



## Marine Data Centre Management Plan: management needs, service requirements, governance for the distributed and federated system



<b>Deliverable number:</b>	D5.1
<b>Work package:</b>	WP5 – Marine Domain
<b>Intermediate Objective:</b>	IO5.2 – (B3): (a) Marine Data Centre Management plan and governance for the distributed and federated system; (b) Implementation plans for technological assets acquisitions at all OU; (c) first personnel hired; (d) WP progress meeting.
<b>Deliverable type:</b>	<input checked="" type="checkbox"/> Document, report
	<input type="checkbox"/> Websites, patent filings, videos, etc.
	<input type="checkbox"/> Other: please specify .....
<b>Dissemination level:</b>	<input checked="" type="checkbox"/> Public
	<input type="checkbox"/> Restricted
<b>Estimated delivery (bimester):</b>	B3
<b>Actual delivery date:</b>	30/04/2023
<b>Author(s) (Partner-OU):</b>	Alessandra Giorgetti (OGS-OCE), Chiara Altobelli (OGS-OCE), Maurizio Azzaro (CNR-ISP-ME), Debora Bellafiore (CNR- ISMAR), Caterina Bergami (CNR-ISMAR), Roberto Bozzano (CNR-ISMAR), Vanessa Cardin (OGS-OCE), Franco Coren (OGS-CGN), Francesca De Pascalis (CNR-ISMAR), Davide Embriaco (INGV), Antonella Gallo (OGS-OCE), Michele Giani (OGS-OCE), Giordano Giorgi (ISPRA), Marcello Magaldi (CNR ISMAR), Elena Mauri (OGS-OCE), Stefano Miserocchi (CNR-ISMAR), Giulio Notarstefano (OGS-OCE), Giorgio Maria Riccobene (INFN-LNS), Katrin Schroeder (CNR-ISMAR), Simona Simoncelli (INGV-BO)
<b>Reviewed by:</b>	ITINERIS Executive Board and WP5 working groups Marine RIs and IOOS Design
<b>Note:</b>	

IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System - CUP B53C22002150006 (D.D. n. 130/2022)

Funded by EU - Next Generation EU

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”

## Table of contents

<i>EXECUTIVE SUMMARY</i> .....	4
<i>LIST OF ABBREVIATIONS AND ACRONYMS</i> .....	5
<i>INTRODUCTION</i> .....	6
<i>THE ITALIAN LANDSCAPE</i> .....	8
<i>STUDY OF MANAGEMENT NEEDS AND SERVICE REQUIREMENTS</i> .....	10
<i>RESULTS</i> .....	11
<i>DATA MANAGEMENT PLAN</i> .....	18
<b>The Marine Data Store</b> .....	<b>19</b>
<b>Four solutions that will have to coexist and converge</b> .....	<b>20</b>
<b>ITINERIS Data and Metadata catalogues</b> .....	<b>23</b>
<i>ITINERIS MARINE DATA POLICY</i> .....	24
<b>Data Use Policy</b> .....	<b>25</b>
<b>User licence</b> .....	<b>25</b>
<b>Access rights</b> .....	<b>25</b>
<i>CONCLUSIONS</i> .....	26
<i>REFERENCES</i> .....	27
<i>ANNEX - QUESTIONNAIRE FOR MARINE RESEARCH INFRASTRUCTURES</i> .....	28

## EXECUTIVE SUMMARY

This document details the Marine Data Centre Management Plan, outlining the management needs, service requirements, and governance for designing and implementing the distributed and federated Marine Data Centre. It should be considered as a “living document” that will evolve through subsequent updates.

**The objective of the Marine Data Centre is** to provide open and shared access to the information and data collected by the Italian Institutions participating to the Marine Research Infrastructures (RIs) and to achieve interoperability with an incremental approach based on consolidated experiences in the European context. The **Marine Data Centre** is designed as a federated system with a single-entry point that will give access to the data collected by the Marine RIs and reported in the ITINERIS catalogue for data and metadata. The Marine Data Centre will be developed based on data themes, with data catalogues and datasets from all participating RIs. Data formats, vocabularies and services will be established in collaboration with WP5.3 and 5.4. Metadata of facilities and data will be collected from all the marine RIs; information will populate the ITINERIS catalogue for data and metadata developed by WP5.3 in compliance to FAIR principles and procedure from WP2.

One important component of the Marine Data Centre will be the **Marine Data Store** in Naples. This marine central facility will store all the RI CNR data and will integrate extra-RIs specific data, selected based on the readiness and complementarity to existing RIs.

The Marine Data Centre will be directly connected to the **ITINERIS HUB**, the main web platform where all users will have access to data and services from all environmental RIs, with a proper access management system and a complete catalogue of data and services. ITINERIS will not develop a new Data Centre, but rather harmonize the already existing ones, which will be reinforced by the project activities. Most of the data-related activities will be devoted to reach the most appropriate level of FAIRness for all the involved RIs. Most importantly, the Marine Data Centre will increase the visibility and FAIRness (Findability, Accessibility, Interoperability, and Reusability) of the Italian contribution to international RIs and support the participation of Italian scientists in pan-European initiatives such as ENVRI, ENVRI-FAIR and EOSC.

As part of the ITINERIS HUB, the Marine Data Centre is an essential element of the Italian Integrated Ocean Observing System (IOOS), integrating and harmonising the observations of the Marine Domain RIs at the Italian level. It also enhances their visibility as well as data FAIRness according to the principles of **Open Science**.

The first two sections (Introduction and the Italian Landscape) summarise the status of policies and strategies to implement the principles of FAIR data at global, European and Italian levels. The section 'Study of management needs and service requirements' explains the survey carried out to design the Marine Data Centre and coordinate its development, starting from the management needs and service requirements at RI level. It also details the main reference documents that were considered to define the Marine Data Centre management plan. In the 'Results' section, the management and service need and gaps at RI level that emerged from the survey are presented. In the section ‘.....’ different scenarios are identified to properly set-up the management plan for ITINERIS data depending on the maturity of the Marine RI and of the Italian research institutes. A proposal for the ITINERIS marine data policy is also included, derived from the data policy developed by the Italian Oceanographic Commission. The conclusions specify the next steps for the establishment of the ITINERIS Marine Data Centre, including a first proposal for the identification of responsibilities and the governance model to be aligned with the ITINERIS governance.

## LIST OF ABBREVIATIONS AND ACRONYMS

APIs	Application Programming Interfaces
CNR	Consiglio Nazionale delle Ricerche
CNR-ISMAR	Consiglio Nazionale delle Ricerche - Istituto di Scienze MARine
CNR-ISP-ME	Consiglio Nazionale delle Ricerche - Istituto di Scienze Polari di Messina
COI	Italian Oceanographic Commission
CTD	Device measuring the electrical conductivity (C), temperature (T), and pressure of seawater to determine depth (D)
DANUBIUS-RI	The International Centre for Advanced Studies on River –Delta – Sea Systems
DMP	Data Management Plan
eLTER	The integrated European Long-Term Ecosystem, critical zone and socio-ecological Research Infrastructure
EMODnet	The European Marine Observation and Data Network
EMSO	The European Multidisciplinary Seafloor and water column Observatory
EOOS	European Oceanographic Observing System
EOSC	European Open Science Cloud
ERIC	European Research Infrastructure Consortium
ESFRI	European Strategic Forum on Research Infrastructures
Euro-Argo	The European contribution to the international Argo Programme for observing the ocean
EUROFLEETS-RI	An Alliance of European Research Fleets
FAIR	Findable, Accessible, Interoperable, Reusable
GeoSciences	Rete italiana per le geoscienze
GOOS	Global Oceanographic Observing System
ICOS	The Integrated Carbon Observation System
INFN	Istituto Nazionale di Fisica Nucleare
INGV	Istituto Nazionale di Geofisica e Vulcanologia
IOC	The Intergovernmental Oceanographic Commission
IODC	The Italian Oceanographic Data Committee
IODE	Oceanographic Information and Data Exchange
IOOS	Italian Integrated Ocean Observing System
ISPRA	Istituto Superiore per la Protezione
ITINERIS	The Italian Integrated Environmental Research Infrastructures System project
JERICO-RI	Joint pan-European Research Infrastructure for Coastal Observations
LNS	Laboratori Nazionali del Sud
N/R Laura Bassi	Nave da Ricerca Polare Laura Bassi
OGS-CGN	National Institute of Oceanography and Applied Geophysics - OGS, Centre for Management of Maritime Infrastructures
OGS-OCE	National Institute of Oceanography and Applied Geophysics - OGS, Section of Oceanography
PNIR	Piano Nazionale Infrastrutture di Ricerca
RI	Research Infrastructures
SeaDataNet	pan-European infrastructure for ocean & marine data management

SIOS	Svalbard integrated arctic earth observing system
SPARQL	SPARQL Protocol and RDF Query Language
UNESCO	The United Nations Educational, Scientific and Cultural Organization
VREs	The Virtual Research Environments
WP	Work Package

## INTRODUCTION

Although Open Science originated several decades ago and has expressed principles that, in theory, should be the basis of shared knowledge and scientific research, in practice, Open Science represents a quite recent revolution in the *modus operandi*. It is an approach for transforming scientific practice to adapt to the changes, challenges, opportunities, and risks of the digital era and to increase the societal impact of science. Making data as open as possible and as close as needed is the Open Science motto, which bases on the FAIR data management principles [Wilkinson 2016]. Data FAIRness means that data need to be Findable, Accessible, Interoperable and Re-usable in order to accelerate data legitimation and mitigate the possibility to apply researchers' bias to knowledge, reducing the dissemination of unreliable information [Burgelman 2019]. Legitimated knowledge must drive towards a thoughtful and far-sighted environmental policy. To this end, the advancement and networking of open web-based tools (e.g., FAIR data discovery and access services), which are the essential prerequisites, and the Open Science engine, are necessary to put its principles into practice [EC 2018]. Open Science has progressively been recognised as a pillar for the United Nations and European policies, evolving to pursue its principles and face its challenges.

