitude	43.6°N
Date	21 September 2013
Time	18:05:00
Depth Range [m]	[4 - 819]
LOCAL_CDI_ID	XO20130921003_136_H13
EDMO_code	136
Bot.Depth [m]	0
Instrument Info	
P01 Codes in O...	SDN:P01::ADEPZZ01   SDN:P01::TEMPET01
P35 Contributo...	SDN:P35::WATERTEMP = [ SDN:P01::TEMPET01 ]
References	

Sample: 739 / 816

1: Depth [m]	742	1
2: ITS-90 water temperature [degrees ...	11.76	4
3: Water body salinity [per mille]		9
drvd: Potential Temperature $\theta$ [degC]		9
drvd: Potential Density Anomaly $\sigma_0$ [kg/ ...		9
drvd: Stability Ratio $R_p$		9
drvd: Dynamic Height-700 [dyn m]		9

Isosurface Values

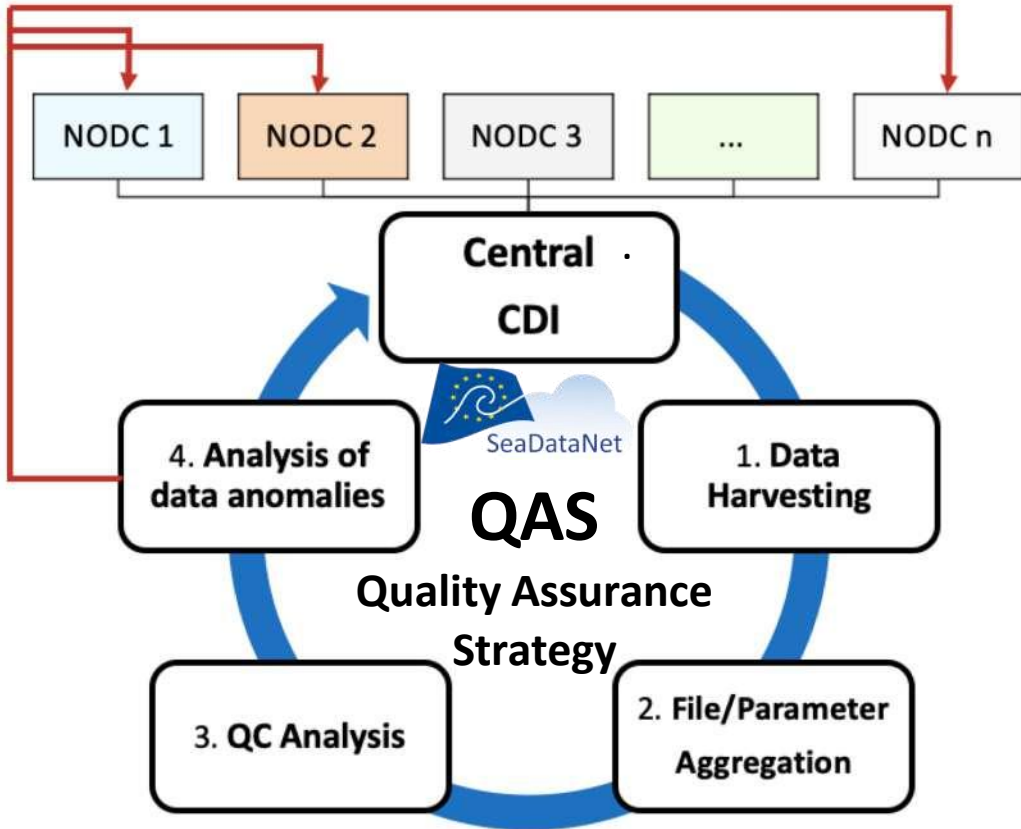
Longitude	8.133
Latitude	43.600
Time [yr]	2013.723
Day of Year	264
ITS-90 water temperature [degrees C] @ Depth [m]=150.00	13.48
ITS-90 water temperature [degrees C] @ Depth [m]=300.00	13.43
ITS-90 water temperature [degrees C] @ Depth [m]=600.00	13.35
ITS-90 water temperature [degrees C] @ Depth [m]=1000.00	
ITS-90 water temperature [degrees C] @ Depth [m]=2000.00	

