MISSIONE 4
ISTRUZIONE
RICERCA

ITINERIS

D5.6 Report on the Design of the Italian Integrated Ocean Observing System (IT-IOOS) based on gap analysis and governance

Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

Deliverable number:	D5.6
Work package:	WP5 – Marine Domain
Intermediate Objective:	IO5.5
Deliverable type:	X Document, report
	☐ Websites, patent filings, videos, etc.
	☐ Other: please specify ....
Dissemination level:	X Public
	☐ Restricted
Estimated delivery (bimester):	B7
Actual delivery date:	31/12/2023
Author(s) (Partner-OU):	Rosalia Santoleri (CNR-ISMAR), Valeria Ibello (CNR-ISMAR), M. Caccavale (CNR-ISMAR), A. Giorgetti (OGS), C. Cantoni (CNR-ISMAR), D. Bellafiore (CNR-ISMAR), F. De Pascalis (CNR-ISAMR), M. Giani (OGS), L. Beranzoli (INGV), A. Fadini (CNR-ISMAR), M. Magaldi (CNR-ISMAR), F. Coren (OGS), E. Mauri (OGS), G. Notarstefano (OGS), A. Gallo (OGS), C. Solidoro (OGS), E. Urbinati (CNR-ISMAR), R. Bozzano (CNR-IAS), G. Buscaino (CNR-IAS), A. Petrocelli (CNR-IRSA), C. Mantovani (CNR-ISMAR), K. Schroeder (CNR-ISMAR), S. Miserocchi (CNR-ISP), G. Riccobene (INFN), S. Simncelli (INGV), D. Embriaco (INGV), C. Altobelli (OGS), V. Cardin (OGS), V. Langella (CNR-ISMAR), F. V. M. Kokoska (CNR-ISMAR), A. Pomaro (CNR-ISMAR), E. Organelli (CNR-ISMAR), G. La Forgia (CNR-ISMAR), C. Bergami (CNR-ISMAR Milano/Bologna), S. Toller (CNR-ISMAR Bologna), A. Oggioni (CNR-IREA Milano)
Reviewed by:	ITINERIS Executive Board
Note:	

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

Index of tables

Index of figures

1. INTRODUCTION

1.1. Purpose of the document

The aim of this document is to design the ITalian Integrated Ocean Observing System (IT-IOOS). Based on user requirements and ITINERIS main objectives, the general IT-IOOS system aims, performance, system and service requirements are identified and used to design the IT-IOOS system architecture. An implementation plan is also presented to achieve the ITINERIS goals to actualize and test the IT-IOOS within the time frame of the project. The short-term and long-term strategy for the governance and maintenance of IT-IOOS within the time leg of the ITINERIS project, and in the next decade is also proposed.

1.2. Applicable and reference documents

This document is applicable to the Italian Integrated Environmental Research Infrastructures System (ITINERIS) Project.

1.1.1 Applicable documents

This document is applicable to the development of the IT-IOOS under the framework of the Italian Integrated Environmental Research Infrastructures System (ITINERIS) Project. On this aim the Applicable documents are:

D1. ITINERIS project full proposal

D2. Attachment B sub 31.1.3 - Economic and financial plan for the operation of the infrastructure(s) as resulting from the project, for at least ten years starting from the final payment

D3. Attachment B sub 33 - Timing of the different work packages and their components

D4. Attachment B sub 34 - WP inter-relation with other WPs.

D5. Piano Nazionale Infrastrutture di Ricerca (PNIR) 2021-2027

D6. ITINERIS Deliverable D2.10 – DMP – Data Management Plan

1.1.2 Reference documents

R1- Intergovernmental Oceanographic Commission of UNESCO. 2023. IOC Strategic Plan for Ocean Data and Information Management (2023–2029). Paris, UNESCO, 18pp. (IOC Manuals and Guides, 92).

R2- European Ocean Observing System (EOOS). Strategy 2023-2027 Advancing EOOS - the foundation of European ocean knowledge.

R3 Giorgetti, A., Altobelli, C., Azzaro, M., Bellafiore, D., Bergami, C., Bozzano, R., Cardin, V., Coren, F., De Pascalis, F., Embriaco, D., Gallo, A., Giani, M., Giorgi, G., Magaldi, M., Mauri, E., Miserocchi, S., Notarstefano, G., Riccobene, G.M., Schroeder, K., Simoncelli, S.. 2023. Marine Data Centre Management Plan: management needs, service requirements, and governance for the distributed and federated system. Deliverable D5.1 in ITINERIS project.

R4 Caccavale M. 2023. Marine Data Store planning with data flows and connection standards for the distributed and federated system interfaced with ITINERIS central hub Deliverables. Deliverable D5.2 in ITINERIS project.

R4- Bellafiore D., De Pascalis, F., Pomaro, A., Schroeder, K., Bergami, C., Magaldi, M., Santoleri, R., Organelli, E., Embriaco, D., Pensieri, S., Buscaino, G., Papale, E., Bozzano, R., Riccobene, G., Giorgi, G., Mauri, E., Cardin, E., Giani, M., Cantoni, C., Petrocelli, A.. 2023. Report on gap analysis

of the national marine facilities and equipment and access procedures. Deliverable D5.3 in ITINERIS project.

R5- M. Azzaro, S. Miserocchi, L. Langone, F. Giglio, P. Giordano, G. Verazzo, F. Filiciotto, V. Sciacca, A. Lo Giudice, F. Paladini de Mendoza, M. Pansera, F. Decembrini, F. Smedile, M. Papale, M. Bensi. Report on gaps identification in data acquisition of the Arctic region. Definition of the new architecture of SIOS infrastructure; procedures for implementation of the observatory platforms. Deliverable D5.4 in ITINERIS project.

R6- M. Magaldi, D. Lagomarsino-Oneto, C. Bergami, S. Toller, D. Bellafiore, F. De Pascalis, E. Organelli, G. La Forgia, J. Pitarch Portero, M. Bellacicco, E. Mauri, V. Cardin, A. Petrocelli, R. Bozzano, S. Pensieri. Report on the gap analysis of biological and ecosystem observations. Deliverable D5.5 in ITINERIS project.

R7- A Roadmap for the Implementation of the Global Ocean Observing System 2030 Strategy. IOC, Paris, 2020, GOOS Report No. 249 (IOC/BRO/2020/6)

R8- Lindstrom, E., Gunn, J., Fischer, A., McCurdy, A., Glover, L., Alverson, K., et al. (2012). A Framework for Ocean Observing. By the Task Team for an Integrated Framework for Sustained Ocean Observing. Paris: UNESCO. doi: 10.5270/ OceanObs09-FOO

R9- COI (Commissione Oceanografica Italiana) agreement for the data exchange and the architecture design of the Italian Ocean Data Center (IODC)

R10- H2020 DANUBIUS-PP Consortium (2019) Science and Innovation Agenda of DANUBIUS-RI - The International Centre for Advanced Studies on River-Sea Systems.

R11.Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). https://doi.org/10.1038/s41597-020-0486-7

R12. 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

R13. IOC Resolution XVI-8. Intergovernmental Oceanographic Commission Reports of Governing and Major Subsidiary Bodies IOC Committee on Ocean Processes and Climate Fourth Session Paris, 27 February-1 March 1991 UNESCO, IOC/OPC-IV/3Paris, 19 April 1991

R14. Lindstrom, E., Gunn, J., Fischer, A., McCurdy, A., and Glover, L. K. (2012). "A Framework for Ocean Observing," in Proceedings of the Task Team for an Integrated Framework for Sustained Ocean Observing, UNESCO 2012 (revised in 2017), IOC/INF-1284 rev.2, Venice

R15. Tanhua, T., Fischer, A., and McCurdy, A. (2019). Successes and challenges in applying the framework for ocean observing. Front. Mar. Sci.

1.3. Norms and standards

N/A

1.4. Glossary, Acronyms and Abbreviations

DANUBIUS: International Centre for Advnced Studies on the River-Sea systems, https://www.danubius-ri.eu/

eLTER-RI: Integrated European Long-Term Ecosystem, critical zone and socio-ecological Research, https://elter-ri.eu/

GeoSciences: Research infrastructure for the Italian Network of Geological Services, https://www.isprambiente.gov.it/en/projects/soil-and-territory/geosciences-ir

EMSO-ERIC: The European Multidisciplinary Seafloor and water column Observatory, https://emso.eu/

EOOS: European contribution to the international Argo Programme, https://www.eoos-ocean.eu/

EURO-ARGO-ERIC: https://www.euro-argo.eu/

EUROFLEETS-RI: An alliance of European marine research infrastructure to meet the evolving needs of the research and industrial communities, https://www.eurofleets.eu/project-information/

EuroGOOS: the European Global Ocean Observing System, https://eurogoos.eu/

FOO, Framework for Ocean Observing

ICOS: Integrated Carbon Observation System, https://www.icos-cp.eu/

INGV: National Institute of Geophysics and Volcanology, https://www.ingv.it/

IR: Interfaces Requirements

ISMAR: Institute of Marine Sciences, National Research Council, https://www.ismar.cnr.it/

ITINERIS: Italian Integrated Environmental Research Infrastructures System, https://itineris.cnr.it/

IOOS: Integrated Ocean Observing System

IT-IOOS: ITalian Integrated Ocean Observing System

JERICO-RI: Joint European Research Infrastructure of Coastal Observatories, https://www.jerico-ri.eu/about/

N/R Laura Bassi: Research Vessel, owner OGS

LNS: Laboratori Nazionali del Sud, https://www.lns.infn.it/it/

MDC: Marine Data Center

MDCR: Marine Data Center Requirements

MDS: Marine Data Store

MDP: Marine Data Portal

OGS: National Institute of Oceanography and Applied Geophysics, https://www.ogs.it/en

PNIR: National Research Infrastructures Program

PR: Portal Requirements

PSO: Public System Outage

RIs: Research Infrastructures

SIOS: Svalbard Integrated Arctic Earth Observing System, https://sios-svalbard.org/

SS: Sub-System

SSR: Sub-System requirements

1.5. Summary

Text here – TO BE ADD in the next version

2. OCEAN OBSERVING SYSTEM INTERNATIONAL CONTEST

2.0. Ocean Observing Systems in the international contest

The world's oceans are a critical component of the Earth system. Sound knowledge and understanding of the ocean is essential to mitigate human impacts on the global environment and to promote the ocean's sustainable use and blue economy. Effective ocean management depends on marine observations, which are generated by existing national or regional ocean observing systems and networks. Marine data are also increasingly relevant to a broad array of policymakers and stakeholders, including the industry sectors engaged with ocean observations.

In order to fulfill the need of information about the ocean and effectively manage ocean data, there is an imperative need for the observation and forecasting of the ocean and its intricate links to weather, climate, and biogeochemical phenomena. Enhanced early warning systems are crucial for mitigating risks associated with these hazards. Notably, the assessment of progress towards the United Nations (UN) Agenda 2030, specifically Sustainable Development Goal 14, which focuses on conserving and sustainably using oceans, seas, and marine resources, necessitates a continual influx of ocean data. Furthermore, the international community recognizes the global significance of systematic ocean observations. The UN Sustainable Development Goals underscore the necessity for increased resources dedicated to sustained ocean and coastal observation. Additionally, the 7th challenge of the UN Ocean Decade which makes up the largest focal area of the Decade is the worldwide reinforcement of the ocean observing capacity (challenge 7: Expanding the Global Ocean Observing System).

This collective effort aims to augment our understanding of the ocean, its correlation with climate dynamics, and the imperative to establish coordinated early warning systems for extreme weather events. The integration of ocean observations aligns with other global goals, such as those outlined by the Convention for Biodiversity, Europe's Marine Strategy Framework Directive, and assessments conducted by organizations like the Intergovernmental Science-Policy Platform for Biodiversity and Ecosystem Services (IPBES) and the World Ocean Assessment.

The UN Framework Convention on Climate Change (UNFCCC) and the Paris Agreement (2015), in particular, stress the need to ensure the integrity of all ecosystems, including those within the ocean. They call upon countries to reinforce systematic observation of the climate system and early warning systems that inform climate services and support decision-making. Ocean observations are also deemed indispensable for validating climate projections assessed by the Intergovernmental Panel for Climate Change (IPCC). In essence, systematic and sustained ocean observations emerge as a cornerstone for addressing global challenges, promoting sustainable development, and advancing climate resilience.
The UNESCO Intergovernmental Oceanographic Commission (IOC) since 1991 has established GOOS (Global Ocean Observing System) program (R13) with the aim to lead the development of a truly global ocean observing system that delivers the essential information needed for our sustainable development, safety, wellbeing and prosperity. In response to the science and user requirements GOOS has developed the Framework for the Ocean Observing (FOO, Lindstrom et al., 2012), an important guidance that has identified key principles and processes to design and maintain ocean observing systems that is fit-for-purpose responding both to scientific inquiry and societal needs. FOO has strived for a balanced integration of research, innovation, and stability, fostering alignment among independent groups and utilizing existing structures. According to FOO principles the international coordination and organization of ocean observing system should center on "essential ocean variables (EOVs)" rather than being based on specific observing systems, platforms, programs, or regions. The consensus is that the introduction of new EOVs should follow their readiness levels,

facilitating the prompt incorporation of components that are already mature. This approach would encourage innovation and concerted efforts to enhance readiness and capacity.
The FOO principles implies an expansion of requirements for GOOS beyond from its traditional focus on the ocean's role in weather and climate. GOOS now also encompasses operational services and marine ecosystem health, from the open ocean into coastal environments where much of the world's population resides (Moltmann et al 2019). This has opened a change in the development of global ocean observing system enhancing the relevance of the development of national integrated observing systems as main pilar of the the global ocean observing development where the GRA (GOOS Regional Alliance) facilitate the cooperation at regional scale.
FOO principles have been applied in the construction of different National Integrated Ocean Observing system, such as the Australian IMOS (https://imos.org.au/) which is an Australian national research infrastructure funded by the government. IMOS has been one of the first Ocean Observing System that integrates data across scales and disciplines, prioritizing Essential Ocean Variables (EOVs) based on national science planning. In Europe, the expansion and integration of ocean observing systems, such as the implementation of the FOO principles is still an ongoing process in many nations.
At European level, several key initiatives were developed to contribute to the development of the Sustained Ocean Observing as defined by GOOS and to FOO principle implementation.
The European Ocean Observing System (EOOS) was designed by EuroGOOS (the European Regional Alliance of GOOS) and the European Marine Board, with the aim to connect the European stakeholders in ocean observation. EOOS is expanding its scope to include marine ecosystem health, climate observations, and applications, alongside traditional real-time oceanographic services. It aims to create a mechanism for users to communicate needs, transforming them into data requirements and optimal measurement strategies. EOOS adopts the feedback loop from FOO to track and assess the implementation of solutions in response to user needs. Looking ahead, EOOS plans to leverage a more mature FOO for effective user feedback mechanisms, contributing to ongoing advancements in ocean observing system design.
The most important initiative launched by Europe to contribute to GOOS objectives was the design, development of operational oceanography and implementation of Copernicus Marine Service (https://marine.copernicus.eu/). The Copernicus Marine Service is the marine component of the Copernicus Programme of the European Union, provides free, regular and systematic authoritative information on the state of the Blue (physical), White (sea ice) and Green (biogeochemical) ocean, on a global and regional scale.
These initiatives have benefited of significant investments made by European nations and European Commission in the last two decades, which recognize the importance of enhancing ocean observation and developing operational services, support and progress research, innovation, data access and management, and technology. However, the existing ocean observing systems have been funded by short-term projects which has prevented their overall long-term continuity. The need to realize a long-term sustained integrated ocean observing system has been long envisioned by EuroGOOS and EOOS, but at the current state, there is no European country that has developed a stable integrated system.