The Intergovernmental Oceanographic Commission (IOC) of UNESCO started the **International Oceanographic Data and Information Exchange (IODE) programme** in 1961 to enhance marine research, exploitation, and development by i) facilitating the exchange of oceanographic data and information between participating Member States, and ii) meeting the needs of users for data and information products. The IODE has always stressed the importance of coordinating data and information management at the national level [UNESCO 2008]. Originally, oceanographic data were mostly managed in a central national facility, i.e., a National Oceanographic Data Centre (NODC). The data usually came from research cruises or research projects and were submitted to the data centres by researchers with a time lag, i.e., after days to years of collection and after publication. Since the development of operational oceanography and the establishment of the Global Ocean Observing System (GOOS), the number of data types and the amount of data have greatly increased, and the NODCs have had to expand towards the new data streams. Data need to be shared, validated, and used in a proper manner to be helpful. To this purpose, IODE states that an efficient coordination at the national level is essential for marine monitoring and research and that a data management plan must be developed as part of each initiative.

To achieve its objectives, the IODE recommends that 1) a National Coordination Committee for Oceanographic Data and Information be established, composed of all stakeholders involved in data collection and management, and 2) a national facility be assigned the role of NODC within the UNESCO IOC/IODE network of Data Centres.

With regard to UNESCO's engagement in data sharing, the following can also be noted 1) the 1999 UNESCO/ICSU Declaration on Science and the Use of Scientific Knowledge and the Science Agenda – Framework for Action, 2) the 2002 Budapest Open Access Initiative, 3) the 2003 Bethesda Statement on Open Access Publishing and the 2003 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, and 4) the IODE guide for establishing a NODC [UNESCO 2022].

More specifically, on 23 November 2021 at the 41st session of the UNESCO General Conference a Recommendation on Open Science was adopted. In this document, UNESCO acknowledged 1) the

importance of science, technology and innovation (STI) to respond to environmental, social and economic challenges, and 2) the opportunities and potential provided by the Open Science cultural revolution and the expansion of information/communication technologies for: i) reducing the existing inequalities in STI; ii) accelerating progress towards the implementation of the 2030 Agenda and the achievement of the Sustainable Development Goals (SDGs) and beyond. It also recognised the existence of international legal frameworks on intellectual property rights of scientists, and practices and open licence schemes that allow free access, re-use and distribution of work under specific conditions, provided that the creator is appropriately credited.

The IOC Oceanographic Data Exchange policy, which was adopted through the Resolution IOC-XXII-6 at the 22nd Session of the IOC Assembly in 2003 and slightly revised by the 30th Session of the IOC Assembly in 2019, was revised in 2022 and the new policy has been submitted for adoption to the IOC Assembly in 2023. The policy promotes free and open access to data, metadata, and products, and aims to maximise the amount of data exchanged without infringing the rights of data originators.

At the EU level, in 2003 the Directive 2003/4/EC on public access to environmental information was adopted to guarantee the right of access to environmental information held by or for public authorities and to ensure that environmental information is progressively made available and disseminated to the public. In 2019, the EU adopted Directive (EU) 2019/1024 to promote the use of open data and stimulate innovation in products and services. At the same time, the EU is funding the development of free, multidisciplinary environments called Open Research Europe (ORE) and European Open Science Cloud (EOSC), including the marine-related initiatives Blue-Cloud. EOSC enables the publishing, finding, and re-use of data, tools, and services for research, innovation, and educational purposes [3], while ORE is an open access platform for the publication of research stemming from Horizon 2020 and Horizon Europe. Concerning open data infrastructures, it is worth mentioning the following **Blue Data Infrastructures (BDI)**: SeaDataNet, EMODnet, and Copernicus.

Since 2006, SeaDataNet has evolved into a robust and state-of-the-art pan-European infrastructure providing open access to up-to-date and high-quality marine metadata, data, and data products. It is based on a pan-European network of National Oceanographic Data Centres (NODCs) and has been adopted by several other EU RTD projects and EMODnet thematic lots.

Since 2009, the philosophy of the European Marine Observation and Data Network (EMODnet) has been "collect once and use many times". Supported by the EU Integrated Maritime Policy, the organisations that are building EMODnet work together to observe the sea, process the data according to international standards, and make this information freely available to all users of marine data - including policy makers, scientists, the private sector and the public - as interoperable data layers and data products. EMODnet provides access to European marine data across seven discipline-based themes: Bathymetry, Biology, Chemistry, Geology, Human Activities, Physics, and Seabed Habitats.

The EMODnet Data Ingestion portal (<https://www.emodnet-ingestion.eu/>) facilitates the processing and publishing of marine datasets as open data thanks to the underlying network of more than 160 organisations that together work on assembling, harmonising, and making marine data, products, and metadata more available to public and private users. It aims to streamline the data ingestion process so that data holders from public and private sectors who are not yet connected to the existing marine data management infrastructures can easily release their data for safekeeping and subsequent distribution through EMODnet. It enriches the data management services for all types of users.

Copernicus is the Earth observation component of the European Union's Space programme. Managed by the European Commission it offers free and openly accessible information services that are based on satellite and in situ observations. In particular, the Copernicus Marine Environment Monitoring Service provides regular and systematic reference information on the physical and

biogeochemical state, variability, and dynamics of the ocean and marine ecosystems for the global ocean and the European regional seas.

## THE ITALIAN LANDSCAPE

To fulfil the functions of the National Coordination Body in formulating proposals for the effective Italian participation in IOC activities and to coordinate the position of the Italian oceanographic scientific community, the **Italian Oceanographic Commission (COI)** was established in 2008. To involve experts per topic, the COI set up working groups in correspondence to IOC programs like GOOS, IODE, Ocean Literacy, etc. The IODE working group of the COI aims to strengthen the coordination in the management of data and information at the national level. Since the beginning, it has been very active, and a number of meetings were organised to answer the IODE programme's requirements to move from a centralised national structure to a distributed data management system and to create the National Oceanographic Data and Information Coordination Committee at the Italian level. The COI has also appointed the IODE National Coordinator for Oceanographic Data Management.

Italy has a long-standing experience in marine data management, participates in the main pan-European data infrastructures, and has been represented in the Intergovernmental Oceanographic Data Exchange program since the 1980s. This effort is also testified by the development of a dedicated marine data management service, hosted by the OGS and called the **National Oceanographic Data Centre for Italy (NODC)**, which holds the expertise of the international framework and can constitute the basis for further developments. For the fulfilment of the IODE programme recommendations to establish a National Oceanographic Data and Information Coordination Committee, the COI through its IODE working group has elaborated an agreement between all Italian research institutes operating in the sea for the establishment of the **Italian Oceanographic Data Committee (IODC)**. IODC's scope is to develop a distributed and federated system of national databases with a single access portal hosted on the COI website and overseen by the IODC. The agreement sets out the obligations to be assumed by the parties and the data policy that defines the ownership and use of the data.

In addition, the document explains the scheme of cooperation between the institutions and proposes possible working methods. The proposed national structure is shown in Figure 1, with the new elements to be developed in support of the structures highlighted in blue and the existing ones in green.