<https://odv.awi.de/>

# Data Products (example)



network of SeaDataNet data centers

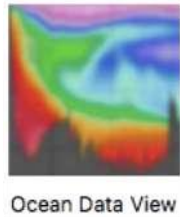
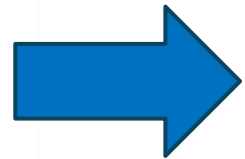
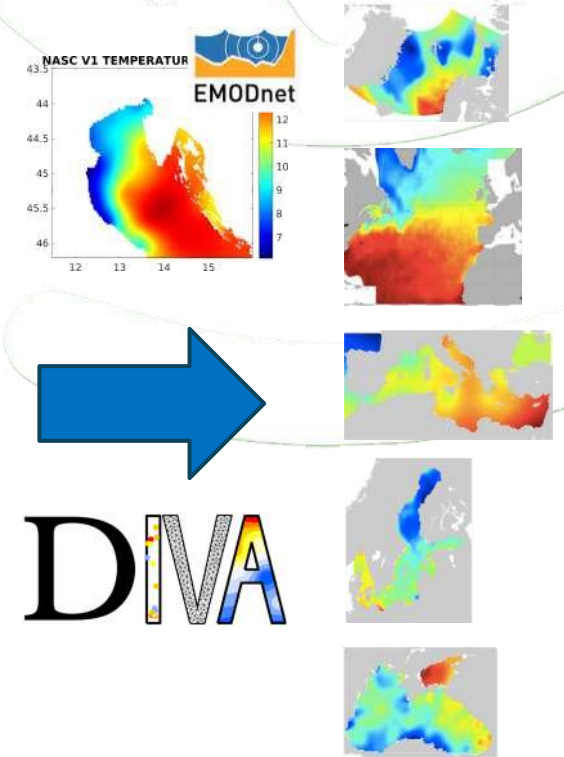
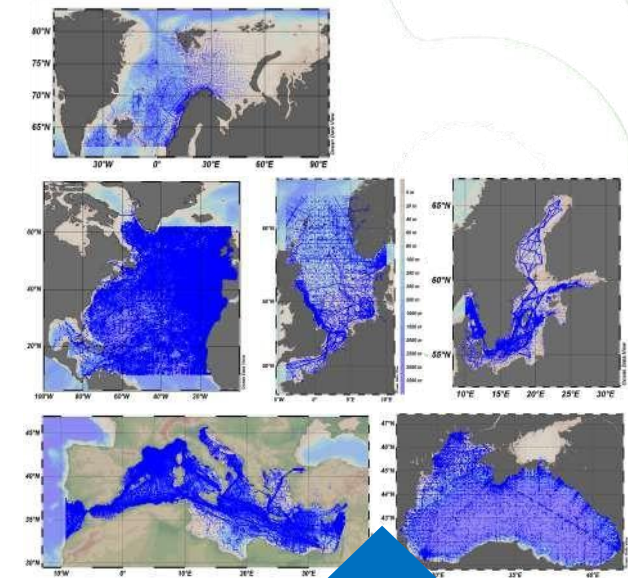


**T&S  
aggregated dataset**

**SDN catalog (DOI)  
+ Product  
Information  
Document (PIDoc)**



**T&S  
gridded  
climatologies**



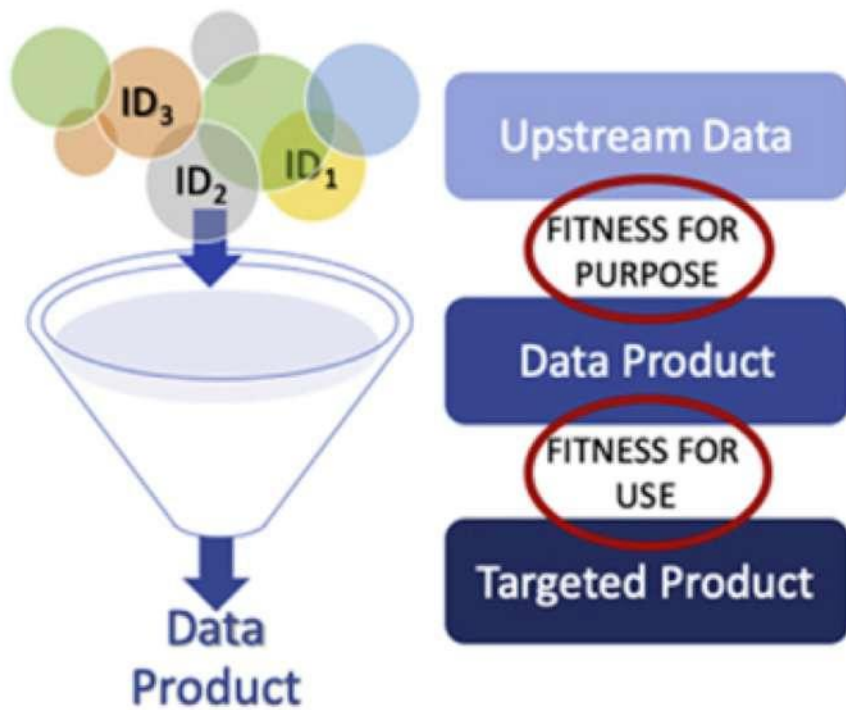
**merging data  
from other BDIs**



**+ OHC, MLD, density,  
stratification  
products**

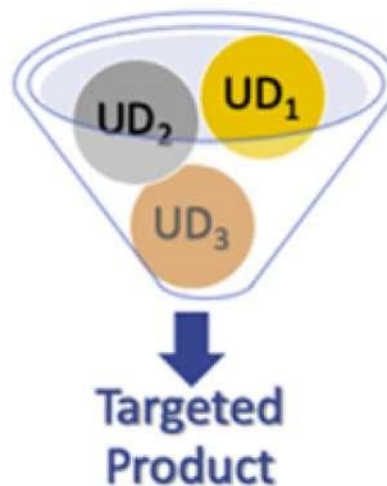
# Quality Assessment

## Input Data



(expected)