2.1. Italian oceanographic contest

At Italian national level, the exploration of the marine environment is fragmented in multiple research centers, universities and consortia, without an overall coordination. It results that ocean data are affected by fragmentation, lack of standardization and limited accessibility, which prevents an effective management of Italian seas and their resources.

To overcome the inefficiency in ocean data management, the new GOOS strategy approved by the UNESCO-IOC assembly in June 2023 requested to member states to work at national level to implement GOOS strategic goals and support the alignment of national activities in a broader international and regional framework. UNESCO-IOC solicited the member states to take actions to coordinate the development and maintenance of the national observing system across inter-disciplinary sectors, establishing the GOOS national committee and nominating a Focal Point.
For the purpose to respond to IOC GOOS strategy, the Italian Oceanographic Commission (COI, the national body of UNESCO-IOC), has established the National GOOS working group.
The National GOOS working group has the role to coordinate the national design and the implementation of national ocean observing system, to enhance the participation of Italy to the international ocean observing community, as GOOS and EOOS. The COI has also the role to respond to IOC recommendation of the International Ocean Data Exchange (IODE) Programme on data sharing. For this purpose, the National Ocean Data committee, established by COI, has prepared an agreement about data exchange and design the architecture of the Italian Ocean Data Center (IODC). This agreement has been approved by COI and is currently under evaluation of the Presidents of COI organizations.
These complementary actions will not only allow a comprehensive and harmonized planning of ocean observation initiatives at national level but will also optimize the management of ocean data for an efficient use from users of academia, society and private sector.

2.2. Research Infrastructures and Italian Integrated Ocean Observing System

The constitution of long-lasting ocean observing systems requires permanent structures sustained by long term funding support. In Europe, permanent structures and funding support can be identified with the Research Infrastructures (RIs) which are pilars of excellence science to support frontier research on long-term, established by the European Commission together with Member States and the scientific community. The ongoing development of Research Infrastructures is set to advance through the Research Infrastructures Programme and the European Strategy Forum on Research Infrastructures (ESFRI) which aims to empower and integrate the European Research Infrastructures. The integration of RIs specifically among the marine RIs can potentially favor the realization and construction of the European segment of the GOOS and guarantee the enduring delivery of robust quality-controlled scientific data. However, the challenge arises from the coordination of these RIs, given that each marine RI operates based on its distinct priorities and development strategy.

Italy participates in all main environmental RIs of pan-European interest and hosts other environmentally relevant national RIs. Further, the strategy of the National Plan for Research Infrastructures 2021-2027, points to empower its national research infrastructure nodes and integrate them into a system of systems. Being the Research Infrastructures long-term supported structures of the European Research System, Italy has chosen to build-up the ITalian Integrated Ocean Observing System (IT-IOOS) on these structures. The Recovery and Resilience Plan (PNRR) call has provided the opportunity to realize, within the ITINERIS project, the IT-IOOS by integrating harmonically Marine Research Infrastructures, including, data, products, facilities and services.

By implementing IT-IOOS, Italy will contribute to European and international efforts on ocean observations (EOOS and GOOS), responding to the major challenges of both the GOOS 2030 strategy (R7) and the Ocean Decade for the Mediterranean Sea.

3. ITALIAN Integrated ocean observing system of systems (IT-IOOS)

3.1. IT-IOOS setup and Goals

The IT-IOOS designed by this document seeks to coordinate national observations for three critical themes: climate, operational services, and marine ecosystem health. IT-IOOS will address immediate needs regarding quality and interoperability of ocean data and will enable Italy to better adapt to changing requirements and a changing environment. Provision of open marine data will help to meet the societal needs for access to ocean information, will benefit the country by assisting in the development of ground-breaking research. IT-IOOS will contribute to the global ocean observation efforts and provide opportunities for Italy to become a leader in multidisciplinary ocean science and pioneers in the development of integrated multi-RIs system of systems.
The implementation of IT-IOOS will be achieved by harmonizing data and products carried out by the different RIs and building up an integrated system of systems able to ensure continuity of data and services and to respond to user needs.
The design and implementation of the IT-IOOS considers the data collected by the Marine Domain RIs and their progressive harmonization and updates. Data acquired from marine RIs, different platforms/systems, such as Eulerian systems (mooring & fixed boys), Lagrangian systems (floats & drifters), ships and Gliders, will be integrated and made visible by a single access point designed by this document. The Marine Data Store, designed following the guidelines of the Italian Oceanographic Commission (COI) as a distributed system, is implemented with the contribution of all Operative Units of ITINERIS Marine domain. IT-IOOS data center is built based on ITINERIS WP2 requirements and interfaced with ITINERIS central HUB to ensure access to data and services.

3.2. IOOS Strategy and implementation plan: ITINERIS framework

In this section, we present the plan for the IT-IOOS implementation. The IT-IOOS system designed in this document will be fully implemented during the ITINERIS project by integrating progressively the Research Infrastructures. The design will consider future upgrade of Italian system of systems as well as long-term maintenance of system, according with the PNRR RI ITINERIS contract requirement. Consequently, the implementation plan will be divided between a short-term implementation limited to the ITINERIS time frame and a long-term strategy for the next decade (ten years road map).

The IT-IOOS implementation, presented in this section, includes the design of the IT-IOOS system architecture, and the definition of IT-IOOS implementation plan and strategy to reach the full integration of the marine domain RIs in a single system that is the main objective of ITINERIS.

3.2.1. Short-term implementation plan

The short-term implementation plan will consider:

- Progressive integration of RIs
- Testing of the first version of the IT-IOOS integrated system
- Realease of the first IT-IOOS version and start of the operations
- Upgrade of the IT systems and its new ITINERIS components
- Testing of the second of the IT-IOOS integrated system
- Release of the second version of the IT-IOOS
- Operational phase of IT-IOOS

The key points leading its realization are listed below:

ASSESSING THE EXISTING OBSERVING SYSTEMS: M1-12

1. Identification of the state of the art of the Italian RIs and facilities, including the gaps in the EOV, ECV and EBV and the bottlenecks in the data access procedures

DESIGNING: M13-16

2. Design of the IT-IOOS overall architecture, identifying the key sub-systems, their functions, their interfaces and components. (M16) JAN 2024
3. Design of the sub-system architecture, dataflow, data products, data access, and interface among sub-systems. Design the single access-point to data, products and services.
4. Scheduling the incremental integration of the SS and their components, identifying the number of SS to be integrated in the first, second and in the third-final IT_IOOS versions. (M16) FEB 2024

FIRST PHASE IMPLEMENTATION: M17-23

5. IT-IOOS metadata catalogue of data and facilities (M18) APR 2024
6. Development of new procedures to overcome the bottlenecks in data access (M17) MAR 2024
7. Develop new procedures to make data available in Real Time and Near Real Time (M17) MAR 2024
8. Creation of the management committee responsible for the data management and IT-IOOS operations (M18) APR 2024
9. Implementation of the first version of IT-IOOS by integrating a limited number of RIs and/or data and products, already existing and in place. The SS component that will be part of this first version will be identified based on the maturity of the RIs (M19) MAY 2024
10. Test of the first IT-IOOS version and production of test report to assess the capability to provide a single point access to data, product and services (M20) JUNE 2024
11. Release of the first version ofIT-IOOS(M21) JULY 2024

SECOND PHASE IMPLEMENTATION: (M22-26) AUG –DEC 2024

12. Implementation of the ITINERIS Marine Data Store and data management and curation (M22) AUG 2024
13. Implementation of second version ofIT-IOOSby integrating the additional sub-systems on top of the first version of IOOS. (M23) SEP 2024
14. Test of the second version ofIT-IOOSand production of test report to provide data, products and services (M24) OCT 2024
15. Release of the second version ofIT-IOOS(M25) NOV 2024
16. Operating the IT-IOOS system, providing access to users and evaluating the IT-IOOS system perfomances (M26) DEC 2024

THIRD PHASE IMPLEMENTATION: (M27-30) JAN-APR 2025

17. Upgrade of the RIs by installing new instrumentations responding to the digital requirements (M27) JAN 2025
18. Realization of the updates of the system integrating the additional components providing test reports (M28) FEB 2025
19. Enhancement of the capability of the RI marine facilities to transfer data from offshore to inshore (M28) FEB 2025
20. Definition of the management and governance structure ofIT-IOOSon short and long-term after ITINERIS. (M27) JAN 2025

21. Nomination of the Executive Board and Advisory Council, in agreement with WP5 partners and defining the governance forIT-IOOSin agreement with key marine institutions for the management ofIT-IOOSafter the conclusion of ITINERIS (M26) DEC 2024
22. Signing the agreement among the Institutions in charge of the management and governance of the IT-IOOS (M28) FEB 2025
23. Finalizing the third-finalIT-IOOSversion: integrating the last SS toIT-IOOS(M27) JAN 2025
24. Final test of the thirdIT-IOOSversion (M28) FEB 2025
25. Release of the final version ofIT-IOOS(M29) MAR 2025

3.2.2. Long-term strategy and sustainability

Principles

The long-term strategy and sustainability of the IT-IOOS shall ensure the effective operation, continuous improvement, and the achievement of overarching goals. The IT-IOOS designed in this document and implemented by ITINERIS will constitute the prototype of the National ocean observing system that should be remain in operation at least ten years and should be able to include additional components that are presently not part of ITINERIS (eg. EMBRC). The
The governance of IT-IOOS will be based on the following key principles:

Collaborative Framework

Collaboration will be fostered among various stakeholders, including government agencies, research institutions, private sector entities, and non-governmental organizations. Partnerships at regional, national, and international levels will be established in order to share resources, expertise, and data.

Clear Leadership and Coordination

The leadership and the coordination will be assigned to a national board elected/nominated in agreement with the Marine Italian Institutions. The board will be responsible for overall coordination, management, and decision-making. The roles and responsibilities of participating organizations will be clarified in order to avoid duplication of efforts and ensure efficiency.

Long-Term Funding Mechanisms

Sustainable funding models will be developed to ensure continuous financial support for the IT-IOOS. A combination of government funding, private sector partnerships and international collaboration will be explored to diversify funding sources.

Stakeholder Involvement

Representation from diverse stakeholders, including scientists, policymakers, industry representatives, and community members will be included. Advisory committees or boards to solicit input, feedback, and guidance from stakeholders will be established.

Data Management and Standards

Robust data management policies will be implemented, including data sharing agreements, quality control measures, and standardized formats. Adhering to international standards will be promoted to facilitate interoperability and compatibility with other data infrastructures.

Infrastructure Maintenance and Upgrades

A plan for the maintenance, upgrade, and replacement of observation platforms and instruments will be developed to ensure the IT-IOOS remains technologically advanced. The plan will be designed in agreement with the responsible operative units and the Research Infrastructures. Incorporation of emerging technologies will be regularly assessed to enhance data collection capabilities.
Infrastructure Maintenance and Upgrades.

Capacity Building

Training programs and capacity-building initiatives will be established to ensure a skilled workforce capable of operating and maintaining the IT-IOOS. Education and outreach activities will be promoted to raise awareness and understanding of the importance of ocean observations and data management principles.

Adaptive Management

An adaptive management approach will be planned in order to allow for adjustments based on changing environmental conditions, technological advancements, and stakeholder needs. Periodic reviews and evaluations will be conducted to assess the effectiveness of the IT-IOOS and make improvements as needed.

International Collaboration

Engagement with international organizations and initiatives will be promoted to align with global standards, share best practices, and contribute to a more comprehensive understanding of the world's oceans.

Policy Integration

IT-IOOS will be integrated into relevant national and regional policies related to marine environment, oceanography, climate change, disaster management, and sustainable development.
The incorporation of ocean observations in policy decisions at various levels of government will be advocated.

4. DEFINITIONS AND DESIGN CONCEPTS

4.0. Architecture of IT-IOOS

The fundamental organization of a system embodied in its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution. The IT-IOOS is designed as a distribution system with a single access point for the users and for the ITINERIS hub.