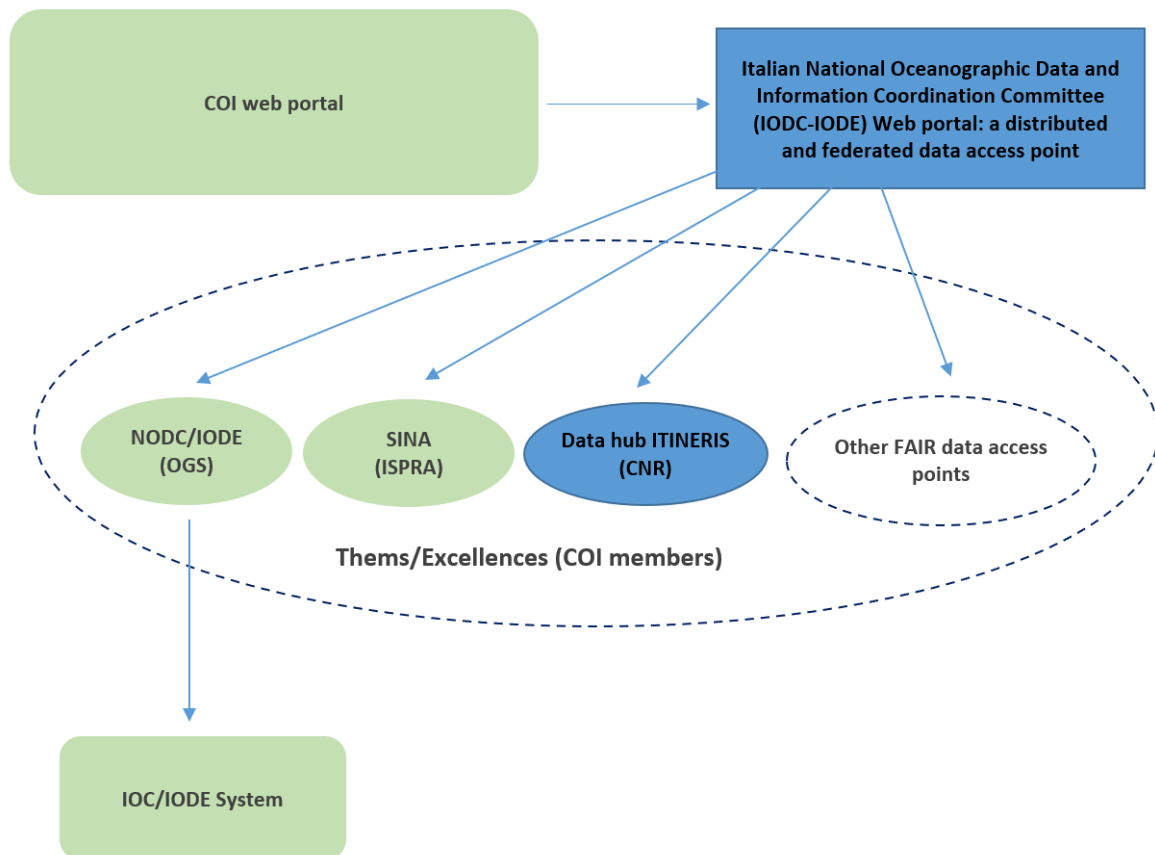


Fig. 1: Schematic of the distributed and federated portal of the IODC-IODE proposed by COI

ITINERIS will adopt the approach proposed by COI.

The diagram highlights the importance of the **ITINERIS Marine Data Centre** as a reference point where all the Italian marine monitoring systems and modelling results are displayed, and all the marine data can be visible and made easily accessible. Data collected by the Marine Domain RIs with different sensors and observation systems, fixed and mobile platforms, ships and gliders will be harmonised, integrated, and made visible into IOOS. The Marine Data Centre, outlined by the COI as a distributed system, will be implemented with the contribution of all ITINERIS WP5 UOs. It will be interfaced with existing RIs' data centres, where planned and already developed, as well as thematic or European data centres, where existing, to provide a single national access point. In detail, the Italian contribution within the following RIs will be considered and connected to the Marine Data Centre:

- DANUBIUS-RI - the International Centre for Advanced Studies on River –Delta – Sea Systems, led by CNR, listed with high priority in the PNIR,
- eLTER-RI - the integrated European Long-Term Ecosystem, critical zone, and socio-ecological Research Infrastructure (eLTER-RI), led by CNR, listed with high priority in the PNIR,
- EMSO-ERIC - the European Multidisciplinary Seafloor and water column Observatory, led by INGV, listed with high priority in the PNIR,
- Euro-Argo ERIC - the European contribution to the international Argo Programme for observing the ocean, led by OGS, with high priority ranking in the PNIR,
- ICOS-ERIC - the Integrated Carbon Observation System, led by CNR, listed with high priority in the PNIR,

- LNS - Laboratori Nazionali del Sud, led by INFN, listed with high priority in the PNIR,
- SIOS - Svalbard integrated arctic earth observing system, led by CNR, listed with high priority in the PNIR,
- EUROFLEETS-RI - An Alliance of European Research Fleets, led by CNR, listed with medium priority in the PNIR,
- GeoSciences - Rete italiana per le geoscienze, led by ISPRA, listed with medium priority in the PNIR,
- JERICO-RI - Joint pan-European Research Infrastructure for Coastal Observations, led by CNR, listed with medium priority in the PNIR,
- N/R Laura Bassi - Nave da Ricerca Polare Laura Bassi, led by OGS-CGN, listed with medium priority in the PNIR.

The establishment of the ITINERIS Marine Data Centre will i) strongly support the Italian role in several RIs, ii) valorise the national investment, and iii) contribute to building a coordinated system of the national nodes for three critical themes: climate, operational services, and marine ecosystem health. It will show how data integration from RIs can improve monitoring of coastal waters (including river discharge) and the 4D view of the ocean state, thus providing essential component for building the digital twin of the ocean. It will contribute to the creation of a coherent and harmonised flow of data, information, and knowledge across all the environmental RIs that will be ready for use in the system of Virtual Research Environments (VREs) developed by ITINERIS, and the enhancement of data FAIRness.

## STUDY OF MANAGEMENT NEEDS AND SERVICE REQUIREMENTS

To design the Marine Data Centre and coordinate its development, starting from the management needs and the service requirements, a survey was conducted by OGS-OCE with the help of and in agreement with the leaders of the working group Marine RIs (INGV and OGS-OCE) and the working group IOOS design (CNR-ISMAR and OGS-OCE). The survey aimed to collect information on the state of the art with regards to data collection and access by all marine RIs. The survey was realised through an online questionnaire entitled “**Questionnaire for Marine Research Infrastructures**” and was described as a “Questionnaire to collect information on marine RIs, variables observed, and protocols for managing and distributing data and metadata within ITINERIS WP5”. It was produced by using the Microsoft app “forms” and sent to all marine RIs on 06/03/2023. The questionnaire is briefly described hereafter and is available in the Annex to this document. The content of the Data Management Plan (DMP) template (version 1.0 of 05 May 2021) provided under the Horizon Europe Model Grant Agreement was used to structure the questionnaire. This choice gives acknowledgement to the great importance Horizon Europe attaches to this mandatory task, which requires the preparation and regular updating of a DMP by the beneficiaries of EU grants.

The questionnaire included 28 questions and was divided into 3 sections as follows: General information, Data summary, and FAIR data. The section ‘General information’ included 4 questions regarding the name of the RI, quantity of facilities, and their possible link with other RIs. The section ‘Data summary’ included 3 questions to obtain information about current and potential variables investigated by the RIs and their acquisition mode. The section ‘FAIR data’ included 21 questions and was structured for topics as follows: findability, accessibility, interoperability, and re-usability. In this section, respondents were asked about the existence and TRL of 1) a data catalogue and 2) data access services also via APIs. It was also asked whether the RI uses a trusted repository. Information about the existence of a data policy and the use of format, vocabularies, and standards to guarantee interoperability was also collected. The last part of this section included questions about data quality assurance, licences for publishing data, and persistent identifiers. The answers were reviewed to remove test answers and duplicates and to harmonise those validated. Although one

response was expected for each RI, in a few cases more than one person provided information for the same RI, e.g. for different facilities.

To define the governance for the distributed and federated system, the content of the **COI agreement** for the establishment of the IODC has been considered. Further considerations are presented in the last concluding section.

Besides, the **IOC Manuals and Guides n. 5** [UNESCO 2022], where the mission and the steps to establishing a data centre as well as its tasks and responsibilities are indicated, was considered.

Overall, the standards and practices implemented in the **European framework** will guide the technical development, which must guarantee data interoperability: the key to distributed data management system success. This is achieved in SeaDataNet by using common vocabularies, adopting the ISO 19115 metadata standard for all metadata directories, using harmonised Data Transport Formats to deliver data sets and using common quality control protocols and flag scales.

## RESULTS

As described in the previous section, a **survey was conducted to determine the state of the art** in data management of each RI. The results of the survey are shown below. The questionnaire was completed by 14 respondents from CNR-ISMAR, CNR-ISP-ME, INGV, INFN, ISPRA, and OGS who provided information on all RIs participating in ITINERIS, namely: DANUBIUS-RI, eLTER-RI, EMSO, EUROFLEETS, Euro-Argo, Geosciences, ICOS, JERICO, Laura Bassi, LNS, SIOS. The following results are elaborated based on the latest questionnaires filled in on 27 April 2023.

The number of facilities varies considerably between RIs, from a few units in most cases to several dozen in Euro-Argo and JERICO. Overall, 64% (7 out of 11) of facilities are shared with other RIs (Figure 2).



Fig. 2: Facilities shared with other RIs (9 respondents of 7 RIs)

The survey showed that the variables that the RIs acquire vary widely (Figure 3), with most RIs collecting data on temperature, salinity (90%), oxygen (81%) and currents (73%), followed by data of particulate matter, ocean sound, nutrients (54%) and wind (45%) (Fig. 3). The potential lack of a standard for ocean sound is worrying, so action is needed to close the knowledge gap. The fact that only half of RIs collect data on known EOVS and ECV variables such as nutrients is surprising, while the presence of sites collecting data on emerging variables such as ocean sound and marine litter is encouraging. The results showed that 59% of RIs provide data in real-time mode.