## available Upstream Data



(realized)

- User requirements define the **Data Product Specification** and the necessary **Input Data** to generate it
- the realized **Targeted Product Description** compared to the preliminary specification provide a measure of fit-for-use
- **Upstream Data** are fit-for-purpose if adequate for the targeted product realization
- Quality assessment process might provide a feedback on the **adequacy of observing system** for specific products/applications

Simoncelli et al. (2022) <https://doi.org/10.1016/B978-0-12-823427-3.00001-3>

**F**indable   **A**ccessible   **I**nteroperable   **R**eusable

*Wilkinson et al. (2016)*



**Findable:** easy to find by humans and computers thanks to metadata and unique persistent identifiers

**Accessible:** stored for easy access and downloading

**Interoperable:** ready to be combined with other datasets by humans and computer systems

**Re-usable:** ready for reuse thanks to detailed, accurate documentation and clear usage license

from Locati & Cacciola (2024) Research Data Management <https://istituto.ingv.it/ufficio-gestione-dati>

# FAIRness Assessment

## Findable

Metadata and data should be findable for both humans and computers

## Interoperable

Data needs to work with applications or workflows for analysis, storage and processing

**F A I R**

## Accessible

Once found, users need to know how the data can be accessed

## Reusable

The goal of FAIR is to optimise data reuse via comprehensive well-described metadata

FAIR proposed by the community <https://force11.org/info/the-fair-data-principles/>

<b>Findable</b>	
F1	(meta)data are assigned a globally unique and eternally persistent identifier.
F2	data are described with rich metadata.
F3	(meta)data are registered or indexed in a searchable resource.
F4	metadata specify the data identifier.
<b>Accessible</b>	
A1	(meta)data are retrievable by their identifier using a standardized communications protocol.
A1.1	the protocol is open, free, and universally implementable.
A1.2	the protocol allows for an authentication and authorization procedure, where necessary.
A2	metadata are accessible, even when the data are no longer available.
<b>Interoperable</b>	
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2	(meta)data use vocabularies that follow FAIR principles.
I3	(meta)data include qualified references to other (meta)data.
<b>Re-usable</b>	
R1	(meta)data have a plurality of accurate and relevant attributes.
R1.1	(meta)data are released with a clear and accessible data usage license.
R1.2	(meta)data are associated with their provenance.
R1.3	(meta)data meet domain-relevant community standards.

# FAIRness Assessment

**Data FAIRness** → useful for data providers to learn and adopt FAIR principles during data generation or to implement data FAIRification on existing datasets.

There exist subjective and objective methods:

1. RDA FAIR Data Maturity Model (FDMM) (self-evaluation)
2. F-UJI tool (automatic evaluation through DOI or URL)

FAIR principles and assessment is expanding to all «digital research objects» (software, semantic artifacts, VRE, ...)

**Software FAIRness** → indicators under development

*Chue et al. et al. (2022). FAIRPrinciples for Research Software version 1.0. (FAIR4RS Principles v1.0). Research Data Alliance. DOI: <https://doi.org/10.15497/RDA00068>*

# Example: CORA product

<https://www.f-uji.net/>

## FAIR assessment

F-UJI is a web service to programatically assess FAIRness of research data objects (aka data sets) based on metrics developed by the [FAIRsFAIR](#) project.

Please use the form below to enter an identifier (e.g. DOI, URL) of the data set you wish to assess. Optionally you also can enter a metadata service (OAI-PMH, SPARQL, CSW) endpoint URI which F-UJI can use to identify additional information.



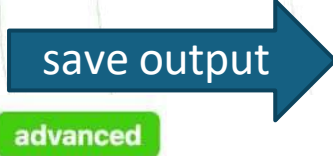
[About](#) [Feedback](#) [Privacy Policy](#) [Terms of Use](#) [Legal Notice](#)

F-UJI is a result of the [FAIRsFAIR](#) "Fostering FAIR Data Practices In Europe" project which received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 (grant agreement 831558).