The system design shall meet the requirements identified in section 4. The design shall consider the existing and already in place sub-systems.

The architecture of IT-IOOS also considers the possible extension and implementation of a new sub-system to allow the IT-IOOS to expand and update its parts.

4.1. Sub-System

A sub-system (SS) is a part of the system having a generic function. The function allocated to each SS shall consider the specific mission of the existing Research Infrastructures that have been identified by the PNIR and that are part of the ITINERIS Marine Domain. In addition to the main RIs sub-system two further sub-systems, having the role of accessing and archiving the data, have been identified as major part of the IT-IOOS architecture.

Specifically, IT-IOOS system is composed by the following the sub-systems:

1. DANUBIUS
2. eLTER
3. EMSO
4. EuroARGO
5. EuroFleet
6. GEOSCIENCE
7. ICOS
8. JERICO
9. Laura Bassi
10. LNS
11. SIOS
12. Marine Data Store
13. Marine Data Portal

The extensive mission and function of the SSs is defined in paragraphs from 5.5 to 5.18 together with SSs architecture, main components and modules.

4.2. Component

Every SS can be made by one or more components depending on if the sub-system:

- Is already managed by one or more organizations (i.e.: EMSO-INGV, EMSO-OGS);
- Is geographically distributed;
- It is characterized by homogeneous data types which require different data-processing (i.e.: JERICO-HF Radar; JERICO-time-series from buoys/fixed platforms, JERICO-AUVs, etc.)

Each component may be organized into ***Production Units (PU)***, which have the mandate to produce a specific set of data/products and correspond to the final endpoint, where necessary.

4.3. Module

Each sub-system and/or component and/or PU shall have different modules, each one characterized by a specific objective or task, that responds to one or more requirements as defined in the section 4 (i.e.: acquisition data module, data processing module, quality control module, data archive module, etc.)

4.4. Internal and external Interfaces

The interfaces share boundaries between at least two units of the system architecture. The fundamental aspect of an interface is functional, and it is defined as the place where inputs and outputs functions occur.

External interfaces identify the source of data used by the SS for its operations. The external interfaces are not part of ITINERIS Marine Domain and are not part of ITINERIS in general. Each SS shall list the external interfaces required to perform data acquisition (input), processing, or providing access to the SS products (output).

The internal interface are the nodes connecting two or more ITINERIS Marine Domain sub-systems. For example, all the RIs sub-systems will have as internal interface the Marine Data Store and Marine data Portal. In addition to that, if the sub-system requires to be connected to another IT-IOOS sub-system to acquire data or access to processing or quality control procedures the related internal interface shall be identified and listed in the sub-system description paragraph.

In the case of the Marine Data Portal, the preliminary list of internal interfaces is:

- all IT-IOOS sub-systems,
- Marine Data Store,
- ITINERIS Hub.
- External interfaces

4.5. Products and Datasets

IT-IOOS will provide a wide variety of in situ data, remote sensing data and model products. These data include temperature, salinity, currents, sea level, waves and various biogeochemical measurements chlorophyll, backscattering ocean properties, acustic observation, ect. In addition to marine RIs observation, as mentioned in the previous section the Marine data Store will be able to host datasets and products including satellite and model data. The data/products can be time series, profiles, bi-dimentional maps, 3D fields.

All this data will be organized in the ITINERIS catalogue of products and datasets, which has been designed in line with existing standards at the EU level. This will ensure the integration of ITINERIS data into the data flow of RIs and their visibility in the EU Data Infrastructures (EMODnet, SeaDataNet, Copernicus). Through its implementation, the ITINERIS catalogue will ensure the findability of ITINERIS data.

Much of the ocean information collected as part of the ocean observing system has a very high value in Near Real-Time (NRT). These data, and their near real-time dissemination, are critical for ocean and weather prediction, disaster response and climate monitoring, detection, and mitigation, i.e., for the end-to-end met-ocean system (Pinardi et al. 2019). Nevertheless, NRT data cannot reach the quality requested by climate study, so NRT data can be re-processed to provide Delayed Mode (DM) and high quality-controlled ocean data/products. Therefore, the data and products can be classified on the basis of their quality level (L0, L1, L2, ...). Different classification on NRT and DM products exist and are adopted by ocean community and RIs.

In IT-IOOS we adopted the following definition:

- Near Real Time (NRT): in situ, remote sensed and satellite quality-controlled observations, hourly updated and distributed by IT-IOOS within 24-48 hours from acquisition in average. Model forecast products will be included in this category.
- Delayed Mode (DM): in situ, remote sensed and satellite observations, which are subject to more thorough quality-control, including human operation, and which are distributed with a delay Model. Model analysis will be included in this category Reprocessed data: in situ, remote sensed and satellite observations, which are subject to more thorough quality-control and reprocessing, and which are distributed with a delay. These data aim to provide consistent time series for climate application. Model reanalysis is included in this category.

5. SYSTEM REQUIREMENTS

This section identifies the general requirements, the performance requirements, and the service requirements of the IT-IOOS aiming to provide a single access point to all RIs. The evaluation of these requirements will consider the potential users that will access to the IT-IOOS portal.

Below there is a preliminary list of user which will be updated after starting the operations:

- Scientific users (internal and external)
- Governmental institutions
- Public bodies
- Agencies
- Private sector stakeholders
- Accademia (students)

The requirements of the system are defined also considering potential future evolution and upgrade of the IT-IOOS. Also, the possibility to integrate other new sub-systems or data sources is not belonging to the network of the RIs has been considered in designing the IT-IOOS general requirements.

5.1. IOOS General Requirements

5.1.1. *IT-IOOS-SR-01: IT-IOOS definition*

Description: IT-IOOS has to provide single access point to all Italian RIs services, including their ocean data produced, and any other additional sub-system that may require to be connected.

The IT-IOOS system shall be designed as a modular system configurable either at system and sub-system levels in order to allow any upgrade and integration of additional sub-systems without requiring any change of the overall system architecture.

5.1.2. *IT-IOOS-SR-01: IT-IOOS Data*

Description: IT-IOOS data will include all available datasets from:

- Research Infrastructures
- ITINERIS integrated products
- Possible additional sub-systems

5.1.3. *IT-IOOS-SR-01:IT-IOOSFunctions*

Description: IT-IOOS shall provide a common platform for the data acquisition, processing, archiving and distribution within all RIs with the aim to provide interoperable data access and management of related services. IT-IOOS will host at least the following functional requirements further described in section 4:

- Acquiring data
- Processing data
- Performing data check and quality control
- Generating products
- Archiving data and products
- Distributing data and products
- Providing data catalogue
- Availability of Near Real Time (NRT) data and products
- Availability of Delayed Mode (DM) data and products

- Providing the list of the Italian facilities belonging to the RIs
- Access to the RIs facilities
- Users' registration
- Monitoring of user's access
- Monitoring of IT-IOOS services
- State of availability of the IT-IOOS services
- Information on IT-IOOS services (ie. downloading, viewing, etc.)
- Users access to IT-IOOS portal should be informed of the IT-IOOS services offered by IOOS, in particular the users should be informed in case of:
    - unavailability of any service
    - availability of new services
    - modifications of existing services

5.2. IT-IOOS Sub-System Requirements

5.2.1. SSR-01: Provide information on SS data products

Description: Each sub-system shall identify the components and/or production units contributing to the sub-system. Each subsystem shall provide the design of sub-system architecture, together with a graphical scheme, using the provided template. It shall also identify who will coordinate the SS itself.

Each sub-system shall identify the list of EOV, ECV and EBV data products that will be provided by the sub-system. For each product, the following basic information shall be specified:

- Sub-system
- Sub-system component
- Operating Facility
- Dataset
- Data provider
- Variables name list
- Type of dataset (existing or new)
- Data format
- Near Real Time Mode (NRT) availability and related delivery time
- Delayed Mode (DM) availability and related delivery time
- End point
- Type of End point
- Updated dataset

The following table shall be implemented for each sub-system with reference to both present status and planned condition after the ITINERIS implementation is completed.

SS	*SS-Component*	*Operating facility*	*Dataset*	*Data provider*	*Variables name list*	*Existing dataset (E)/New dataset (N)*	*Data format*	*NRT (add "x")*	*DT data (add "x")*	*NRT Delivery Time*	*DT Delivery Time*	*End point*	*Type of End point*	*Updated dataset*

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component or additional data providers

Normal flow: Identification of sub-system components and variable list, together with the above-mentioned metadata information shall be provided to ITINERIS Marine Domain Management Team and to the Marine Data Portal.

5.2.2. SSR-02: Acquire data

Description:IT-IOOSsystem shall acquire data from Italian facilities and/or from identified interfaces (upstream system or RI sub-system).

Who: IT-IOOS sub-systems (RI) or IT-IOOS sub-system components

Preconditioning: upstreamIT-IOOSfacilities and/or IR databases are operational and expected products are available: i.e. raw data from Italian facilities and/or data access through interfaces and/or sub-system catalogue already in place.0

Normal flow: Data acquisition shall be automated. The sub-system level must describe the activities to be performed by the support operator and automated part of the system.

5.2.3. SSR-03: Perform input data checks and Pre-processing

Description: Each sub-system shall perform input data check and pre-processing on the data acquired from the Italian facilities and/or from identified interfaces (upstream system or RI sub-system). Missing data shall be recognized, and warning system and recovery procedures shall be activated. The database of missing data shall be updated and transmitted to the Marine Data Portal. Data after input data check will be pre-processed (if required) to produce input data for QC.

Who: IT-IOOS sub-system (RI) or IT-IOOS sub-system component

Preconditioning: Input Data Check and pre-processing shall be performed on data that has been already acquired by the sub-system.

Normal flow: Data input check and pre-processing shall be automated. The sub-system level must describe the activities to be performed by the support operator and automated part of the system.

5.2.4. SSR-04: Perform QC and Process data

Description: Each sub-system shall perform QC on the data acquired from the Italian facilities and/or from identified interfaces (upstream system or RI sub-system). Data processing shall be performed if required. Data output should include data and information on QC level (quality flag). Data should be provided in one of the format standards recognized at international level and agreed as RIs data formats.

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component

Preconditioning: QC and processing shall be performed on data that have been already acquired by the sub-system.

Normal flow: Data QC and data processing shall be automated. Sub-system level must describe the activities to be performed by support operator and automated part of the system.

5.2.5. SSR-05: Generate products

Description: Each sub-system can generate additional products from original data acquired and quality controlled from the Italian facilities. ITINERIS Marine Domain integrated products (activities 5.18, 5.19, 5.20, 5.21) shall be generated and provided toIT-IOOSMarine Data Store. Products from external sources can be made available to theIT-IOOSMarine Data Store. Data products should be provided in a format standard recognized at international level and agreed asIT-IOOSdata product formats. Proper metadata shall always be provided together with data products.

For in situ observations the minimum metadata information shall include:

- Reference RI (Program Name) RI Start date (Program Start date)Authors, name (Platform Owner name)
- Authors, institute (Platform Owner - EDMO reference)
- Platform class (L06 - SEAVOX PLATFORM CATEGORIES)
- Data provider (Collate Centre - EDMO reference)
- Operating facility (Platform Name)
- Date and Time or time domain (Start date)
- Latitude, longitude or spatial domain (C19 Sea areas and Geographic coverage (bounding box)
- Variables name list included in the file (Parameters, P02 - SEADATANET PARAMETER DISCOVERY VOCABULARY)
- Sensor name
- QC type (Quality control procedures)
- Data type: NRT or DM
- Funding source, Project in EDMERP reference (e.g. PNRR-ITINERIS)
- Creation time
- Version
- Data licence (L08 - Data Access Restriction)
- DOI
- Abstract

For in generated or external products the minimum metadata information shall include:

- Owner
- Authors
- Date and Time or time domain
- Latitude, longitude or spatial domain
- Variable name list included in the file
- QC type
- Data type: NRT or DM
- ReferenceIT-IOOSdataset used for product creation
- Reference external dataset used for product creation
- Founding source (e.g. PNRR-ITINERIS)
- Creation time
- Version

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component or additional data providers

Preconditioning: Data products shall be performed using only QC data that has been already acquired by the sub-system. ITINERIS marine domain integrated products shall use data generated within the ITINERIS project.

Normal flow: Generation of the data products shall be automated where possible. Access to the products generated outside IT-IOOS shall acknowledge the data source. All the products generated by IT-IOOS shall be delivered with associated to proper metadata. Sub-system level must describe the activities to be performed by the support operator and/or automated by the system.

5.2.6. SSR-06: Perform output check and QC

Description: Quality control operations shall be performed also on output files to be stored into the archive to verify the completeness of metadata information and use of standard nomenclature, correctness of the data file format, presence of quality control information, quality check on file integrity and completeness on data transferred to the archive.

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component or additional data providers should perform quality control on output files and provide evidence through the Marine Data Portal before starting the integration. The marine Data Store will check for completeness of information and quality flags.

Normal flow: SS operator shall perform the quality check on data before the integration. Output quality check shall be automated.

5.2.7. SSR-07: Archive data and products

Description: IT-IOOS and/or its sub-systems shall archive all data and products of the IT-IOOS sub-systems and integrated marine data products and, on request, additional marine dataset. Each sub-system shall identify the physical system or systems responsible for the data archive and provide the end point information to the Marine Data Center. The archive shall ensure accessibility of data in real time, the maintenance of the data for at least 10 years and the backup of online data, including previous versions, if any. The Marine Data Store shall be designed to ensure possibility to host sub-system data and/or sub-system data back-up on request.

The Marine Data Store shall archive and provide access for theIT-IOOSintegrated products (5.18, 5.19, 5.20, 5.21). The Marine Data Store, located in Naples, shall provide archive service for marine data on request.

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component or additional data providers shall archive the product datasets. The Marine Data Store shall be designed to ensure the archive of any sub-system data and/or sub-system back-up. The Marine Data Store shall provide archive service for additional marine data on request. The Marine Data Store shall provide a backup service for all products available on IOOS.