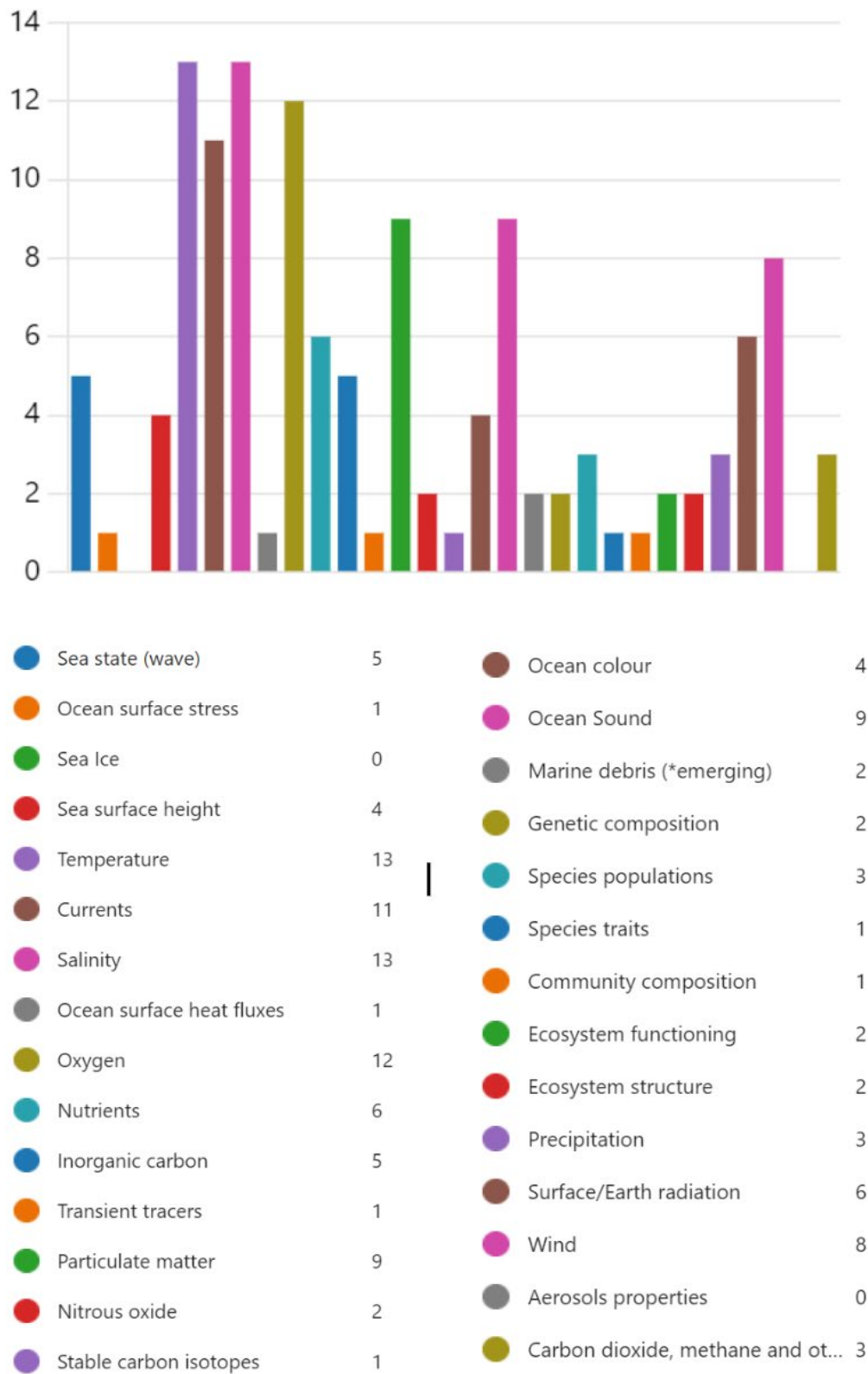


Fig. 3: Variables that the RIs acquire data on

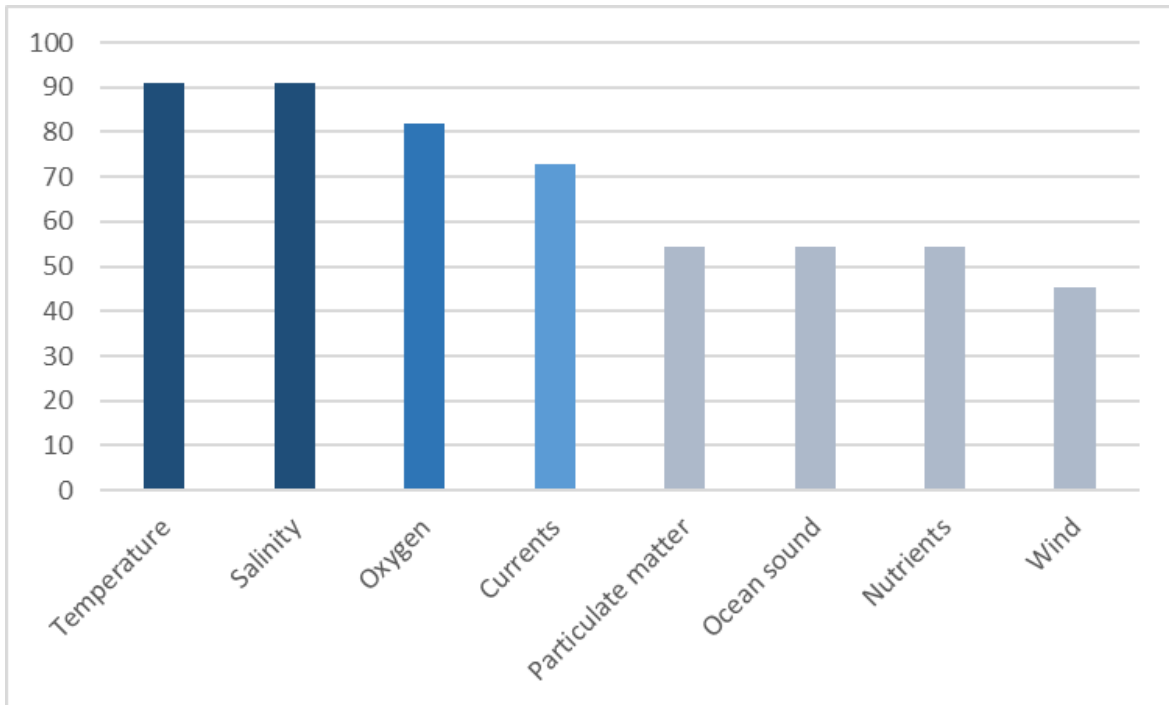


Fig. 4: Variables most monitored by the 11 RIs (percentage)

36% of respondents answered that the RIs they represent collect or intend to collect data on pH, followed by photosynthetically active radiation (PAR, 27%) (Figure 5).

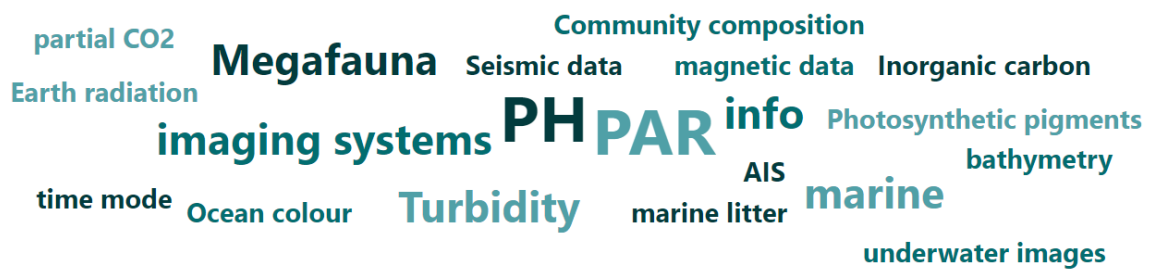


Fig. 5: Word cloud showing the non-standard parameters captured or planned to be captured

55% of respondents reported that the RI has an online data catalogue (Figure 6). However, the situation is not consistent across all facilities (e.g., EMSO), most of the URLs given do not refer to a data catalogue, and no reference to standard catalogue management systems was given. This is a clear indication that more competence in data management is needed.

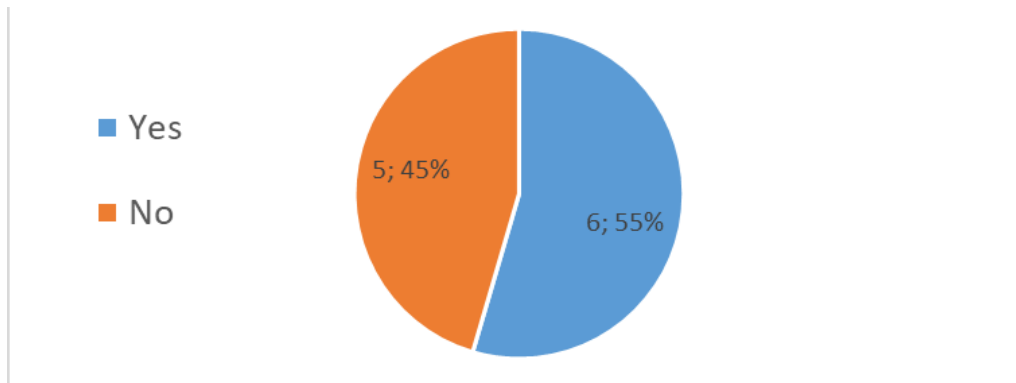


Fig. 6: Percentage of RIs with an online data catalogue according to respondents

61% of respondents said that the RI deposits data in a trusted repository, with Copernicus and EMODnet being the most used (Figure 7). Again, the situation is not the same in all EMSO facilities.