# Example: CORA product

F-UJI requires the data identifier (URL, PID, or metadata endpoint service) → results are given according to FAIRsFAIR FAIR Data Object Assessment metrics

## CORA, Coriolis Ocean Dataset for Reanalysis



FAIR level: ?	
Resource PID/URL:	<a href="https://doi.org/10.17882/46219">https://doi.org/10.17882/46219</a>
DataCite support:	enabled
Metric Version:	metrics_v0.5
Metric Specification:	<a href="https://doi.org/10.5281/zenodo.6461229">https://doi.org/10.5281/zenodo.6461229</a>
Software version:	3.2.0
Download assessment results:	<a href="#">{JSON}</a>
Save and share assessment results:	Saved assessments: <ul style="list-style-type: none"> <li>FAIR 75% <a href="#">2023-04-28 (2.2.5)</a></li> <li>FAIR 75% <a href="#">2023-09-16 (2.2.5)</a></li> </ul>



### Summary:



	Score earned:	Fair level:
Findable:	6 of 7	advanced
Accessible:	3 of 3	advanced
Interoperable:	4 of 4	advanced
Reusable:	7 of 10	moderate

# FAIR Data Maturity Model (FDMM)



**FDMM Working Group** develops a common set of core assessment criteria for FAIRness and a generic and expandable self-assessment model for measuring the maturity level of a dataset <https://www.rd-alliance.org/groups/fair-data-maturity-model-wg/activity/>

Guidelines and excel table <https://doi.org/10.15497/rda00050>

FDMM 41 indicators

<https://www.rd-alliance.org/>



	Findable	Accessible	Interoperable	Reusable	TOT
essential	7	8	0	5	20
important	0	3	7	4	14
useful	0	1	5	1	7
TOT	7	12	12	10	41

## Guidelines and excel table <https://doi.org/10.15497/rda00050>

ID	PRINCIPLE	INDICATOR_ID	INDICATORS	PRIORITY	METRIC	VIZ	SCORE/I	SCORE/E	SCORE/A	SCORE/U
1	F	F1	RDA-F1-01M	Metadata is identified by a persistent identifier	Essential	0 – not applicable	0	0	0	0
2		F1	RDA-F1-01D	Data is identified by a persistent identifier	Essential	0 – not applicable	0	0	0	0
3		F1	RDA-F1-02M	Metadata is identified by a globally unique identifier	Essential	0 – not applicable	0	0	0	0
4		F1	RDA-F1-02D	Data is identified by a globally unique identifier	Essential	0 – not applicable	0	0	0	0
5		F2	RDA-F2-01M	Rich metadata is provided to allow discovery	Essential	0 – not applicable	0	0	0	0
6		F3	RDA-F3-01M	Metadata includes the identifier for the data	Essential	0 – not applicable	0	0	0	0
7		F4	RDA-F4-01M	Metadata is offered in such a way that it can be harvested and indexed	Essential	0 – not applicable	0	0	0	0
8	A	A1	RDA-A1-01M	Metadata contains information to enable the user to get access to the data	Important	4 – fully implemented	4	1	0	1
9		A1	RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)	Essential	4 – fully implemented	4	1	1	0
10		A1	RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)	Essential	3 – in implementation phase	3	0	0	0
11		A1	RDA-A1-03M	Metadata identifier resolves to a metadata record	Essential	4 – fully implemented	4	1	1	0
12		A1	RDA-A1-03D	Data identifier resolves to a digital object	Essential	4 – fully implemented	4	1	1	0
13		A1	RDA-A1-04M	Metadata is accessed through standardised protocol	Essential	4 – fully implemented	4	1	1	0
14		A1	RDA-A1-04D	Data is accessible through standardised protocol	Essential	4 – fully implemented	4	1	1	0
15		A1	RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)	Important	4 – fully implemented	4	1	0	1
16		A1.1	RDA-A1.1-01M	Metadata is accessible through a free access protocol	Essential	4 – fully implemented	4	1	1	0
17		A1.1	RDA-A1.1-01D	Data is accessible through a free access protocol	Important	4 – fully implemented	4	1	0	1
18		A1.2	RDA-A1.2-01D	Data is accessible through an access protocol that supports authentication and authorisation	Useful	4 – fully implemented	4	1	0	0
19	A2	RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available	Essential	4 – fully implemented	4	1	1	0	
20	I	I1	RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	Important	4 – fully implemented	4	1	0	1
21		I1	RDA-I1-01D	Data uses knowledge representation expressed in standardised format	Important	4 – fully implemented	4	1	0	1
22		I1	RDA-I1-02M	Metadata uses machine-understandable knowledge representation	Important	4 – fully implemented	4	1	0	1
23		I1	RDA-I1-02D	Data uses machine-understandable knowledge representation	Important	4 – fully implemented	4	1	0	1
24		I2	RDA-I2-01M	Metadata uses FAIR-compliant vocabularies	Important	4 – fully implemented	4	1	0	1
25		I2	RDA-I2-01D	Data uses FAIR-compliant vocabularies	Useful	4 – fully implemented	4	1	0	1
26		I3	RDA-I3-01M	Metadata includes references to other metadata	Important	4 – fully implemented	4	1	0	1
27		I3	RDA-I3-01D	Data includes references to other data	Useful	3 – in implementation phase	3	0	0	0
28		I3	RDA-I3-02M	Metadata includes references to other data	Useful	4 – fully implemented	4	1	0	0
29		I3	RDA-I3-02D	Data includes qualified references to other data	Useful	3 – in implementation phase	3	0	0	0
30		I3	RDA-I3-03M	Metadata includes qualified references to other metadata	Important	4 – fully implemented	4	1	0	1
31	I3	RDA-I3-04M	Metadata include qualified references to other data	Useful	3 – in implementation phase	3	0	0	0	
32	R	R1	RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse	Essential	4 – fully implemented	4	1	1	0
33		R1.1	RDA-R1.1-01M	Metadata includes information about the licence under which the data can be reused	Essential	3 – in implementation phase	3	0	0	0
34		R1.1	RDA-R1.1-02M	Metadata refers to a standard reuse licence	Important	4 – fully implemented	4	1	0	1
35		R1.1	RDA-R1.1-03M	Metadata refers to a machine-understandable reuse licence	Important	3 – in implementation phase	3	0	0	0
36		R1.2	RDA-R1.2-01M	Metadata includes provenance information according to community-specific standards	Important	4 – fully implemented	4	1	0	1
37		R1.2	RDA-R1.2-02M	Metadata includes provenance information according to a cross-community language	Useful	3 – in implementation phase	3	0	0	0
38		R1.3	RDA-R1.3-01M	Metadata complies with a community standard	Essential	4 – fully implemented	4	1	1	0
39		R1.3	RDA-R1.3-01D	Data complies with a community standard	Essential	4 – fully implemented	4	1	1	0
40		R1.3	RDA-R1.3-02M	Metadata is expressed in compliance with a machine-understandable community standard	Essential	4 – fully implemented	4	1	1	0
41		R1.3	RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable community standard	Important	2 – under consideration or in planning phase	2	0	0	0