Normal flow: Data archive shall be automated.

5.2.8. SSR-08: Interface with Marine Data Store

Description:IT-IOOSsub-systems and/or components or additional data providers shall provide an open protocol interface with data and/or products endpoints to the Marine Data Store, allowing data and/or products retrieval according to the sub-system described procedure.

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component or additional data providers.

Normal flow: Interface with data and/or products endpoint shall be provided and accessible to the Marine Data Store. Catalogue import to the Marine Data Portal shall be enabled.

5.2.9. SSR-09: Update IT-IOOS data and/or product catalogue

Description: AllIT-IOOSsub-systems and/or components or additional data providers should be listed and properly described in the IT-IOOS data and/or product catalogue as well as in the ITINERIS HUB catalogue. Therefore, they are requested to provide metadata as comprehensive as possible on two levels: the facility and datasets.

Who:IT-IOOSsub-system (RI) orIT-IOOSsub-system component or additional data providers.

Preconditioning: Guidelines and forms for metadata compilation will be provided by the Marine Data Portal.

Normal flow: Data and/or product catalogues shall be provided and made accessible on a planned schedule to the Marine Data Portal.

5.3. Marine Data Portal requirements

The Marine Data Portal components will take into account the ITINERIS Hub and IT-IOOS SS Requirements. It will follow the international standards and the guidelines for Open Data.

5.3.1. MDPR-01: Register for IT-IOOS service

Description: Before accessing IT-IOOS services and data, users shall be registered, and a password shall be provided. IT-IOOS system shall install a software registration interface in the Marine Data Portal (ITINERIS or IOOS). IT-IOOS system shall maintain a users' database, including records of the users' access to the services. The minimum information to be included into theIT-IOOSuser database are:

- Name, surname, username
- Password
- User authorization level
- Affiliation
- Access location
- Number of accesses
- Date of the accesses
- Data/services requested
- Data/services downloads

Different profiles and authorizations shall be assigned to internal contributors and external users.

Who: IT-IOOS Marine Data Portal and/or ITINERIS Hub

Preconditioning: IT-IOOS Marine Data Portal (or ITINERIS Hub) shall provide a registration page to allow the user to log into the system. The user shall have requested to be registered.IT-IOOSMarine Data Portal shall be designed to implement Personal Data Protection protocols (GDPR).

Normal flow: User registration and password management shall be automated.

5.3.2. MDPR-02: Search for data/products/services

Description: Users shall be enabled to search for data, products and services without registration. All metadata should be freely available and search for data should be allowed without registration.

Who: All users and internal contributors shall be free to access the IT-IOOS Marine Data Portal.

Preconditioning: Users and contributors shall be already connected to the system, but do not need to be already registered. Availability of the data, products and services. Sub-setting masks shall be enabled for search criteria.

Normal flow: Enable search operations for one or more search criteria. The minimum criteria to be provided for the search are:

- Type of data/products/services
- Geographical area of interest

- Variable name
- Data type (NRT/DM)
- Date or interval time range
- Name of RI (Program Name)
- Authors, name and institute (Platform Owner name and EDMO reference)
- Data provider (Collate Centre - EDMO reference)
- Operating facility (Platform Name)
- Type of facility

5.3.3. MDPR-03: Get Project & product information

Description: Project information, data, products and services information shall be provided for all available data and products or to the selected items after the search tool. The ITINERIS metadata catalogue, with detailed information on allIT-IOOSsub-systems (RI) and/orIT-IOOSsub-system components and/or additional active and/or planned monitoring sites (buoys, facilities, etc.) should be published in the Marine Data Portal.

Who:IT-IOOSMarine Data Portal

Preconditioning: Users shall have access to a list and/or catalogue of all available data and products. Alternatively, users shall have access to the information for the selected items after the search tool.

Normal flow: Enable selection of information on available or searched data, products and services

5.3.4. MDPR-04: Visualize product

Description: Users shall be able to visualize data and products by using dedicated visualization tools. Proper visualization tools shall be implemented on theIT-IOOSMarine Data Portal.

Who:IT-IOOSMarine Data Portal

Preconditioning: Users shall select the data and products to be visualized.

Normal flow: Enable full access and visualization of the selected data and products.

5.3.5. MDPR-05: Get product(s)

Description: Data and/or products provided by theIT-IOOSsystem shall be made available for download after user registration. Sub-setting realizations shall be enabled.

Who:IT-IOOSMarine Data Portal with link to Marine Data Store

Preconditioning: Data and products shall be available in the Marine Data Store and/or in the archive of the sub-system component/production unit. The user shall be registered in order to be enabled for the download of data and products. Sub-setting masks shall be enabled through search criteria.

Normal flow: User shall be enabled to get the products automatically.IT-IOOSMarine Data Portal and Marine Data Store shall manage user requests for download including the realization of specific sub-setting.

5.3.6. MDPR-06: User support

Description: User support queries and questions shall be managed. Tutorials and FAQs shall be provided through the unique contact point of theIT-IOOSMarine Data Portal. Where needed, theIT-IOOSMarine Data Portal may request additional information toIT-IOOSsub-systems to deliver information support to the users.

Who:IT-IOOSMarine Data Portal, also viaIT-IOOSsub-systems

Preconditioning: User shall perform an information request

Normal flow: User shall get the required information support

5.3.7. MDPR-07: Visualize measure and indicators

Description:IT-IOOSshall provide statistical information about the users' access and product search and/or downloads. Metrics for the evaluation of sub-systems functioning and interfaces operational activity shall be defined as indicators and performed at regular schedule to monitor theIT-IOOSservice quality. A preliminary list of key performance indicators comprises of:

- Percentage of data availability at the due time for the single component of each sub-system
- Percentage of data availability at the due time for each sub-system
- Percentage of failure of single component of each sub-system
- Percentage of failure of each sub-system
- Number of registered users
- Number of users access
- Number and type of data/products search
- Volume and type of data/products download

Each indicator shall contain information of the following relevant information, in order to enable data analysis also with reference to different sub-settings (i.e. sub-system, component, facility, variable, data type, etc.)

Who:IT-IOOSMarine Data Portal

Preconditioning: Marine Data Store has performed indicators and metrics analysis.

Normal flow:IT-IOOSMarine Data Portal shall provide access to operability indicators to theIT-IOOSITINERIS Marine Domain manager, Sub-systems manager and ITINERIS Hub coordinator.

5.3.8. MDPR-08: Extract data from Marine Data Store

Description: Perform data and/or products extraction according to users request on the Marine Data Portal

Who: Marine Data Portal

Preconditioning: Request from the user and Marine Data Portal connection to the Marine Data Store.

Normal flow: Manage the request, extract data and/or products from the Marine Data Store archive or IT-IOOS sub-system/component archive and, eventually, from external interfaces (e.g. Argo floats data from Euro-Argo Data Portal).

5.3.9. MDPR-09: Output preparation and delivery

Description: Data package creation and delivery of datasets

Who: Marine Data Portal

Preconditioning:

Normal flow:

5.3.10. MDPR-010: Manage Machine to Machine Connections

Description: this function will allow to interconnect IT-IOOS sub-systems/component endpoints with IT-IOOS Marine Store and Marine Data Portal in order to interconnect all the IT-IOOS SS

Who: Marine Data Portal

Preconditioning: Interface with RI data and/or products endpoint shall be provided and accessible to the Marine Data Portal

5.3.11. MDPR-11: Cyber-security requirements

The Marine Data Portal will be compliant with the cyber security requirements stated by the National Agency – Agenzia per la Cybersicurezza Nazionale for qualified Cloud Services and infrastructures used byPublic sector institutions in Italy. The qualification criteria are listed in attachements of Determinazione ACN 307/2022. [1]

The Marine Data Portal refers to the part A2 of the attachemnt of Determinazione 307/2022

Since the data provided from the Marine Data portal are low critical for the , the Marine Data Portal will accomplish the "Ordinary" level of requirements.
With reference to the *National Framework for Cybersecyrity and data protection* FNCS a special attention on the Identity Management, Authentication and Access Control (PR.AC) requirements that will be managed with the adoption of standard and federated authentication protocols (Oauth2, SAML, LDAP...) that allow the system to rely on d federated identity systems (e.g, SPID, EIDAS, Idem GARR/Edugain, ORCID....) that already include multi factor authentication in most cases

The use of wide adopted standards and open source software will be helpful in detect and quickly correct vulnerabilities with the support of the community and developers.

5.4. Marine Data Store requirements

This section will identify the Marine Data Store requirements allowing provide single point access to the IT-IOOS data/products and services, monitor the IT-IOOS system and its SS and components. This section includes the Interface requirements between the Marine data Store and SS t come estrarre i dati e definire le M2M connections

5.4.1. Storage of ITINERIS Marine data integrated products

Description: products created from existing datasets are part of the ITINERIS project and not

5.4.2. MDSR-01: Dynamic temporary archive creation

Description: Creation of a dynamic archive for the last 30-days requests in order to make data transfer faster/easier for the users. Temporary arrchives are deleted based on time information from the Marine Data Portal request.

Who: Marine Data Store based on download request received from the Marine Data Portal.

Preconditioning: Availability of the requested data/products on the Marine Data Store or on IT-IOOS sub-systems/components archive or through external interfaces.

Normal flow: User download request through the Marine Data Portal. Check on creation date of temporary data/products in the dynamic archive. Delete procedure activation

---

[1] https://www.acn.gov.it/strategia/strategia-cloud-italia/qualificazione-cloud

5.4.3. MDSR-02: Do measure and build indicators

Description: Marine Data Store shall build the archive for all information related to IT-IOOS operability and users/contributors access and related history. Marine Data Store shall perform statistical analysis about the users' access and product search and/or downloads. Metrics for the evaluation of sub-systems functioning and interfaces operational activity shall be implemented to provide a measure for each indicator and performed at regular schedule to monitor the IT-IOOS service quality. A preliminary list of key performance indicators is defined within the Marine Data Portal PR-07 requirement.

Each indicator shall contain the following relevant information, i.e. sub-system, component, facility, variable, data type, etc., in order to enable data analysis also with reference to different sub-settings Who: IT-IOOS Marine Data Store

Preconditioning: Database record of any access and user operation on IT-IOOS. Database record of operability of IT-IOOS and its sub-systems.

Normal flow: IT-IOOS Marine Data Store performs indicators and metrics analysis and provides access to the results to the Marine Data Portal in dedicated dashboards. IT-IOOS Marine Data Store keeps a continuous record of IT-IOOS operability indicators and metrics, users/contributors access and related history.

5.4.4. MDSR-03: Monitor IT-IOOS service

Description: check data flow from all PUs of different subsystems and in case of failure sends alerts to MDS manager and PU contact point. Where possible data gaps should be filled with automatic procedure in case of temporary interruption of data flow.

Who: Marine Data Store

Preconditioning: data flow mapped and described from PU

Normal flow: For each PUs and for each dataset the system regularly check the availability and the health of the endpoint

5.4.5. MDSR-04: Users and contributors management

Description: the Marine Data Store evaluate requests from users and contributors external to the IT-IOOS subsystem to integrate useful datasets in the catalog.Who: Marine Data Store

Preconditioning: compatible data format and license

Normal flow:

5.4.6. MDSR-05: Cyber-security protocols

Description: The Marine Data Store shall be compliant with the cyber security requirements stated by the National Agency – Agenzia per la Cybersicurezza Nazionale for qualified Cloud Services and infrastructures used byPublic sector institutions in Italy.

Who: Marine Data Store

Preconditioning: Implementation of Marine Data Store

Normal flow:

6. IOOS ARCHITECTURE

In this section the IT-IOOS Architecture is described.

6.1. Marine Data Store sub-system

The Marine Data Store is the dynamic archive for all ITINERIS marine data, metadata and products. It is operated by CNR ISMAR and OGS in synergy with the Marine RIs. The Marine Data Store will host storage of ITINERIS Marine data integrated products

CNR ISMAR will oversee its development by providing the civil infrastructures, storage, computing and transmission facilities and by developing the services for archiving, harvesting, sharing and accessing all ITINERIS data.

The system will implement a federated approach, connecting with existing IT-IOOS sub-systems/components. It will also provide storage capacity initially for CNR data, but is designed to meet the needs of all Marine RIs.

The OGS will oversee the development, taking into account the state of the art solutions implemented at EU level and the requirements of the Commissione Oceanografica Italiana (COI) as expressed in the Italian National Oceanographic Data and Information Coordination Committee (IODC).

Figure 1 below provides an overview of the overall architecture of IT-IOOS Marine Data Store and its connections with Marine Data Portal, with IT-IOOS RI sub-systems and the external interfaces. The arrows indicate the functional flow of data from the production units of the sub-systems and external interfaces to the Marine Data Store. The data flow from the Marine data Store to the Marine Data Portal and its interconnections with ITINERIS Data Portal is also displayed.