Fig. 7: Word cloud with the most frequently used trusted repository

50% of RIs allow programmatic access to all or some of their data via APIs (Figure 8), while % of respondents reported that their RI data are made openly available through an operational data access service (Figure 9).



Fig. 8: Percentage of RIs allowing programmatic access via APIs



Fig. 9: Technology readiness level of the RI data access service making data openly available

Only 36% of the RIs have a national/Italian data access service (Figure 10), and only 4 valid URLs of data access services were provided (1 hosted by CNR-ISP for SIOS Italy RI data, 1 hosted by ISPRA for Geoscience RI data, and 2 hosted by OGS for EMSO RI and ICOS RI data (1 facility and 2 platforms, respectively)). This information is further expanded in the next section, where the need to enhance the national capabilities with the support provided by the Marine Data Store is highlighted. Most of these cases relates to the CNR and in fact the Marine Data Store was initially developed to support them in improving data FAIRness.



Fig. 10: Percentage of RIs having/not having a national/Italian data access service

50% of the RIs do not have an identity provider to access data and services (Figure 11).

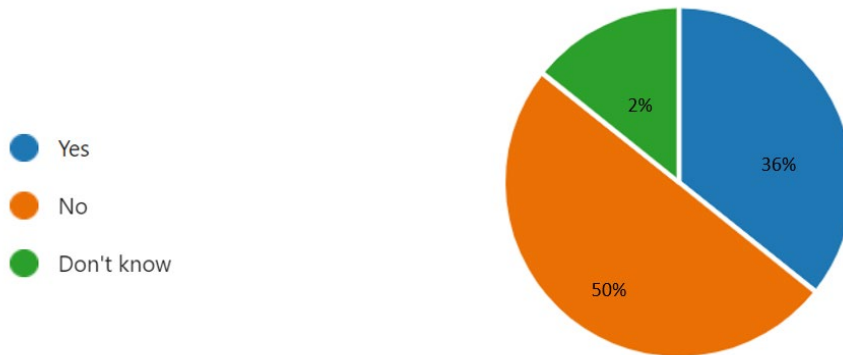


Fig.11: Use of an identity provider by RIs

50% of RIs do not have a general policy for user authorisation; permissions are granted on an individual basis or have not been decided yet (Figure 12).



Fig. 12: Adoption of general policy for user authorisation by RIs

Concerning interoperability, 70% use NERC data and metadata vocabularies (Figure 13), as well as data and metadata formats, standards, and methodologies adopted by SeaDataNet to make data interoperable.

In addition, 70% of respondents reported that the RIs do not have a SPARQL endpoint exposed on the web for semantic metadata access.

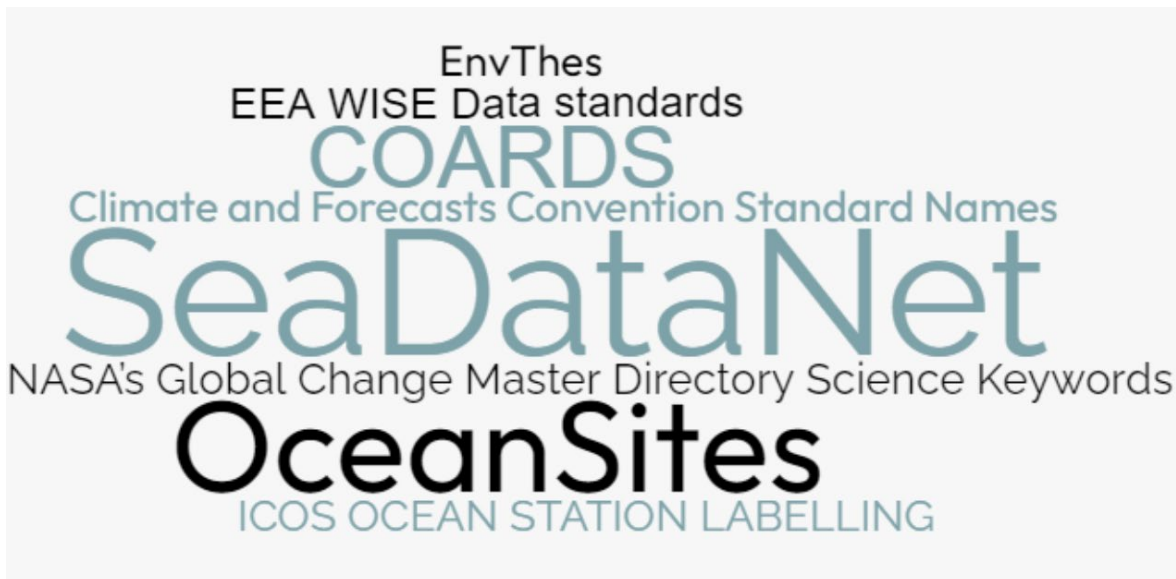


Fig. 13: Interoperability: most used data and metadata vocabularies.

In terms of data validation, almost all RIs (79%) follow processes to ensure data quality. In addition to well-known protocols such as SeaDataNet quality control protocols and flagging scales, there was also a frequent response that each dataset contains information on the quality assurance protocol applied (Figure 14).



Fig. 14 RIs following processes to ensure data quality.

86% of RIs have a preferred licence for publishing data (Figure 15); the Creative Commons Attribution-Non-commercial 4.0 International Public License and the Creative Commons Attribution 4.0 International were mentioned. Two data policies were indicated and are available at

- 1) <https://nodc.ogs.it/data/datapolicy> and
- 2) <https://www.icos-cp.eu/data-services/about-data-portal/data-license>



Fig. 15: Percentage of RIs having a preferred licence for publishing data.

Data reusability is facilitated by the production and circulation of specific documentation, such as readme files, for 57% of RIs.

Lastly, regarding persistent identifiers, RIs use them in 71% of cases (for all datasets or some), with Digital Object Identifier (DOI) being the most common type (Figure 16).

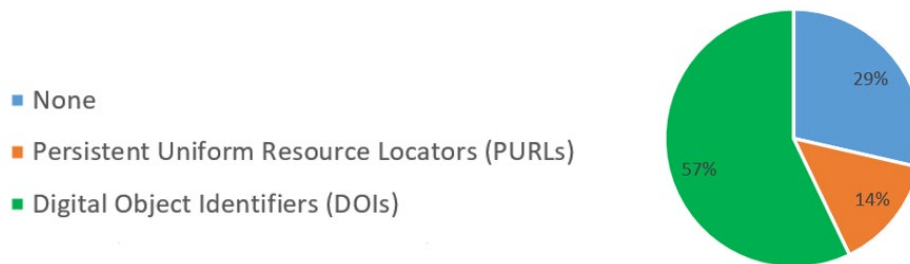


Fig. 16: Most used identifiers.

## DATA MANAGEMENT PLAN

Given the extensive information provided by the Marine RIs on the level of maturity of the discovery, access, and delivery services, in order to streamline the proper management of the new ITINERIS datasets complementing existing one and contributing to the Marine Data Centre, a first version of the Data Management Plan (DMP) is being developed. This document can be considered as a living reference to be updated during the project lifetime.

Data management is an integral part of ITINERIS, and the various aspects are spread across the different activities, resulting in a few outputs related to data management. The final DMP will be the result of the combination of the set of deliverables listed below, covering important aspects such as data flows, metadata, data, and service catalogue. While the data management practices are described below, the technical solutions to be adopted will depend on the maturity of the marine RIs and will be defined in coordination with the responsible organisations. Therefore, this DMP is considered as a first milestone to be complemented by the following project deliverables:

**D5.1** Marine Data Centre Management Plan: management needs, service requirements, governance for the distributed and federated system (Activity 5.3; B3; OGS\_OCE),

**D5.2** Report on Marine Data Store planning with data flows and connection standards for the distributed and federated system interfaced with ITINERIS central hub (Activity 5.2; B3; CNR-ISMAR-Napoli)

**D5.11** Report and release of IOOS metadata catalogue of data and facilities (Activity 5.3; B9; OGS\_OCE)

**D5.12** Report and release of IOOS catalogue of services (Activity 5.2; B12; CNR-ISMAR-Napoli)

**D5.18** Test report on the connection among RI distributed systems with IOOS Marine Data Store (Activity 5.2; B14; CNR-ISMAR-Napoli)

**D5.20** IOOS integration with ITINERIS central Hub (Activity 5.2; B14; CNR-ISMAR-Napoli)

**D5.21** Report on demo of the prototype web applications (Activity 5.4; B14; INGV-BO)

**D5.22** Report on common FAIR metadata catalogues, vocabularies, services, and data for the marine data centre, with the involvement of all participating RIs (Activity 5.3; B14; OGS\_OCE)

**D5.27** Final report and release of datasets (monitoring modelling) for specific environmental issues (Activity 5.8; B14; ISPRA)

**D5.28** Final report and release of new modelled and integrated products provided in the data centre (Activity 5.17, 5.19; B14; CNR-ISMAR-Venezia)

Data management involves many aspects: adequately describing the data and metadata, ensuring long-term appropriate storage and retention of data and metadata, documenting opportunities for dissemination and reuse through licences, and sharing data and handling procedures for sensitive data. These activities need to be documented so that all interested parties know their responsibilities and to ensure that they are followed.