# RDA - FAIR Data Maturity Model

## Findability

Indicator	Description	Example
<b>F1-01M:</b> Metadata is identified by a persistent identifier	Commitment of the organisation that assigns and manages the identifier, thus ensuring findability over time and reducing broken links.	<i>MetaData identifier is assigned by a PID resolver (DOI Foundation), a Research organisation, or any organisation reliable, that won't disappear over time.</i> □ PID such as DOI
<b>F1-01D:</b> Data is identified by a persistent identifier	Same as above	<i>Data identifier is assigned by a PID resolver (DOI Foundation), a Research organisation, or any organisation reliable, that won't disappear over time.</i> □ PID such as DOI
<b>F1-02M:</b> Metadata is identified by a globally unique identifier	Description of how the identifier is assigned should make it clear that the mechanism cannot possibly assign the same identifier to different resources, or assign an identifier that has already been assigned via some other mechanism/organisation	<i>It is the case for PID resolver Identifier built with only a reference is not considered unique (e.g. 1235354RD)</i>
<b>F1-02D:</b> Data is identified by a globally unique identifier	Same as above	
<b>F2-01M:</b> Rich metadata is provided to allow discovery	Verify the presence of metadata, but also how much metadata is provided and how well the provided metadata supports discovery.	<i>Metadata standards such as ISO 19115 allows for rich description of geospatial datasets, and therefore provides rich discovery metadata.</i>
<b>F3-01M:</b> Metadata includes the identifier of the data	Reference/Identifier of the data is included in the metadata element that is specified for that purpose in the metadata standard used. It ensures access to the data.	<i>For example, in ISO 19115, verify if the metadata field «MD_Distribution/TransferOption » contains the identifier of/the reference to the dataset.</i>
<b>F4-01M:</b> Metadata is offered in such a way that it can be harvested and indexed	MD is offered in such a way that it can be indexed/harvested, by a general search engine or in a more restricted index, such as an institutional repository or a discipline-specific portal.	<i>Portal allowing metadata harvesting through OAI-PMH, API REST using JSONAPI, or Catalog Service for the Web (CSW)...</i>