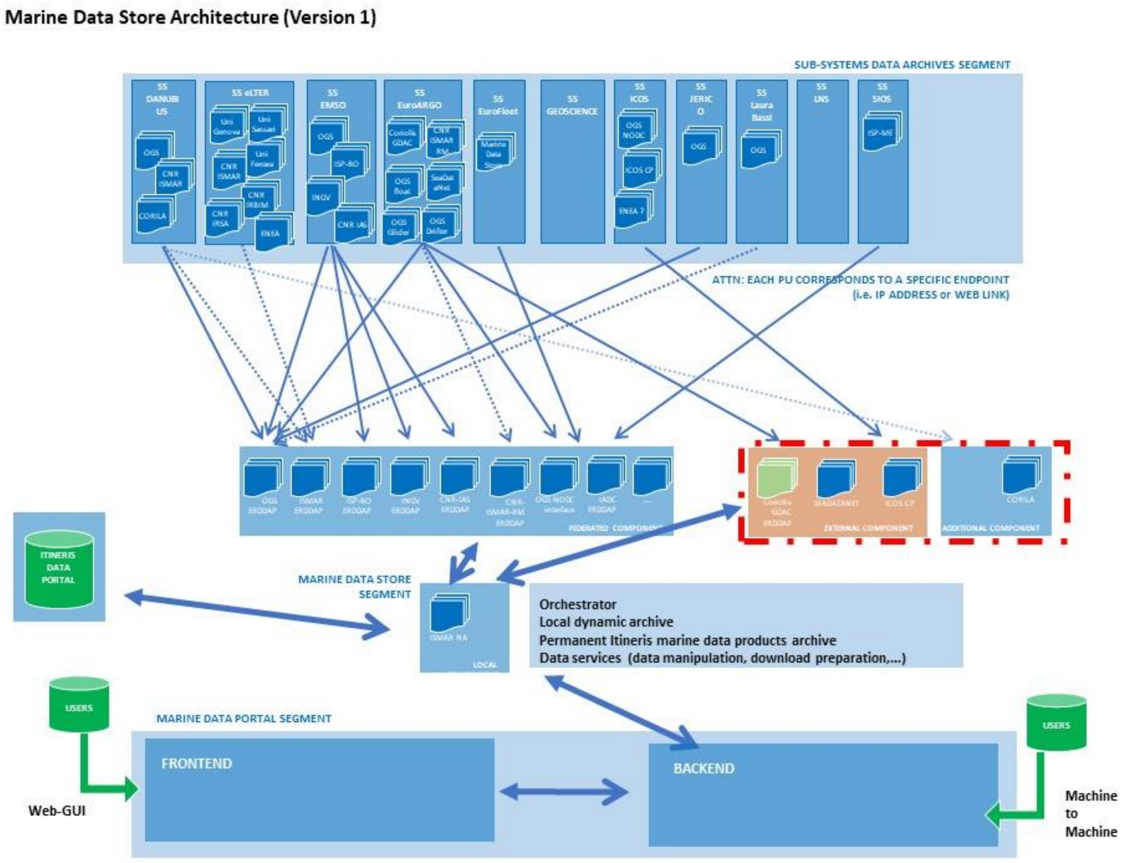

Figure 1. Overall architecture of IT-IOOS Marine Data Store and its connections with Marine Data Portal, with IT-IOOS RI sub-systems and the external interfaces.

6.1.1. Purpose of the sub-system

The purposes of the Marine Data Store are:

- Physically contains data and metadata.
- Connects to a federated system.
- Provides data and metadata to WebGUI through backend operations.
- Catalogs data from other data stores for access and distribution.

6.2. Marine Data Portal sub-system

The Marine Data Portal is the single-entry point of IT-IOOS and the interface between the users and the IT-IOOS sub-systems, including the IT-IOOS Marine Data Store.

Marine Data Portal will be carried out by CNR ISMAR and OGS in synergy with the Marine RIs. CNR ISMAR will oversee the planning and development of the IT-IOOS Marine Data Store, designed as a federated system, and of the IT-IOOS Marine Data Portal, both located in Naples. OGS will oversee the definition and implementation of the ITINERIS catalogue dedicated to 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.

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. The SeaDataNet Common Data Index (CDI) service, giving users a highly detailed insight in the availability and geographical spreading of marine data sets, will be considered.

The ITINERIS Catalogue gathers all metadata describing both NRT and DM data collected by Marine RIs and the other SS that will be part of IT-IOOS. It will be used to develop the Marine Data Portal.

6.2.1. Purpose of the sub-system

The Marine Data Portal sub-system is the first contact to the users. Its main objective is to provide access to IT-IOOS the data and services of IOOS

6.2.2. Local logical/functional architecture

The Marine Data Portal will be composed by the following main components:

- Website
- Web GUI
- Backend (Orchestrator)

Website Component

A first interaction with ITIT-IOOSSystem is done through the Web Portal which must propose a allowing users or contributors to log into the System.

The website will host the following

- Project information [PR-03]
- Main IT-IOOS products and service catalogue [PR-03]
- User registration function [PR-01]
- FAQs. [PR-03]
- Statistics of e users' access and product search and/or downloads [PR-07]

Web GUI (Web Graphical User Interface) component

The Web GUI system will be able to manage user request and present the results. It has the following functions:

- Textual data search. [PR-02]
- Geolocated data search on a map. [PR-02]
- Data preview. [PR-04]
- Data plotting where applicable. [PR-04]
- Data filtering in search. [PR-02]
- Data download [PR-05]

Backend (Orchestrator)

- Manages requests via Web GUI. [PR-01], [PR-02], [PR-05], [IR-05], [IR-03]
- Manages metadata catalogue. [IR-01]
- Provides machine-to-machine data interface. [IR-05]
- Manages user database. [PR-01]
- Interfaces with WebGUI via API. [IR-05]
- Accesses data from the Data Store. [IR-05], [IR-01]

6.3. Sub-system DANUBIUS

6.3.1. *Purpose of the sub-system*

The DANUBIUS sub-system aims at enlarging the research capability to investigate in a multidisciplinary way river-sea systems, through a regular and updated provision of data and products produced by the DANUBIUS-RI Italian facilities. The mission of the DANUBIUS SS is connected to the three main thematic cores of the DANUBIUS SRIA: Water sufficiency, sediment management, ecosystem health.

6.3.2. *Sub-system architecture*

Coordination, SS-components and Production Units

The SS- DANUBIUS is coordinated by CNR-ISMAR and is structured in 2 internal functional components called DANUBIUS-ISMAR and DANUBIUS-OGS and external link, outside the IT-IOOS with DANUBIUS-CORILA component. Below, details of components and production units are reported:

- DANUBIUS-ISMAR. This component collects all four production units belonging to ISMAR, namely *PU ISMAR-VE1*, *PU RM*, *PU ISMAR-TS*, *PU ISMAR-BO*, and manage the facilities AAOT, Turbidimeters VE, SWAMP, S1-GB, PALOMA, Mareografica- TS, METEOMARE TS
- DANUBIUS-OGS. This component consists of one production unit *PU OGS* and manages the facility MAMBO

The diagram in Figure 2. Logical and functional architecture of SS-DANUBIUS resumes the logical/functional architecture of the SS-DANUBIUS.

The list of variables, within each component, with defined end points is provided reported in the annex 1. This list will be a living document that will also be updated with new data from planned ITINERIS equipment. Within SS-DANUBIUS, the catalogue of DANUBIUS-RI provides a common DANUBIUS metadata catalogue for all facilities. SS-DANUBIUS will provide a common metadata catalogue, coordinated among components.

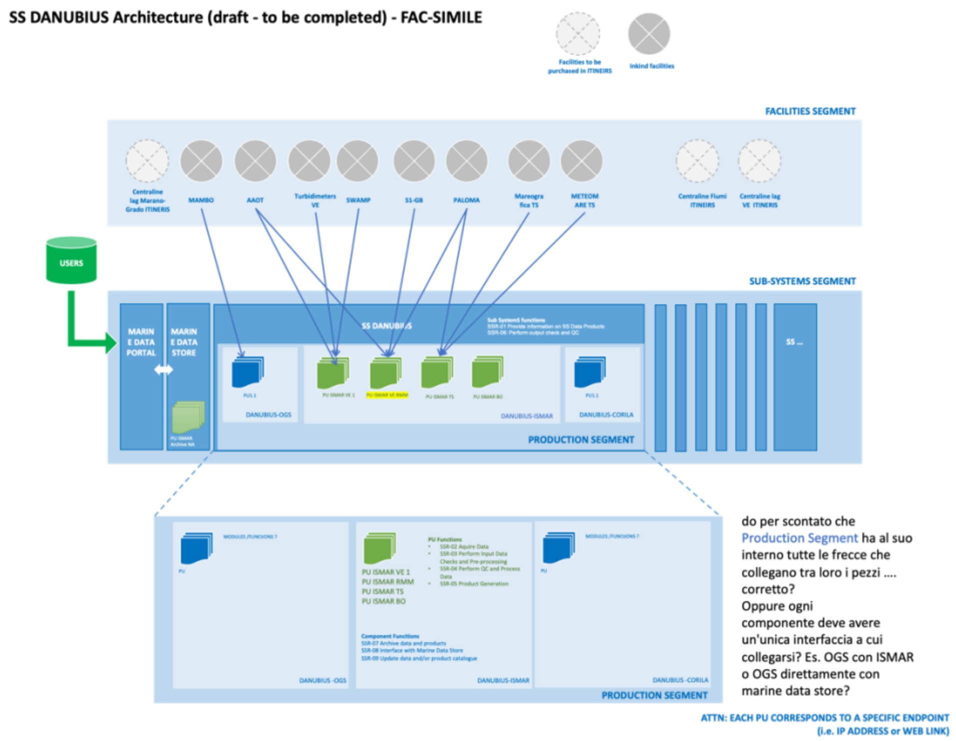

Figure 2. Logical and functional architecture of SS-DANUBIUS

6.3.3. Local physical architecture and functional modules

The local physical architecture of DANUBIUS facilities is described hereafter with details on the data flow from the data acquisition to the data dissemination. The compliance of each SS facility and PU with the sub-system requirements is reported in annex 2.

DANUBIUS-SS

The DANUBIUS SS is in charge with providing the information on SS data products (SSR01). The two components will actively contribute to the update and check of the metadata DANUBIUS catalogue (SSR-09). The SS will collect this information and transmit to the MDS (SSR-09).

DANUBIUS SS will provide a quality check at the DANUBIUS-RI level (SSR-06): given for granted the data quality control, provided within each PU on each variable collected to adhere to the DANUBIUS-RI commons, at the component level the QC will be done on metadata and on the end point persistence.

The two sub-system components will provide the variable listed in the dataset and products file in the annex 1, together with the above-mentioned metadata information defined by the deliverable D5.11, to ITINERIS Marine Domain Management Team and to the Marine Data Portal. The corresponding data are acquired by each of the component's PUs.

The SS-DANUBIIUS will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-DANUBIUS into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the ITINERIS deliverable D5.11. The update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (SSR-09)

In the following sections the detailed description of each DANUBIUS SS Component and PU is provided.

DANUBIUS-ISMAR component

DANUBIUS-ISMAR component manages the data and production archiving (archives SSR07), interfacing with Marine Data Store to provide access to the data (SSR-08) and updating data and/or product catalogue (SSR-09).
In addition to this, the component keeps also a mirror function of archive, relying on a dedicated part of the Marine Data Store SS (embedded in the ISMAR Archive).

ISMAR PUs

The DANUBIUS-ISMAR component has presently 4 Production Units (PU), but the internal organization of data in ISMAR is in an upgrade phase, therefore the PUs will probably change in the next future and also the connected SSRs. The present 4 PUs are:

- PU ISMAR-VE1,
- PU ISMAR-RMM-VE,
- PU ISMAR-TS,
- PU ISMAR-BO,

The 4 Production Units will have SSR from 2 to 5, while the SSR from 7 to 9 are directly connected to the component as a whole. All functions connected to SSR-02 Acquire Data, SSR-03 Perform Input Data checks and pre-processing, SSR-04 Perform QC and process data are functions of each DANUBIUS-ISMAR PUs. Below PUs functional description:

PU ISMAR-VE1

PU ISMAR-VE1 is carrying out the processing of the data from the facilities belonging to CNR-ISMAR, namely AAOT, Turbidimeters VE and SWAMP. In addition to data from facilities, this PU will manage also products, such as Numerical Model Datasets, MSP Tools, etc.

PU RMM-DANUBIUS

PU RMM is carrying out the NRT processing of the data from the facilities AAOT, PALOMA and S1-GB Buoy.

PU ISMAR-BO

PU ISMAR-BO is carrying out the processing of laboratory data mainly of geological data provided by several facilities belonging to CNR-ISMAR and field sampling. The data are in delayed mode.

PU ISMAR-TS

PU ISMAR-TS is carrying out the processing of the data of facilities belonging to CNR-ISMAR for the delayed mode data. The facilities are:

1. PALOMA Platform,
2. RETE METEOMARINA TS,
3. RETE MAREOGRAFICA TS

Data flow

DANUBIUS-SS is a RI under development. Therefore the data flow is still in progress and not completely set up and validated for all type of data. The specific modules and procedures will be detailed when the instrumentation will be fully available and operative. Below the general description of the SSRs and main modules and procedures applied to the data is presented based on the design,:

SSR-02 Acquire Data:

Within the PU, raw data in delayed and real-time mode recorded by the automatic sensors and in situ sampling procedures are acquired. Data acquisition are independent from the manufacturer software to avoid lock-in. In case that the sensor's data are available only into the manufacturer cloud there will be a software component inside PU infrastructure that will collect data through API.
Automatically data are acquired through three types of links:
Direct link to internet (e.g. with IRIDIUM network)
Data are sent direct to internet on the manufacturer cloud system and made available on the web
Sensor is direct connected to PU LAN manly through a converter from serial to TCP/IP and using a management software installed on a VM, a special case of this connection is the LoRa Network that we plan to extend in the future (don't need aa converter for serial protocol)
In case of fault of the network link or the acquisition of data is possible with direct access to the sensor (memory card or connection to management interface).
Data in delayed mode are recorded and stored within the sensors, and after the sensor retrieval, they are downloaded by the operator. Real-time data, after measurements, is collected in a datalogger that handles file storage and transfer or directly sent via connections (e.g GSM, radio, cellular network, satellite) to the shore. Data obtained by analysis of samples, usually by laboratory instruments, are stored in spreadsheet files (excel or separated value).