Each Marine RI has its own data policy and DMP. The Marine Data Centre Management Plan defines the data flow for the ITINERIS data, from their creation to their dissemination, and the integration into the marine RI's workflow. Different approaches are taken depending on the maturity and consolidation of the data management processes. This work plan ensures long-term data storage while maintaining interoperability by relying on the distributed approach of federated systems.

### The Marine Data Store

Data management will be carried out by **two complementary data centres** in synergy with the Marine RIs. **CNR ISMAR** will oversee the planning and development of the ITINERIS **Marine Data Store**. It will i) provide the civil infrastructures, storage, computing, and transmission facilities, ii) develop the services for the harvesting, exchange, and access to all ITINERIS data, and iii) connect to the main ITINERIS HUB. It will focus at first on CNR data; however, it is designed to support the needs of all Marine RIs. The ITINERIS Marine Data Store together with the online data catalogue used to develop the single-entry point constitute the ITINERIS Marine Data Centre, designed as a federated system. **OGS** will oversee the definition and implementation of the dedicated **ITINERIS catalogue** for data and metadata, exploiting its experience as the Italian reference within the IOC/IODE network of UNESCO Data Centres. The ITINERIS Catalogue will collect information from all maritime RIs, acknowledging all ITINERIS data and data products. It will be designed by adopting the consolidated standards established at the EU level by SeaDataNet. This will also ensure the visibility of the ITINERIS data into the **BDIs**. By its implementation, the ITINERIS catalogue will guarantee the *findability* of ITINERIS data. ITINERIS Marine domain node interface with

SeaDataNet could (or will) be installed. *Accessibility* of the ITINERIS data will be ensured for both delayed mode and real-time data. Standard API and services could (or will) be implemented to achieve *interoperability*. The integration of ITINERIS data into SeaDataNet/EMODnet/Copernicus will be discussed and faced taking into account the maturity level of the Marine RI. Technical meetings will be organised by CNR ISMAR and OGS to identify the most appropriate technical solution and plan the complementary tasks together, taking into consideration the technical set-up already operated by the Marine RIs. Maximum visibility of the ITINERIS data, at both Italian and European levels, will be ensured. To streamline the collection of historical datasets, the EMODnet Data Ingestion portal (<https://www.emodnet-ingestion.eu/>) will also be used.

#### Four solutions that will have to coexist and converge

Depending on the maturity of the Marine RI and of the Italian research institutes, different scenarios should be identified to properly set-up the management plan for ITINERIS data. In fact, **different solutions shall be implemented simultaneously** depending on the maturity of the Marine RI, with a solution where both the Marine RIs that already have a data storage system with a distribution service (e.g., Euro-Argo, EMSO) and the less structured with be linked. If the data flow towards the EU Marine RIs or the Blue Data Infrastructures is not established, ITINERIS action will be taken to store, make visible, and distribute ITINERIS data. The figures below show **4 main scenarios that will be considered**. In blue the state of the art is visualised, in yellow the new components to be developed by ITINERIS have been inserted.

1. Scenario 1 - The observing facilities send data directly to the EU RI Data Collection Centre, which deals with data integration, validation, access, and redistribution (Figure 17). Euro-Argo ERIC is a good example of this. The Argo floats send their data to the French Coriolis Data Assembly Centre (DAC) via satellite communication every cycle once the floats are on the surface. As part of the ENVRI-FAIR project, Euro-Argo ERIC has been working to strengthen data FAIRness and improve its data access services.

In this scenario, **the ITINERIS marine data centre will retrieve ITINERIS and all other Italian data from the EU RI Data Collection Centre through operational services** and without duplicating the data. To this end, the RI's reference staff will need to take action to provide metadata for the ITINERIS catalogue and to acknowledge the ITINERIS contribution. Since ITINERIS needs to demonstrate its impact on RI observations and the use of ITINERIS data, the RI's reference staff will have to take action and modify the metadata stream so that data streams funded by ITINERIS can be easily identified and user accessing national data can be monitored, even when they access data from the EU RI Data Collection Centre.

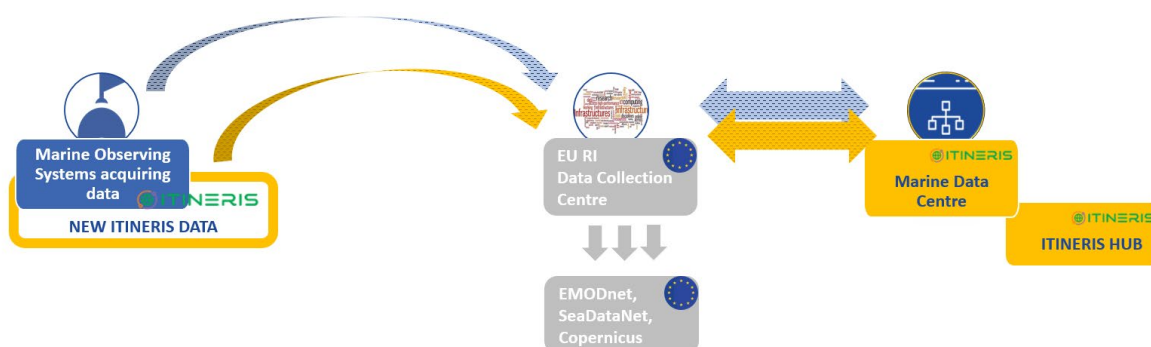


Fig. 17: Scenario 1 dealing with a mature EU RI Data Centre without an Italian data access service.

- Scenario 2 - Each observing facility sends data to a specific structured Italian research institute that collects, processes, and shares data with the EU RI Data Collection Centre and the BDIs (Figure 18). EMSO ERIC is a good example. Each observation facility constantly acquires and transmits various measurements, mostly in near-real and real time, to the owner Italian research institute, which in turn gives access to RI data through the connection with the EU EMSO ERIC data portal and BDIs.

In this scenario, **the ITINERIS marine data centre will retrieve ITINERIS and all other Italian data from the Italian research institutes through operational services** and without duplicating the data. To this end, the RI's reference person will have to take measures to provide metadata for the ITINERIS catalogues and to adapt the data services, if necessary.

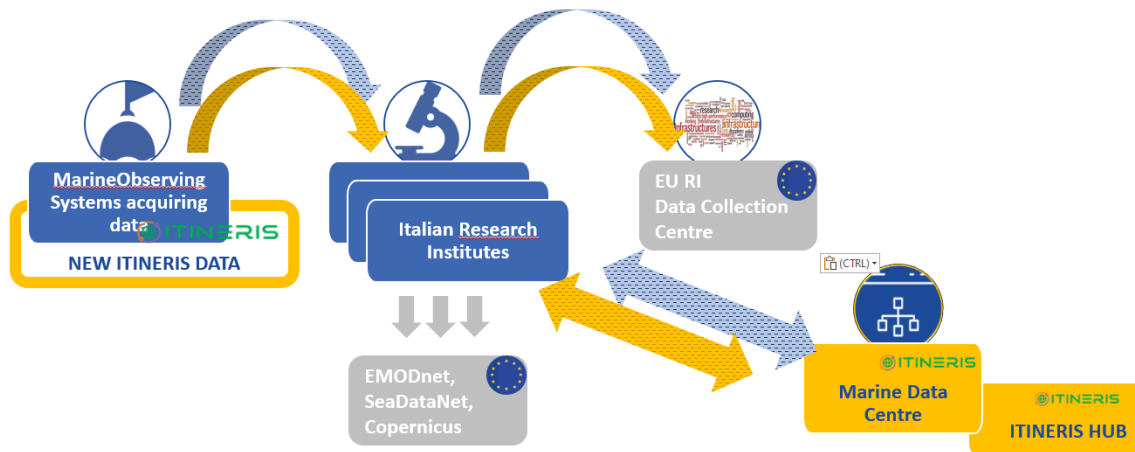


Fig. 18: Scenario 2 dealing with different Italian data access services sharing data with an EU RI Data Centre.

- Scenario 3 - The observing facilities send data to an unstructured Italian research institute that collects the data. DANUBIUS-RI is a good example of this. The observing systems might collect marine data in both real-time and delayed mode, which are i) not deposited in a structured and trusted repository, ii) not accessible through a web service, and iii) not connected to the EU RI Data Centre and the BDIs.