# RDA - FAIR Data Maturity Model

## Accessibility

Indicator	Description	Example
<b>A1-01M:</b> Metadata contains information to enable the user to get access to the data	MD should describe access conditions (open, restricted, closed) and possible requirements & actions to be taken to access data (e.g. sign-posting/registration, access request through forms, etc.).	<i>Access level and conditions described in metadata (e.g. in « Access restriction » for ISO 19115)</i>
<b>A1-02M:</b> Metadata can be accessed manually (i.e. with human intervention)	Any human interactions needed to access metadata. May be important when MD itself contains sensitive information.	<i>Information in the landing page or documentation about e-mailing author, etc.</i>
<b>A1-02D:</b> Data can be accessed manually (i.e. with human intervention)	Any human interactions needed to access the digital object : clicking on a link on a landing page, sending an e-mail to the data owner, ...	<i>Same</i>
<b>A1-03M:</b> Metadata identifier resolves to a metadata record	The MD identifier is associated with a resolution service that enables access to the metadata record.	<i>Metadata can be accessed with its identifier</i>
<b>A1-03D:</b> Data identifier resolves to a digital object	Data identifier should be associated with a formally defined retrieval/resolution mechanism that enables access to the digital object or provides access instructions	<i>Data identifier allows access to the dataset file (HTTP GET)</i>
<b>A1-04M:</b> Metadata is accessed through standardised protocol	the protocol through which the metadata is accessed is defined in a standard.	<i>MD are accessible using HTTP, FTP, etc.</i>
<b>A1-04D:</b> Data is accessible through standardised protocol	the protocol through which the digital object is accessed is defined in a standard.	<i>Data is accessible using HTTP, FTP, DAP, etc.</i>
<b>A1-05D:</b> Data can be accessed automatically (i.e. by a computer programm)	Evaluate the way machines interacts & grant access to data. Resolving the link to the data (e.g. PID) and verifying the correct data file is reached.	<i>If data identifier is an HTTP URI, a HTTP GET method should grants access.</i>
<b>A1.1-01M:</b> Metadata is accessible through a free access protocol	The protocol to access metadata can be freely used (free of charges)	<i>Free access protocols : HTTP, FTP</i>
<b>A1.1-01D:</b> Data is accessible through a free access protocol	The protocol to access data can be freely used (free of charges)	<i>Free access protocols : HTTP, FTP, etc.</i>
<b>A1.2-01D:</b> Data is accessible through an access protocol that supports authentication and authorisation	If the data indicates a degree of additional protection, then the access protocol must support authentication and authorisation of people and/or machines	<i>AAI process (e.g. EduGAIN, MarineID) HTTP protocol support AAI</i>
<b>A2-01M:</b> Metadata is guaranteed to remain available after data is no longer available	MD available even after the data has been deleted.	<i>Use of repository/catalog that have memorial pages and ensure continuous access to MD</i>

# RDA - FAIR Data Maturity Model

## Interoperability

Indicator	Description	Example
<b>I1-01M:</b> Metadata uses knowledge representation expressed in standardised format	An appropriate standard is used to express knowledge, for example, controlled vocabularies for subject classifications.	<i>MD accessible in RDF, OWL</i>
<b>I1-01D:</b> Data uses knowledge representation expressed in standardised format	An appropriate standard is used to express knowledge, in particular the data model and format.	<i>Data described with NetCDF CF, or else Check if available in <a href="#">RDA-endorsed FAIRsharing registry</a></i>
<b>I1-02M:</b> Metadata uses machine-understandable knowledge representation	metadata should be readable and thus interoperable for machines without any requirements (translators or mappings)	<i>MD accessible in RDF, JSON-LD, SKOS through a SparQL Endpoint. Check if available in <a href="#">RDA-endorsed FAIRsharing registry</a></i>
<b>I1-02D:</b> Data uses machine-understandable knowledge representation	Same but for data	<i>Data described with RDF, JSON-LD, NetCDF CF Check if available in <a href="#">RDA-endorsed FAIRsharing registry</a></i>
<b>I2-01M:</b> Metadata uses FAIR-compliant vocabularies	The vocabulary used for the MD is conform to the FAIR principles, and at least is documented and resolvable using globally unique and persistent identifiers.	<i>Standard &amp; community vocabulary is used, with PID &amp; MD for each terms: e.g. NERC Vocabulary SDN P01 collections</i>
<b>I2-01D:</b> Data uses FAIR-compliant vocabularies	Same but for data	<i>E.g. NetCDF Climate &amp; Forecast convention, etc.</i>
<b>I3-01M:</b> Metadata includes references to other metadata	links to information about organisations, people, places, projects or time periods related to the data described	<i>E.g. ORCID for persons, ROR for organisations, Geonames for places, etc.</i>
<b>I3-01D:</b> Data includes references to other data	Link to other resources in Excel spreadsheet or RDF-based data	<i>E.g. DOI or identifier to source data or related datasets</i>
<b>I3-02M:</b> Metadata includes references to other data	Links to previous or related data providing additional context	<i>E.g. DOI or identifier to related datasets in dedicated MD field</i>
<b>I3-02D:</b> Data includes qualified references to other data	references with specification of the relationship role that the related resource has with the data	<i>E.g. skos:broader in skos language</i>
<b>I3-03M:</b> Metadata includes qualified references to other metadata	Same with metadata to metadata	<i>E.g. person Y is the author of dataset X.</i>
<b>I3-04M:</b> Metadata include qualified references to other data	Same with metadata to data	<i>E.g. dataset X is derived from dataset Y</i>