SSR-03 Perform input data checks and pre-processing.
Within the PU, for NRT data a module of quality check will be implemented in order to automatically verify the arrival of the data at the due time from the facility to the shore. This module will also check the integrity and completeness of the file and its content. In case of errors, an alarm will be sent to the operator of the PU and to the MDP. After data transfer, both NRT and delayed mode data are checked and pre-processed. Sensors management (configuration, power supply, calibration etc..) is performed using the software and interface provided by each manufacturer or reseller. These first steps consist of an initial gross check of the data, visual inspection, and graphing of the data to assess the reliability of the measurements. These operations usually are performed by humans, only few operations can be executed by scripts or algorithms. Important at this stage is the creation of the metadata of the dataset, which also includes the method or instrumentation by which it was acquired, and the labelling of the variables and units of measurement by controlled vocabularies. All information on how the data check and pre-processing was performed should be stored in a file containing information on provenance.

SSR-04 Perform QC and Process data
Within the PU, data are processed and subjected to QC procedures. Each PU will perform QC and processing depending on the measured variable and how it is acquired (NRT, DT, and sample base). All information on how the QC and processing is performed should be stored in a file containing information on provenance and be part of the metadata info, because each typology of data can have a specific QC procedure or algorithm applied.

SSR-05 Generate products:
Within the PU there is the production of derived products from data, data from field campaign and cruises, the production of numerical model products and any other product derived from geographical management tools.
In case of data collected from field campaign and cruises are downloaded by the operator, processed and evaluated, then stored in a NFS network share of the PU. The network share could be accessed from other PU of the same component. Data that are produced in laboratory activity are also saved in a NFS network share and make available to the publishing systems. GIS Products like raster and vector maps or numerical models products are also saved in a NFS network share and make available to the publishing systems and to the Marine Data Store.

SSR-06: Perform output check and QC

The output files will be quality checked before to be stored in the archive and provided to the MDS. Before the integration of the single product into the IT-IOOS a verification of the use of the standard nomenclature and the correctness of the data file format will be carried out and evidence of the compliance will be notified to the MDS. After integration a quality check will be carried out automatically to all the files to verify the completeness of metadata information and correctness of the data file format, presence of quality control information, quality check on file integrity and completeness on data transferred to the archive. In case of errors an alarm will be sent to the PU operator, MDP and the coordinator of the SS-DANUBIUS.

SSR-07: Archive data and products

The PU will transfer/provide access to the data and products to the Marine Data Store and the ISMAR archive that will act as an endpoint for the component interfaced with the Marine Data Store. The physical interfaces of the sub-system are several web api through LAN or VPN or publicly avialable on the Internet. There will be one metadata catalog with the resource to web interfaces.

DANUBIUS-OGS component

The DANUBIUS-OGS component consists in a Production Unit (PU) supported by an existing facility MAMBO (PU OGS-MAMBO) and new planned facilities.

The OGS Production Unit can provide SSR from 2 to 6. About SSR from 7 to 9, they will be implemented according to IT-IOOS requirements and will be made available at OGS-NODC end point in communication with the Marine Data Store. In addition to data from facilities, there are also products, such as Numerical Model Datasets.

PU-OGS-MAMBO

The PU OGS-MAMBO is carrying out the processing of the data from the facility MAMBO-Miramare Buoy, that is part of the Gulf of Trieste Observing Platform. This facility operates in near real-time. Data collected by sensors are sent, through GSM Modem, to the National Oceanographic Data Center (NODC) at OGS, where they are stored and makes available through an ERDDAPP Server. Access to the ERDDAPP server and thus to the data is free and accessible to anyone who requests it.

As mentioned, above, the PU OGS-MAMBO, through the NODC facilities can provide the following services:
SSR-02 Acquire Data. This facility operates in near real-time. Data collected by sensors are sent, through GSM Modem, to the National Oceanographic Data Center (NODC) at OGS.
SSR-03 Perform Input Data Checks and Pre-processing
N/A
SSR-04 Perform QC and Process Data
N/A
SSR-05 Product Generation
N/A
SSR-07- SSR-08. The data managed by PU OGS MAMBO are stored at the National Oceanographic Data Center (NODC) at OGS and made available through an ERDDAPP Server and interfaced with Marine Data Store. Access to the ERDDAPP server and thus to the data is free and accessible to anyone who requests it.

As part of the ITINERIS project, it is planned to expand the observational capabilities of the DANUBIUS-OGS component by laying a new observational network in the Grado-Marano lagoon. The data collected by this new network will also be processed according to a similar data flow.

6.3.4. Computing and archiving facilities

The computing and archiving facility that hosts SS-DANUBIUS data and products is MDS, located at CNR-ISMAR Napoli. Its storage capacity is 750 TB. The archive will ensure real-time accessibility of data and maintenance of data for at least 10 years and backup of online data, including any previous versions.

6.3.5. Physical interfaces of the sub-system (external interfaces)

SS-DANUBIUS will have an external interface with DANUBIUS-RI data portal, currently in development.

SS-DANUBIUS will have an external interface with DANUBIUS-CORILA component.

6.3.6. Design constraints

Since the full DANUBIUS SS is not built, given the status of DANUBIUS, as RI in implementation, there are not constraints related to already existing SS. However, since the DANUBIUS SS relies on the internal data organization of CNR-ISMAR and OGS, there could be technical constraints due to firewalls, network and operation management.

6.3.7. Relationships between sub-systems (internal interfaces)

Since some facilities are providing data to the different RI SS and components, there are shared PU among RIs. Therefore, the modality of data flow between the PUs and the Data Storage will be performed in a harmonized way for the DANUBIS SS (as an example: DANUBIUS-ISMAR Component and JERICO-ISMAR and eLTER-ISMAR component what concerns data flow from S1GB facility).

The internal interfaces od SS-DANUBIUS will be eLTER, JERICO and MDS.

6.3.8. Overall architecture growth

SS-DANUBIUS has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. Furthermore, the setup and deployment of even high-cost, relatively large, and power-consuming instruments are feasible within this system. In respect to dataset and catalogue, the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM.

The endpoints will be managed by application servers that will allow always the addition of new datasets. Specific care will be devoted on the data size, regarding model data products, since bigger than any other dataset (from GB to TB).

6.3.9. Internal failures analysis

The Internal failures strategy will be in place only after DANUBIUS Operative phase will start. Meanwhile the reference procedure to analyse failures will refer to partners DMPs (CNR and OGS).

6.4. Sub-system eLTER

6.4.1. *Purpose of the sub-system*

The eLTER sub-system aims at expanding the research capability to study long-term ecological changes in marine and transitional ecosystems through a holistic 'whole system' approach, based on the integration of different environmental disciplines, to understand the role and interactions of multiple and complex ecosystem variables. eLTER responds to the challenge of understanding the complex interactions between people and nature over the long term. The mission of eLTER is to facilitate high impact research and catalyse new insights about the compounded impacts of climate change, biodiversity loss, pollution, and unsustainable resource use on a range of European ecosystems and socio-ecological systems, representing the "critical zone" in which we live.

6.4.2. *Sub-system architecture*

Coordination, SS-components and Production Units

The eLTER SS is coordinated by CNR-ISMAR and is composed by eight logical/functional components: eLTER-ISMAR, eLTER-Uni Ferrara, eLTER-IRBIM, eLTER-Uni PM, eLTER-IRSA, eLTER-Uni Sassari, eLTER-Uni Genova, eLTER-ENEA.

The diagram in Figure 3 resumes the logical/functional architecture of the SS-eLTER.

The list of variables, within each component, with defined end points is provided reported in the annex 1. This list will be a living document that will also be updated with new data from planned ITINERIS equipment. Within SS-LTER, the catalogue of eLTER-RI provides a common eLTER metadata catalogue for all facilities.

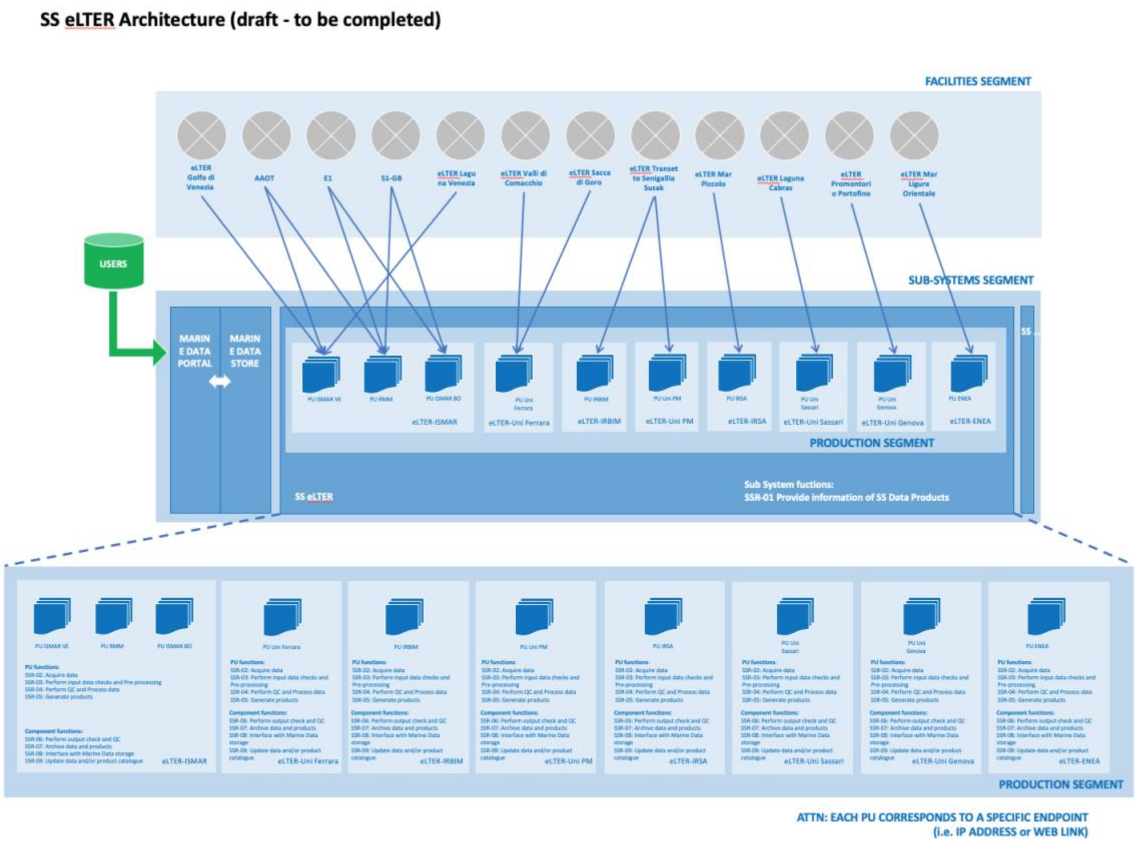

Figure 3. Logical and functional architecture of SS-eLTER

6.4.3. Local physical architecture and functional modules

The local physical architecture of eLTER facilities is described hereafter with details on the data flow from the data acquisition to the data dissemination. The compliance of each SS facility and PU with the sub-system requirements is reported in annex 2.

eLTER SS will centrally provide Information on SS data products (SSR-01). The eight components will actively contribute to Subsystem Requirements, by acquiring data (SSR-02), performing pre-processing (SSR-03) and quality check (SSR-04), generating new products (SSR-05), performing output check and QC (SSR-06), archiving data and products (SSR-07), interfacing with Marine Data storage (SSR-08), updating data and/or product catalogue (SSR-09).

Details of each requirement for the eLTER PUs are given in the description below:

SSR-02: Acquire data

Within the PU, data can flow in near-real time - NRT (sensor based) or in delayed mode (sensor and sample based), depending on the different type of observations and monitoring activities. In detail, the facilities segment of eLTER comprising the 5 fixed-point observing system (AAOT, S1-GB, E1 and TeleSenigallia, MEDA2 Promontorio di Portofino) acquire data in NRT with release time < 2 hours. The other facilities (research sites) acquire data in delayed mode on a number of variables.
Real-time data, after measurements, are collected in a datalogger that handles file storage and transfer or directly sent via connections. (e.g. GSM, radio, cellular network, satellite, LoRa network) to the shore, to the servers located in PUs RMM, IRBIM and Uni Genova. Data are stored in a local geodatabase and exposed using dedicated web services (i.e.: Marine Coastal Information System – MACISTE for the Promontorio di Portofino). Raw data in delayed mode are recorded and stored within the sensors, and after the sensor retrieval, they are downloaded by the operator. Data obtained by analysis of samples, usually by laboratory instruments (e.g. spectrophotometer, microscope), are stored in spreadsheet files (excel or separated value).

SSR-03: Perform input data checks and Pre-processing

Within the PU, once the data have been transferred, they are checked and pre-processed. These first steps consist of an initial gross check of the data, visual inspection, and graphing of the data to assess the reliability of the measurements. These operations usually are performed by humans, only few operations can be executed by scripts or algorithms. Important at this stage is the creation of the metadata of the dataset, which also includes the method or instrumentation by which it was acquired, and the labelling of the variables and units of measurement by controlled vocabularies. All information on how the data check and pre-processing was performed should be stored in a file containing information on provenance.

SSR-04: Perform QC and Process data

Within the PU, data are processed and subjected to QC procedures. Each PU will perform QC and processing depending on the measured variable and how it is acquired (NRT, D, and sample base). All information on how the QC and processing was performed should be stored in a file containing information on provenance.

SSR-05: Generate products

Within the PU there is provision of derived products from data. All information on how products were derived should be stored in a file containing information on provenance.

SSR-06: Perform output check and QC

Within the component quality control operations will be performed on output files to be stored into the archive to verify the completeness of metadata information and use of standard nomenclature, correctness of the data file format, presence of quality control information, quality check on file integrity and completeness on data transferred to the archive.