In this scenario, **the ITINERIS Marine Data Centre will support the ITINERIS data storage, standardisation, integration, access and sharing by implementing operational services** connected with the EU RI Data Collection Centre and the BDIs, following the consolidated EU standards (Figure 19). To this end, the RI's reference person will need to take action to describe the data (type, size, format, parameters, mode of collection and delivery), and to provide metadata for the ITINERIS catalogues.

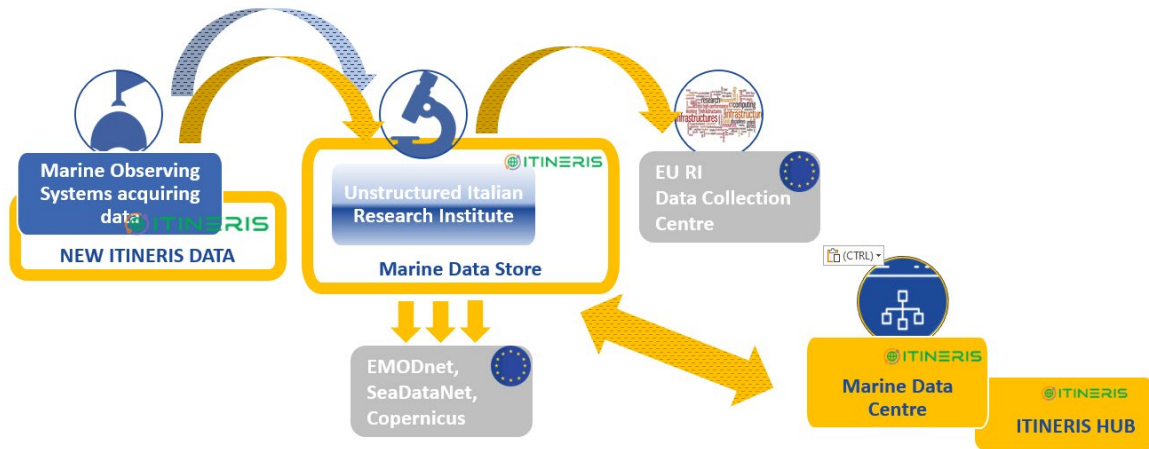


Fig. 19: Scenario 3 dealing with unstructured Italian research institutes not connected to the EU RI Data Centre.

- Scenario 4 - The observing systems send ITINERIS data to the reference unstructured Italian research institute that collects and validate them. These data might be new experiments as well as the ITINERIS Data Products result of Activity 5.19, 5.20, 5.21 or others. They will be handled by the ITINERIS Marine Data Store.

In this scenario, **the ITINERIS Marine Data Centre will support the storage, standardisation, integration, access, and sharing of the data and the derived products** with the BDIs (Figure 20). To this end, the RI's reference person will need to take action to describe the data (type, size, format, parameters, mode of collection and delivery) and provide metadata for the ITINERIS catalogues.

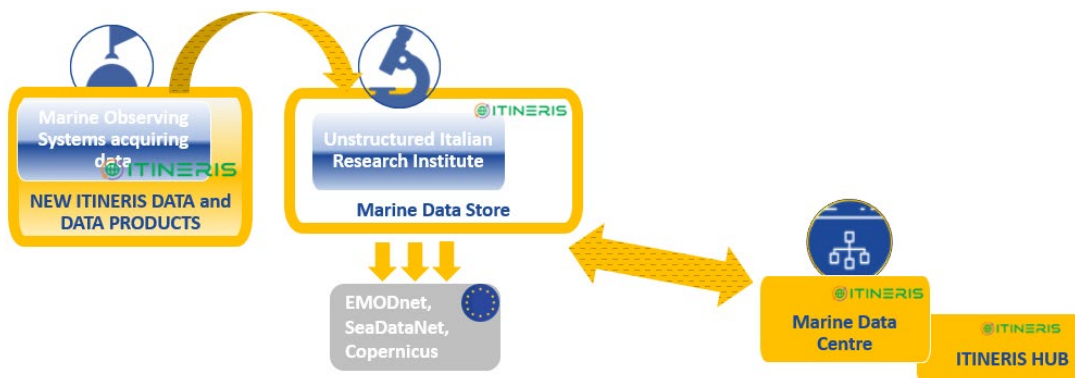


Fig. 20: Scenario 4 dealing with new ITINERIS data and data products.

## ITINERIS Data and Metadata catalogues

All Marine RIs should be listed and properly described in the **ITINERIS metadata catalogue**, with detailed information on active and planned monitoring sites (buoys, facilities, etc.). When a new platform or instrument is approved, the facility is created or updated in the ITINERIS metadata catalogue, which is maintained by the OGS and published on the ITINERIS portal hosted by the CNR.

The ITINERIS catalogue will be based on the European Directory of Marine Organisations (**EDMOs**), on the European Directory of Marine Environmental Research Projects (**EDMERP**), and on the European Directory of the Initial Ocean-observing Systems (**EDIOS**), all maintained by SeaDataNet and making use of SeaDataNet common vocabularies. ITINERIS catalogues will provide an overview of the i) Italian organisations operating the RIs, ii) monitoring systems, and iii) data and data products generated, acknowledging the ITINERIS contribution. In particular, the EDIOS directory contains information on the location, parameters measured, availability of data, responsible institutes, and links to data-holding agencies, as well as some technical information on the instruments, such as sampling frequency. EDMO, EDMERP and EDIOS entries makes use of XML formats and exchange schema's (XSD) based upon the ISO 19115 content model. EDIOS is hierarchical and includes three schemas for (1) observing programmes, (2) series and (3) platforms. Observing programmes may connect to series and platforms in a variety of ways. The observing programme may correspond to the EU RI, with more than one organisation contributing. Although the observing programme links to a number of series (i.e. data profiles), some observations may be collected by one platform and others by a different one [SeaDataNet 2013]. MIKADO software tool can be used to generate and edit XML metadata files for the EDMERP and EDIOS directory. More information on this Java tool can be found on the software section of [SeaDataNet web portal](#).

Mandatory Observing Programme information:

- EDIOS program identifier (local identifier)
- Program acronym
- Program Name
- Start date
- Sea-areas
- Abstract
- Data Access Restriction
- Collate Centre (EDMO reference)
- Principal investigator/laboratory
- Coordinating institute (EDMO reference)

Example of an observing programme: GeoSciences - Rete italiana per le geoscienze (National tide gauge and wave networks)

Mandatory Series information:

- EDIOS series identifier (local identifier)
- Series acronym
- Series Name
- Geographic coverage (bounding box)
- Start date
- Abstract
- Parameters
- Quality control procedures
- Collate Centre (EDMO reference)

Example of series names connected to the GeoSciences programme: PSMSL101, PSMSL2098, PSMSL2136, PSMSL104, PSMSL2089, etc.

Mandatory Platform information:

- EDIOS platform identifier (local identifier)
- Platform Name
- Platform class
- Platform Owner (EDMO reference)
- Collate Centre (EDMO reference)

Example of a platform connected to the GeoSciences programme through the Series, PSMSL101: sea level station, Ancona (GeoSciences Sea Level Station)

The representatives of the Marine RIs are requested to provide metadata as comprehensive as possible on two levels: the facility and datasets. Guidelines for metadata creation, including a webinar will be dedicated to train Marine RIs operators on the use of MIKADO software and on the compilation of the ITINERIS catalogue entries. The training activities should immediately start with **the webinar to be held by B4**. By the end of B6, **the representatives of all marine RIs must create the metadata entries for the ITINERIS metadata catalogue** giving the picture of the operational services.

All the metadata related to the new facilities and to the additional dataset (trajectories, time series, CTD profiles, sampling measurements and observations) must be provided by the RI managers within 2 months (together with the datasets or a data access point).

D5.11 will report detailed information on the IOOS metadata catalogue for data and facilities and its publication (Activity 5.3; B9; OGS\_OCE).

An ITINERIS dataset catalogue (D5.11) is developed by the OGS and its information is retrieved by the Marine Data Centre and displayed on the ITINERIS portal.

## ITINERIS MARINE DATA POLICY

The ITINERIS marine data policy is derived from the COI data policy proposed and recently revised by the IODE working group. It sets the framework for accessing the data and information provided through the operational data access portal and is the reference at the Italian level.

The data policy aims to strike a balance between the rights of data producers and the need for free, open, and unhindered sharing and exchange. The aims are to serve the scientific community, public organisations, and environmental authorities and facilitate the dissemination of data and information, as well as their access and use. **The ITINERIS marine data policy applies to all ITINERIS data collected and managed by the marine RIs** and made available by the Marine Data Centre through the single-entry point connected to the main ITINERIS hub. Data will be managed in a distributed system (i.e. the data will be archived on different geographically distributed hubs, but will still be accessible through a single portal).

This data policy is in line with the European Directive 2003/4/EC on public access to environmental information [EC 2003], with the European Directive 2019/1024 on open data [EU 1024] and the implementing documents [EU 2023/138][DL 200], with the Directive INSPIRE [EC 2007], with the principles of the SeaDataNet data policy [SeaDataNet 2006], with the principles of IOC [IOC 2023] and EuroGOOS [EuroGOOS 2023].