# RDA - FAIR Data Maturity Model

## Reusability

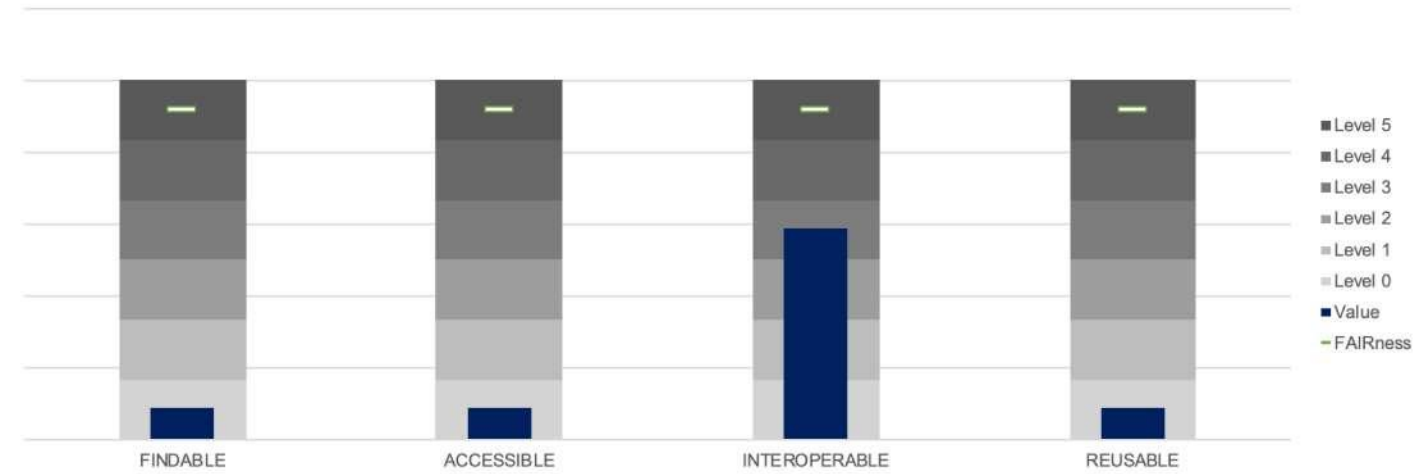
Indicator	Description	Examples
<b>R1-01M:</b> Plurality of accurate and relevant attributes are provided to allow reuse	quantity but also quality of MD provided to enhance data reusability.	<i>Use of MD standards such as ISO 19115 allows for rich description (variables, quality processes, provenance info, dataset format, type, project, etc.)</i>
<b>R1.1-01M:</b> Metadata includes information about the licence under which the data can be reused	information that is provided in the MD related to the conditions (e.g. obligations, restrictions) under which data can be reused.	<i>MD provides info about licence (e.g. in ISO 19115 « Legal Constraints / AccessConstraints/ AccessRestrictions »)</i>
<b>R1.1-02M:</b> Metadata refers to a standard reuse licence	conditions of reuse are described using a standard licence, rather than a locally defined licence.	<i>E.g. Creative Commons Licences, Open Data Commons</i>
<b>R1.1-03M:</b> Metadata refers to a machine-understandable reuse licence	Rather than being a human-readable text, the licence should be expressed in such a way that it can be processed by machines, without human intervention	<i>E.g. licence expressed in RDF – CC, serialisations of Open Data Rights Language</i>
<b>R1.2-01M:</b> Metadata includes provenance information according to community-specific standards	MD includes information about the provenance of the data, i.e. information about the origin, history or workflow that generated the data, in a way that is compliant with the standards that are used in the community	<i>E.g. provenance info described in ISO 19115 « Lineage » or other community standards (Check references in <a href="https://fairsharing.org/">https://fairsharing.org/</a>)</i>
<b>R1.2-02M:</b> Metadata includes provenance information according to a cross-community language	metadata provides provenance information according to a cross-domain language.	<i>E.g. use of <a href="#">PROV-Q</a> to describe provenance</i>
<b>R1.3-01M:</b> Metadata complies with a community standard	MD follows a standard used by the community	<i>E.g. ISO 19115, GEO-DCAT AP, EML, ...</i>
<b>R1.3-01D:</b> Data complies with a community standard	Data follows a standard used by the community	<i>E.g. NetCDF CF, ASCII TSV, ...</i>
<b>R1.3-02M:</b> Metadata is expressed in compliance with a machine-understandable community standard	MD follows a community standard that has a machine-understandable expression	<i>E.g. XML, RDF</i>
<b>R1.3-02D:</b> Data is expressed in compliance with a machine-understandable community standard	data follows a community standard that has a machine-understandable expression	<i>E.g. NetCDF CF,</i>