SSR-07: Archive data and products

Within the component all data and products will be stored and archived in a Network File System (NFS) network of each PU. The network could be accessed from other PU of the same component. Data acquired by sensors will be archived in Sensor Observation Service (OGC-SOS), while data derived from the analysis of samples and products will be available in spreadsheets (separated value). Data and products could be associated with a PID (Persistent IDentifier). Spatial data are shared and made available through a GeoServer System by using Web Map Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS).
Both sensors and research sites are described by specific metadata schemas provided by eLTER-RI (SensorML and Environmental Monitoring Facilities – EMF, respectively) and the information is stored in the official eLTER-RI catalogue DEIMS-SDR (https://deims.org). The information is also shared as a Linked Open Data (LOD) through a triple store and SPARQL endpoint.
The metadata of the datasets produce by the facilities is stored in the catalogue of eLTER-RI.

SSR-08: Interface with Marine Data Store

A single SOS and data endpoint for data and products through to the Marine Data Store will be provided within the component, enabling retrieval of data and products according to the sub-system described procedure. The information on data availability and the corresponding endpoints is provided in the metadata catalogue. The Marine Data Store will be able to reach data and products by API interfaces.

The SS-eLTER will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-eLTER into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the D5.11. The Update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (**SSR-09**)

6.4.4. Computing and archiving facilities

The computing and archiving facility that hosts SS-eLTER data and products is MDS, located at CNR-ISMAR Napoli. Its storage capacity is 1000 TB. The archive will ensure real-time accessibility of data and maintenance of data for at least 10 years and backup of online data, including any previous versions.

6.4.5. Physical interfaces of the sub-system (external interfaces)

The IT-IOOS SS-eLTER has as external interface the eLTER-RI connected with the central eLTER ERDDAP portal to provide access to Italian data.

To avoid data duplication and mismatches, EMSO ERIC provides a single ERDDAP endpoint for end-users through a distributed/federated network of ERDDAP servers, referencing datasets served from the ERDDAP server of each facility. Thus, the unique data repository where the most up-to-date data files from each EMSO facility are stored is the ERDDAP server of the facility. Such configuration also allows to automatically provide the latest version of the data.

The information on data availability and the corresponding endpoints is provided in the eLTER metadata catalogue. The Marine Data Storage could reach data and products by API interfaces.

6.4.6. Design constraints

Until the full implementation of the eLTER SS, given the status of eLTER, as RI in the preparation phase, there are not constraints related to already existing SS. However, since the eLTER SS relays on the internal data organization of different institutions, there could be specific technical constraints due to firewalls, network specifications and management.

6.4.7. Relationships between sub-systems (internal interfaces)

Since some facilities are providing data to the different RI SS and components, there are shared PU among RIs. Therefore the modality of data flow between the PUs and the Marine Data Store will be performed in an harmonized way for the SS-eLTER. Specifically, the metadata and data flow from the AAOT and S1-GB facilities will be exchanged among eLTER, JERICO and DANUBIUS.

SS-eLTER will have internal interface with the SS-JERICO and SS-DANUBIUS to exchange data from AAOT and S1-GB facilities.

SS-eLTER will have an internal interface with MDS to provide access to the data though the IT-IOOS web portal.

6.4.8. Overall architecture growth

SS-eLTER has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. All facilities have the capability to accommodate extra instrumentation for atmospheric, air-sea interface and oceanographic observations. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. Furthermore, the setup and deployment of even high-cost, relatively large, and power-consuming instruments are feasible within this system. In respect to dataset and catalogue, the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM

In the contest of ITINERIS, the components of the SS-eLTER will be enriched with new measured Essential Variables acquired by new sensors. Specifically, the implementation plan includes:

- Incorporating innovative instrumentation to measure biological variables, such as real-time imaging and Automatic Identification System (AIS) for mapping and monitoring fish and gelatinous macroplankton.
- Improving automated acquisition of biological variables through innovative instruments like phytoplankton/cytometer and Zooscan, enabling near-real-time data collection with high time resolution and broad taxonomic classification.
- Establishing synergies among eLTER sites to ensure comparability between different areas and basins. This involves acquiring instrumentation like cytometers, Guard-1, and Underwater Video Profiler-UVP for observations across multiple sites.
- Integrating instrumentation for acquiring physical-chemical variables in the marine environment, such as Photosynthetically Active Radiation (PAR), solar net radiation, and Turbidity. This integration aims to enhance coordination between oceanographic and biological/ecological research efforts

The physical architecture will be improved and expanded to ensure an adequate storage capacity and transmission efficiency for the new products/variables that will be supplied.

6.4.9. Internal failures analysis

Each component will be responsible for the internal failure prevention, basing on their specific operating procedures and Data Management Plans. eLTER SS will prevent server shutdown by using UPS systems and server redundancy.

6.5. Sub-system EMSO

6.5.1. *Purpose of the sub-system*

EMSO, the European Multidisciplinary Seafloor and water column Observatory, aims to support scientific research in gaining a better understanding of phenomena occurring within and below the oceans through a multidisciplinary approach, and in unravelling the role of these phenomena and their mutual interactions in the broader Earth System.

EMSO consists of a system of regional facilities with underwater monitoring functions, located at key sites around Europe, from the Northeast to the Atlantic, through the Mediterranean, to the Black Sea. The facilities are platforms equipped with multiple sensors, placed along the water column and on the seafloor. The facilities provide continuous time-series of measurements of different biogeochemical and physical observables that address natural hazards, climate change and marine ecosystems.

6.5.2. *Sub-system architecture*

Coordination, SS-components and Production Units

The SS-EMSO is coordinated by INGV and it composed by 3 components:

- EMSO-INGV. This component contains a PU (P3.1-INGV) which uses data from the facility Western Ionian Sea
- EMSO-OGS. This component contains one PU (P1.1-OGS) which uses data from the facility South Adriatic Sea
- EMSO-CNR. This component contains 3 PU (P2.1 CNR-ISP, P2.2 CNR-IAS, P2.3 CNR-IAS) which use, respectively, data from the facilities South Adriatic Sea, GranSea, Western Mediterranean Sea

The list of EOVs, EBVs and ECVs produced by SS-EMSO, their datasets and related present and planned access points are listed in annex 1 at the end of this document.

The diagram in Figure 4 resumes the logical/functional architecture of the EMSO sub-system.

Facilities of the SS

The EMSO regional facilities managed by Italian research organizations are three: the **South Adriatic Sea (SAS)**, the **Western Ionian Sea (WIS)** and the **Western Mediterranean Sea (W1M3A)**.

South Adriatic Sea (SAS):
SAS it is jointly managed by **CNR-ISP** and **OGS** and consists of moorings and a surface buoy for data transmission via satellite; the facilities acquire measurements addressing mass properties, biogeochemical cycles, and cascading in the Southern Adriatic, ecosystem function, especially concerning carbon sequestration dynamics and acidification processes in deep waters;

Western Ionian Sea (WIS):
WIS it is jointly managed by **INGV** and **INFN**; it incorporates a set of seafloor observation modules cabled to a shore station for power supply and real-time data transmission and conventional geophysical, oceanographic and acoustic signals are acquired to address geohazards, ocean noise, deep water circulation in a key site for water mass exchange between the Tyrrhenian Basin and the Ionian Basin;

Western Mediterranean Sea (W1M3A)

The W1M3A is managed by **CNR-IAS** and acquires physical, biogeochemical, wave and underwater sound measurements for assessing climate changes and ocean acidification, and for identifying anthropogenic, geophysical and biological sound sources.

Other facilities are expected to be officially acknowledged in EMSO such as the **GranSea**, managed by **CNR-IAS**, in the shallow water of the south-west Sicily coast and the mooring in the Sicily Channel.

EMSO ERIC provides data to end-users through a single ERDDAP endpoint (https://erddap.emso.eu/erddap/index.html) that federates a network of ERDDAP servers managed at each single EMSO facility. The central EMSO ERDDAP has the capability to reference datasets served from other ERDDAP servers. Such a configuration is based on a simple URL that leads to the desired dataset hosted by the remote ERDDAP server located at the facility. Data files are accessed by means of the remote servers managed by each facility to avoid data duplication and mismatches. Thus, each EMSO facility manages its data repository where the latest data version is stored.

EMSO ERIC provides interoperable data and metadata in a coherent and standardized manner. The distributed datasets follow the **EMSO ERIC Data and Metadata Specifications**, a superset of OceanSITES specifications (http://www.oceansites.org/docs/oceansites_data_format_reference_manual.pdf), compliant with CF standards, with some additional metadata coming from SeaDataNet (https://www.seadatanet.org/), adding meaningful information to the distributed datasets. The metadata specification is written in a way that can be read by humans and processed by machines (software processes). Based on machine-actionable definitions of the EMSO Metadata Specifications, within the EMSO, a metadata report tool was implemented, which assesses the alignment of a dataset with the EMSO Metadata specifications. This tool checks if the metadata of a dataset is aligned with EMSO Metadata Specifications and provides a summary with a harmonization score. Additionally, it also provides tips to improve each of the individual metadata terms. It can check standalone NetCDF files and datasets hosted in ERDDAP services (**SSR-06**).

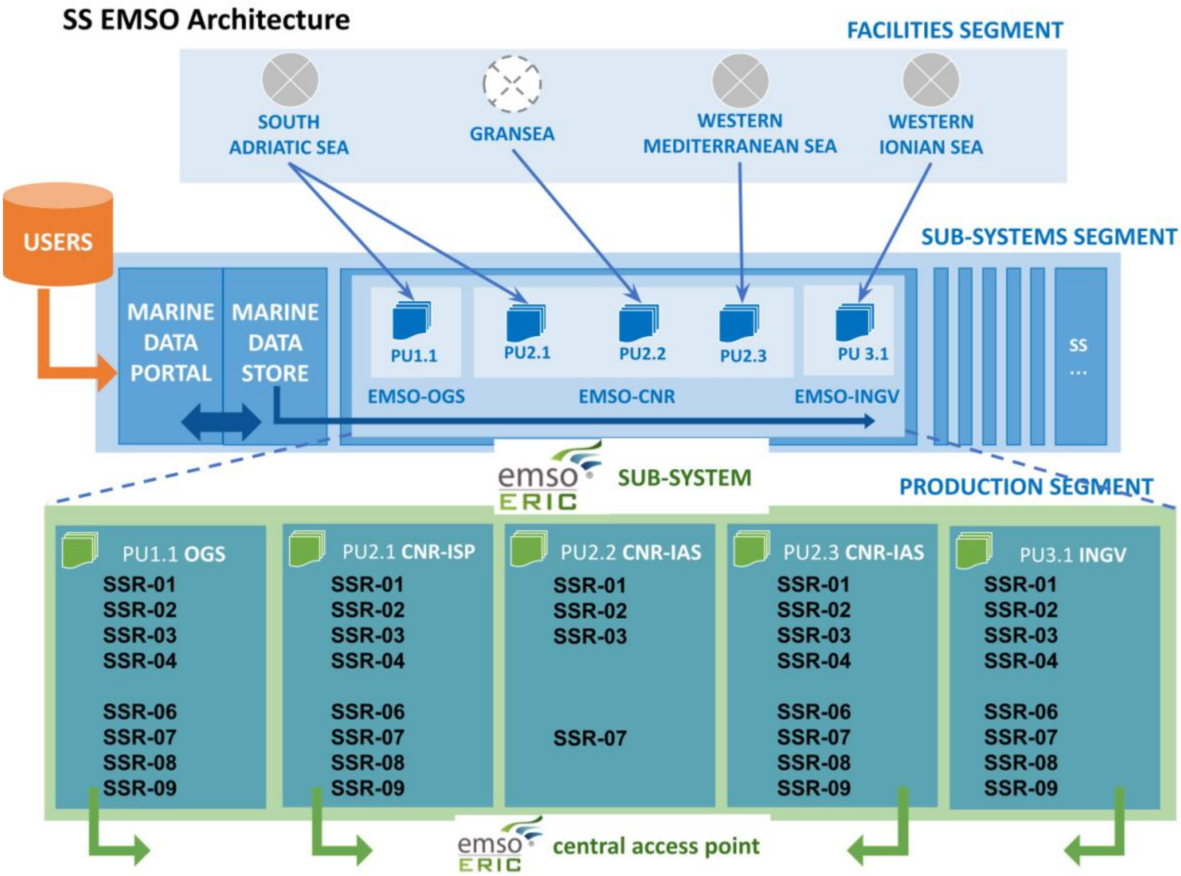

Figure 4. Logical and functional architecture of SS-EMSO

6.5.3. Local physical architecture and functional modules

The local physical architecture of each EMSO facility is described hereafter with details on the data flow from the data acquisition to the data dissemination. The compliance of each SS facility and PU with the sub-system requirements is reported in annex 2.

South Adriatic Sea

The South Adriatic Sea facility has different characteristics in terms of data transmission and reception due to its composition of four production segments:

- E2M3A_B: it is located in the southern Adriatic Pit and consists of a surface buoy and it is currently the only site that transmits data in real-time;
- E2M3A_M: it is located in the southern Adriatic Pit and consists of a mooring transmitting data in delayed mode (DM);
- BB: it is located in the Bari canyon and consists of a mooring transmitting data in DM;
- FF: it is located on the slope and consists of a mooring transmitting data in DM.

The first two segments (E2M3A_B and E2M3A_M) are managed by OGS, while the others (BB and FF) are managed by CNR-ISP.

Data can flow from the instrument to the ground segment in NRT or DM, depending on whether the instrument is connected to the on-board controller or operating in stand-alone mode. Specifically, E2M3A_B can transmit in NRT while E2M3A_M, BB and FF can provide data just in DM.

Data flow for E2M3A_B

SSR2: Acquire data

For E2M3A_B, data are acquired by the sensors and transmitted ashore in NRT. The bidirectional data flow between the offshore site and the onshore infrastructure is powered by the Google Cloud via the Iridium Certus broadband network. The on-board controller developed by the technology unit of the OGS Oceanography department (TEC) handles both NRT data uploads and software updates. In this way, data flows through a robust platform that implements the latest security criteria.