## Data Use Policy

In principle and subject to the exceptions explicitly stated in the access rights, **the Marine Data Centre provides data freely and without restriction**: "freely" means that the data itself is free of charge; "without restriction" means that there is no discrimination against, for example, individuals, research groups, nationalities, or on the basis of purposes that may be commercial in nature, and in any case in accordance with the applicable legislation on the re-use of public sector information (Directive 2019/1024 of the European Parliament). Specific requests will be forwarded to and processed by the data provider.

The Marine Data Centre single-entry point provides data to users in a timely and simple manner, using state-of-the-art communication tools, but within the limits of the data released by the parties. The portal is the central point for the dissemination of the data. Depending on the different types of information, the access conditions vary:

- 1) the metadata are freely accessible without any conditions and associated with a Creative Commons licence CC BY 4.0.
- 2) the data and products are freely accessible according to the COI Data Policy. User registration and acceptance of a specific licence for use are required.

Registration is requested and then used to generate usage statistics.

## User licence

1. The Licensor grants to the Licensee a non-exclusive and non-transferable licence to receive and use data and products accessed through the Marine Data Centre single-entry point in accordance with this Licence. It is suggested that all data is provided under the CC-BY 4.0 licence unless otherwise stated by the Marine RI data policy.
2. Receipt of data and information by electronic download and use thereof is free of charge, unless otherwise provided for in the access rights.
3. Whether or not data quality is controlled, the Marine Data Centre and its individual constituents do not accept responsibility for the accuracy and/or reasonable interpretation of the data. The correct and appropriate interpretation of the data must follow scientific rules and is always the sole responsibility of the users.
4. Users are expected to acknowledge that ownership of the metadata rests with the Marine RIs, while ownership of the data remains with the data provider. It is unethical to publish data downloaded from the Marine Data Centre portal without proper attribution to the owner. To support good citation practises, each dataset listed on the Marine Data Centre portal is provided with a digital identifier and appropriate bibliographic citation.
5. Users are required to inform the Marine Data Centre of any problems with the data provided in order to correct and improve the quality of such data.

## Access rights

Specific access restrictions for individual datasets and products might be assigned by the Marine RI, and ultimately by data provider, for each ITINERIS data or data product entered into the portal. The list of access rights can be found in list L08 (SeaDataNet Data Access Restriction Policies) of the SeaDataNet Vocabularies (<https://www.seadatanet.org/Standards/Common-Vocabularies>).

## CONCLUSIONS

**The primary tasks of ITINERIS Marine Data Centre** are to provide open access to and manage ITINERIS marine data resources following the FAIR principles, building on community data management standards, tools, and infrastructures. The Marine Data Centre demands effort and in-depth analysis of the existing landscape and technological developments.

**Its functions include:** 1) capture ITINERIS data originated by the Marine RIs; 2) ensure archiving where necessary and required; 3) provide outputs to various research and engineering users, forecasters, experiment managers, or to other WPs within the project, in particular to the VREs; 4) disseminate data on the internet through dedicated (cloud) servers and other means; and 5) provide indicators for the different types of data being exchanged in order to track the progress.

The most important **prerequisite for the establishment** of an effective Marine Data Centre is the support and cooperation of the Marine RIs that collect data. Without this cooperation, the new centre will face great difficulties in obtaining access to and managing ITINERIS data. Thanks to the cooperation of the Marine RIs involved in the realisation of the ITINERIS HUB, the Marine Data Centre will not duplicate data and efforts but will harmonise the existing ones. Most of its activities will aim at achieving the most appropriate level of FAIRness for all RIs involved.

**The first outcome** of the survey showed that more than half of the Marine RIs uses a trusted repository to store their data, pointing to the standards and data infrastructures of Copernicus, SeaDataNet, and EMODnet. The majority of Marine RIs allow programmatic access via APIs, with an operational service. It is important to note that a small minority has a national/Italian data access service, suggesting that the ITINERIS Marine Data Centre should be capable of connecting the national resources and supporting each RI to bridge the gap to fulfil the requirements of data FAIRness.

To this end, **the Centre will use** the IOC IODE programme instructions and **will leverage** the expertise of the Italian NODC in building the ITINERIS catalogues, managing and making available ocean data, information, products, and services, and **will consider** the structure and the governance determined in the COI agreement for the establishment the Italian Oceanographic Data Committee. The Marine Data Centre Management Plan will therefore be consistent with the 1) objectives of the ITINERIS project, 2) structure of cooperation defined in the COI agreement, aimed at developing a distributed and federated system of national databases, its data policy, and obligations, and 3) existing pan-European data services.

Given the above ideas and principles the following steps are suggested as an outline on proceeding to establish the Marine Data Centre:

**Step 1. Engagement of the parties** (data managers, ICT experts) to ensure that effective consultation occurs while proposing a mission and organisational model for the centre. This team should represent all the Marine RIs, starting with the representatives of the organisations who replied to the questionnaire. Workshops may provide a suitable forum for the generation of ideas and directions. These also may suggest ICT solutions that might be useful to the development.

**Step 2. Identify responsibilities and the governance model** for the Marine Data Centre. Since its conception, the centre has been designed as a distributed and federated system supported by a Marine Data Store, to be developed in Naples, and by a single-entry-point for the marine community, to be interfaced with ITINERIS central hub. The roles should be clearly defined. It might be beneficial to set up a **Steering Committee** with representatives from the Marine Data Store, the Italian NODC, at least one RI representative per scenario, and the co-chairs of WG IOOS e WG RIs.

**Step 3. Identify a range of skills and experience** required in the process of development, management, and future maintenance. Consider discussing the options with interested parties. A possible initial model would include a data centre manager, with some administrative support, 2–3 data managers with a marine science background, 2–3 software developers, and a communications person to promote the data centre, respond to data requests, and provide advice to data centre users [UNESCO 2022]. This can be refined as necessary in the planning stages and again as the centres become established when the workload may increase. Assuming the data centre will be located within CNR-ISMAR Naples, then the hosting institute could provide the administrative support and IT infrastructure.

**Step 4. Landscape analysis** starting from the results of the questionnaire on the state of the art of RIs management and service needs and requirements. The in-depth length will help identify the standard in use by the Marine RIs to be adopted for the development of ITINERIS connection to data and the access services (D5.12 Report and release of IOOS catalogue of services, D5.18 Test report on the connection among RI distributed systems with IOOS Marine Data Store). From this study, review the opportunities for useful synergies with other Italian and EU supported activities and projects. Also discuss the structure of the metadata and data catalogues describing data from the Marine RIs and the interface with the ITINERIS Marine Data Store. Then, review the needs of the centre for specific types of expertise and make a first attempt to define a client base.

**Step 5.** Before starting the development, it is important to consider the entire process while keeping in mind the final goal. Aim for **defining a draft action plan with intermediate deadlines in synergy with D5.2.. Step 6.** Leading the release of IOOS metadata catalogue of data and facilities (D5.11). This will be used to implement the ITINERIS Marine Data Access portal.

## REFERENCES

[Burgelman et al., 2019] Burgelman JC, Pascu C, Szkuta K, Von Schomberg R, Karalopoulos A, et al. (2019) Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in Big Data* 2(43).

[DL 200] DECRETO LEGISLATIVO 8 novembre 2021, n. 200. Attuazione della direttiva (UE) 2019/1024 del Parlamento europeo e del Consiglio, del 20 giugno 2019, relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico (rifusione). GU n.285 of 30-11-2021 - Supplemento Ordinario n. 42/L.

[EC 2003] Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC.

[EC 2007] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

[EC 2018] Directorate-General for Research and Innovation (European Commission) (2018) Turning FAIR into reality. Final report and action plan from the European Commission expert group on FAIR data.

[EU 1024] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

[EU 2023/138] Commission Implementing Regulation (EU) 2023/138 of 21 December 2022 establishing a list of specific high-value datasets and modalities for their publication and re-use.

[EuroGOOS 2023] EuroGOOS Data Policy (2023).

[IOC 2023] Annex to Resolution XXII -6 IOC Oceanographic Data Exchange Policy (2023).

[SeaDataNet 2006] SeaDataNet I3 Pan-European infrastructure for ocean and marine data management - SeaDataNet Data Policy, 2006.

[SeaDataNet 2013] SeaDataNet Guidance for preparation of entries for the European directory of the ocean-observing system (EDIOS), SeaDataNet Deliverable 4.6 B, 29 pp., v2.0, 2013.

[UNESCO 2008] Guide for establishing a National Oceanographic Data Centre, Second revised edition. Oostende, Intergovernmental Oceanographic Commission of UNESCO, 27 pp. (IOC Manuals and Guides No. 5, 2nd. rev. ed.) (English), Electronic publishing 2008.

[UNESCO 2022] Intergovernmental Oceanographic Commission of UNESCO. 2022. Guide for Establishing an IODE National Oceanographic Data Centre, IODE Associate Data Unit or IODE Associate Information Unit (3rd revised edition). Paris, UNESCO, 26 pp. (IOC Manuals and Guides, 5, 3rd revised edition) (English).

[Wilkinson 2016] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.

## ANNEX - QUESTIONNAIRE FOR MARINE RESEARCH INFRASTRUCTURES