# FDMM outcome

## FAIRNESS LEVELS PER AREA

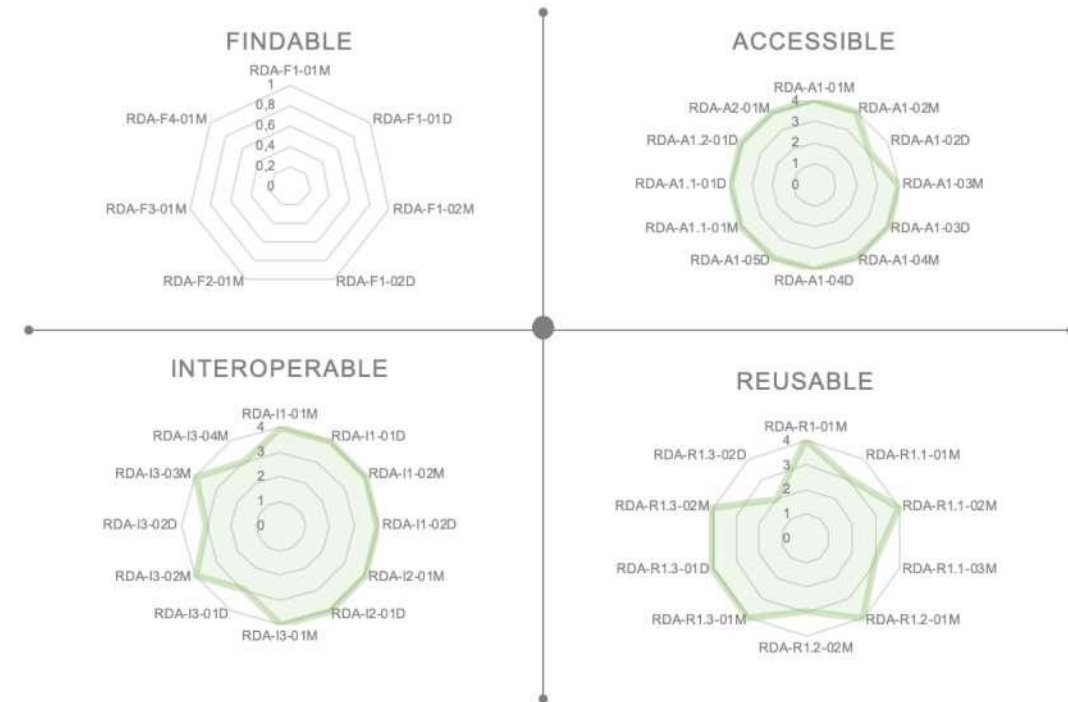
Level	Description
Level 0	Not FAIR
Level 1	FAIR essential criteria only
Level 2	FAIR essential criteria + 50% of important criteria
Level 3	FAIR essential criteria + 100% of important criteria
Level 4	FAIR essential criteria + 100% of important criteria + 50% of useful criteria
Level 5	FAIR essential criteria + 100% of important criteria + 100% of useful criteria

What level describes your digital object?



## FAIRNESS PROGRESS PER INDICATOR

Maturity level per indicator (per FAIR area)	Description
0	not applicable
1	not being considered this yet
2	under consideration or in planning phase
3	in implementation phase
4	fully implemented



Self assessment is important to identify weak points and define FAIRification strategies

# References

Simoncelli, S., Manzella, G.M.R., Storto, A., Pisano, A., Lipizer, M., Barth, A., Myroshnychenko, V., Boyer, T., Troupin, C., Coatanoan, C., Pititto, A., Schlitzer, R., Schaap, D.M.A., Diggs, S., 2022. A collaborative framework among data producers, managers, and users. <https://doi.org/10.1016/B978-0-12-823427-3.00001-3> In: Manzella, G., Novellino, A. (Eds.), **Ocean Science Data: Collection, Management, Networking and Services**. Elsevier, pp. 197–280. ISBN: 9780128234273

UNESCO-IOC (2023). Ocean Decade Data & Information Strategy. Paris, UNESCO. (The Ocean Decade Series, 45)

Calewaert et al. (2024). Ocean Decade Vision 2030 White Papers – Challenge 8: Create a Digital Representation of the Ocean. Paris, UNESCO-IOC. (The Ocean Decade Series, 51.8.). <https://doi.org/10.25607/bxhy-ra59>

Miloslavich et al. (2024). Ocean Decade Vision 2030 White Papers – Challenge 7: Sustainably Expand the Global Ocean Observing System. Paris, UNESCO-IOC. (The Ocean Decade Series, 51.7.). <https://doi.org/10.25607/brxb-kr45>

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3:60018 <https://doi.org/10.1038/sdata.2016.18>

[https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/PDF/EUInSituMarine\\_EMODnet\\_CMEMS\\_FINAL.pdf](https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/PDF/EUInSituMarine_EMODnet_CMEMS_FINAL.pdf)

European Commission: Directorate-General for Research and Innovation, The digital twin ocean – An interactive replica of the ocean for better decision-making, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2777/343496>

Schaap, D., Le, H. M., Iona, S., Diez, S., Partescano, E., Sagen, H., Buga, L., Schepers, L., Hedblom, M., O'Grady, E., Van Iseghem, S., Krijger, T., & Giorgetti, A. (2024). AQUARIUS Data Management Plan (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.13642869>

Locati & Cacciola (2024) Research Data Management <https://istituto.ingv.it/ufficio-gestione-dati>



# THANKS!

**IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System**  
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-  
Mission 4 "Education and Research" - Component 2: "From research to business" - Investment  
3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures"