SSR3: Perform input data checks and Pre-processing

Pre-processing and inputs data check is performed by technology unit of the OGS Oceanography department (TEC). Pre-processing usually involves converting the raw data and averaging the recorded samples. The converted data is then formatted as comma-separated ASCII strings in NMEA style. Each string begins with a label uniquely identifying the instrument, followed by the controller timestamp. The formatted strings are stored as daily csv text files in the controller's internal memory and uploaded to the Google Cloud every hour.

SSR-04: Perform QC and Process data

Processing procedures consists in four steps: 1. Messages received by the instruments are decoded and made available in XML format to the personnel involved in the activity via the resources provided by Google Drive and Google App Scripts; 2. Files are converted, provided with additional metadata and entered into a relational PostgreSQL database after visual inspection. 3. Data are transmitted and entered into the database in NRT. 4.Structured Query Language (SQL) procedures are executed on an hourly basis to check the quality of the data according to the OceanSITES guidelines and assign the relative quality flag.

SSR-05: Generate products

The SS-EMSO does not perform additional products. Therefore, this SSR is not satisfied

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

The ground segment of the installation consists of equipment for storing, processing and distributing the received data. NRT Data from E2M3A_B that can be telemetered from the observation system to the ground segment receiving system, are processed by special computers that perform various interaction services such as decoding, storage, quality control and quality assurance, formatting and distribution.

The NRT data of South Adriatic (E2M3A) facility are stored in a PostgreSQL relational database.

SSR-08: Interface with Marine Data Store

The data and the associated quality flags are made immediately publicly available through the ERDDAP service (https://nodc.ogs.it/erddap/info/E2M3A_TS/index.html), following the EMSO ERIC specifications (https://github.com/emso-eric) for data and metadata formatting.

Data accessible through ERDDAP servers are visible on the OGS-NODC geoportal (https://nodc.ogs.it/geoportal) and are also available through the Copernicus in situ-TAC, EMODnet Physics and the metadata are visualised by MonGOOS (Mediterranean Oceanographic Network for the Global Ocean Observing System, http://oceanobs.mongoos.eu/) and JCOMMOPS (https://www.ocean-ops.org/board).

SSR-09: Update IT-IOOS data and/or product catalogue

N/A

Data flow for E2M3A_M, BB and FF moorings

SSR-02: Acquire data

Autonomously collected data, acquired by instruments positioned at E2M3A-M, BB and FF moorings, are provided in Delayed Mode as they are processed after the instruments recovery, which takes place at least once a year: files from these instruments are manually entered into the processing chain on the ground segment, just as they were received via the telemetry system.

SSR-03: Perform input data checks and Pre-processing

Pre-processing is not required for data deriving from E2M3A_M, BB and FF moorings, as the data are made available in DM, after the processing and quality check is carried out.

SSR-04: Perform QC and Process data

Data from E2M3A-M are quality controlled according to OceanSITES best practices. Data Quality Control (QC) is carried out using Matlab routines. Data sets provided in ODV format are given a unique Digital Object Identifier (DOI), which can be accessed via https://nodc.ogs.it/catalogs/doi. The data are then entered into an Oracle database in DM via special scripts, validated (according to Medar Medatlas specifications), standardized (through BODC vocabularies) and redistributed via data infrastructures such as SeaDataNet using standard distribution formats.

Data from sites BB and FF are subjected to QC tests according to SeaDataNet QC procedure (https://www.seadatanet.org/Standards/Data-Quality-Control). After QC, performed using Matlab routines, each data element is assigned a quality flag code according to the L20 SeaDataNet Flagging Scheme. The dataset is prepared in NetCDF format with appropriate metadata according to the SeaDataNet format and EMSO-compliant tests to create EMSO-compliant data archives based on the FAIR principles.

SSR-05: Generate products

The SS-EMSO does not perform additional products. Therefore, this SSR is not satisfied

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

DM data from E2M3A_M are managed following a different data flow, they are available with a data policy defined in accordance with data originator at SeaDataNet infrastructure (https://cdi.seadatanet.org/search), at the OGS-DOI catalogue (https://nodc.ogs.it/catalogs/doi) and relative landing pages.

The datasets for BB and FF moorings are conveniently stored and accessible via the ERDDAP data server, located at https://bo.isp.cnr.it/erddap/index.html, following the EMSO ERIC specifications (https://github.com/emso-eric) for data and metadata formatting.

SSR-08: Interface with Marine Data Store

In order to access to DM data from E2M3A_M, SS-Marine Data Store will be interfaced with SeaDataNet infrastructure (https://cdi.seadatanet.org/search), at the OGS-DOI catalogue (https://nodc.ogs.it/catalogs/doi) and relative landing pages.

In order to access to data from BB and FF moorings, SS-Marine Data Store will be interfaced with ERDDAP data server, located at https://bo.isp.cnr.it/erddap/index.html

SSR-09: Update IT-IOOS data and/or product catalogue

Data flow for Western Mediterranean Sea (W1M3A)

SSR-02: Acquire data:

Data collected from the instruments located on the W1M3A platform have two different types of transfer to the ground segment: manual and automatic, depending on whether the instrument is operating in autonomous mode (flow of data in DM) or it is connected to the onboard data logger (flow of data in NRT). Manual transfer is applied to data collected autonomously, which are processed after the instrument's recovery. Data files from such instruments are entered manually in the processing chain ashore as soon as they are received through the telemetry system. Automatic transfer, instead is applied to data from instruments connected to the onboard data logger. These data are transferred ashore using a satellite (IRIDIUM) and standard Long-Term Evolution (LTE) system. The ground segment of W1M3A consists in equipment for receiving the data (antenna and modem of the IRIDIUM global communication system) and for storing, processing and distributing the received data files. The satellite telemetry system adopts a Stop-and-Wait Automatic Repeat Request (ARQ) protocol with a cyclic redundancy check (CRC, 16 bits wide) to avoid burst errors through the communication channel to upload the data file to the receiving station. When using the satellite system, data transmissions occur as soon as the onboard data logger has a minimum number of files to be transferred to optimize (in time and money) the data transfer. On the contrary, the LTE system uses the standard FTP protocol to upload the data to a server ashore. Independently from the telemetry system used, all data files of a day are usually available at the ground station by the early morning of the successive day after the measurement. The ground segment of W1M3A consists in equipment for receiving the data (antenna and modem of the IRIDIUM global communication system) and for storing, processing and distributing the received data files.

SSR3: Perform input data checks and Pre-processing

The data files automatically transferred ashore are divided into an array of packets and sent sequentially to the ground station, which acknowledges the packets were correctly received and asks for retransmission of corrupted ones. Such protocol ensures that information is not lost due to dropped packets and that packets are received in the correct order.

Data telemetered from the automatic observing system to the receiving system ashore are processed through dedicated computers running different interplaying services for decoding, storing, quality controlling, formatting, and distributing. Data is telemetered in a datalog format, a file type property of National Instruments Corporation in which data are stored as a sequence of identically structured records, similar to a spreadsheet: each row represents a record, each record in a datalog file must have the same data types associated with it and each record is written to the file as a cluster containing the data to store. Datalog files can be processed only using LabVIEW, a graphical programming language developed by National Instruments Corporation.

For the data files manually transferred ashore, the same software is used to program the data logger on board the infrastructure offshore. Several routines devoted to data processing and QC are performed using Matlab. The NRT and DM data streams undergo checks, pre-processing and QC procedures.

SSR-04: Perform QC and Process data

QC is applied to the acquired data following checks defined by the OceanSITES data management team. Quality flags are included in the distributed data products for each parameter (both EOVs and ECVs).

SSR-05: Generate products

The SS-EMSO does not perform additional products. This SSR is not applicable

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

Within the EMSO sub-system architecture EOVs and ECVs datasets by the Western Mediterranean (W1M3A) research facility are provided in NetCDF data files stored in an ERDDAP data server at the following address http://erddap.w1m3a.cnr.it/erddap/index.html on a daily basis.

The following name format is used to publish the data on the ERDDAP server: OS_W1M3A_yyyymmdd_R.nc. “OS” stands for OceanSITES, “W1M3A” is the code of the research facility, “yyyymmdd” is the basic format of the ISO 8601 date format, whereas “R” stands for “near real-time data”. A single data file in NetCDF format contains all data successfully received for the day indicated in the file name.

Data sets are organized by deployment, meaning that bounces of data files are organised and locally stored in separate folders: a new deployment occurs when the configuration of the scientific payload changes (i.e., when even one instrument in the observatory is recovered/replaced or added). Each dataset visible in the ERDDAP server is identified by a title with the following convention: “W1M3A data (yyyymm-yyyymm)”.

Data from the Western Mediterranean research facility are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

The generation and archiving of such data products is totally automatic.

SSR-08: Interface with Marine Data Store

In order to access to data from W1M3A, SS-Marine Data Store will be interfaced with ERDDAP data server at the following address http://erddap.w1m3a.cnr.it/erddap/index.html

Data flow for Western Ionian Sea (WIS)

The WIS facility is located in the central Mediterranean basin, offshore Eastern Sicily Island (Southern Italy). WIS has been developed over the last decades and undergoes continuous enhancements.

An electro-optical cable, approximately 28 km long, connects the land laboratory with the infrastructure on the seabed. At about 20 km east from the coast, the cable splits into two branches, each one long approximately 5 km, that extend respectively to the south with the South Test Site (TSS) managed by INFN and to the north with the North Test Site (TSN) managed by INGV. In TSN, the electro-optical cable terminates in a watertight Cable Termination Frame (CTF), designed to withstand great depths and to operate for long time periods. The CTF provides two submarine ROV-operable electro-optical connectors which allow the connection of two units at the terminal part of the electro-optical cable. A Junction Box (JB) is installed at the TSN subsea frame and can supply power (1kW, 350 VDC) and fibre optic data connection up to a total of 4 independent observatories. Two seafloor multisensor platforms, Calipso and Dione, and a 21 km long instrumented Smart Cable are currently connected to the JB and are operating at the seafloor providing oceanographic, chemical, geophysical, acoustic and seismic data acquired in real-time at the land station.

SSR-02: Acquire data

The ground segment, thanks to the agreements between INFN and INGV, includes two laboratories located in the port of Catania and in Portopalo di Capo Passero. This latter lab hosts the WIS Data Center. The lab in Catania hosts the ground terminal part of the electro-optical cable which supplies energy to the submarine instrumentation and manages the data acquisition system acquired on dedicated servers. The laboratory workstations are used for on-site monitoring of all the observables acquired by the submarine platforms. The acquisition and management servers have an ultra-fast internet connection to transfer the data in almost real-time to the servers, guaranteeing data availability to users. In the same way, the connection makes it possible to monitor the parameters and completely manage the platforms remotely.

SSR-03: Perform input data checks and Pre-processing

The power supply and the control of each single output is performed via dedicated software running on the server of the ground station from which it is possible to monitor the parameters of each single submarine observatory system connected to the output port of the JB. Each connector provides two fibre optic connections and two electrical conductors for each of the four output ports via a ROV-operable.

All data stream timestamps are:

- synchronized with GPS time
- distributed directly to the sea systems for high-rate devices (mainly acoustic and seismic data)
- provided to land acquisition systems for low-frequency data.

SSR-04: Perform QC and Process data

The historical data set undergoes a QC following the QARTOD (Quality Assurance/Quality Control of Real-Time Oceanographic Data) protocols from U.S. IOOS.

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

The Western Ionian Sea datasets are stored and subject to backup in Portopalo Data Center, then inserted in a dedicated data management system (www.moist.it) and identified by a DOI. Data and metadata are standardizsed into NetCDF files following the format conventions and standards compliant with EMSO ERIC Metadata Specifications. Each dataset contains EOVs acquired by NEMO-SN1 seafloor observatory during the period 2002-2013 with missing years.

SSR-08: Interface with Marine Data Store

In order to access to WIS datasets, SS-Marine Data Store will be interfaced the Western Ionian Sea data portal www.moist.it. A subset of datasets is also available through the WIS ERDDAP server (http://erddap.western-ionian.emso.eu:8080/erddap/index.html). An ERDDAP server exposes the historical data and connects with the central EMSO ERDDAP. The same procedure will be provided for NRT data.

Data flow for Gransea

SSR-02: Acquire data

The GranSea facility component allows data acquisition and real-time transmission via sea bottom electro-optic cable. The marine station consists of a power and telemetry module interfaced with sensors by a wet mate panel. The electro-optical cable connecting the underwater station and the land station transfers electrical power from land to sea and provides an optical link for data transmission to and from the land. The electro-optical cable is made up of 500 meters of armored sections, terminated at the ends with two 4 PIN electro-optical connectors.

SSR-03: Perform input data checks and Pre-processing

A ground fault monitor cable insulation control system constantly monitors any ground leakage of the umbilical connection cable between the land station and the sea station. In the event of a loss of insulation of the cable due to water infiltration (breakage of the cable, malfunction of a connector) the system will automatically interrupt the power supply to the cable, as well as signaling the incident via a visual alarm.

A software interface manages the connections to the sensors and the power supply.

SSR-04: Perform QC and Process data

Each sensor data acquisition is managed by specific software, and data is stored in a folder on the servers of the IAS-CNR laboratory in Capo Granitola .

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

N/A

SSR-08: Interface with Marine Data Store

N/A

The SS-EMSO will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-EMSO into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the D5.11. The Update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (**SSR-09**)

6.5.4. Computing and archiving facilities

The data will be stored at the EMSO server which has a larga capability to store data and back-up for a period of 10 years.

6.5.5. Physical interfaces of the sub-system (external interfaces)

The IT-IOOS SS-EMSO has as external interface the EMSO ERIC connected with the central EMSO ERDDAP portal to provide access to Italian data.

To avoid data duplication and mismatches, EMSO ERIC provides a single ERDDAP endpoint for end-users through a distributed/federated network of ERDDAP servers, referencing datasets served from the ERDDAP server of each facility. Thus, the unique data repository where the most up-to-date data files from each EMSO facility are stored is the ERDDAP server of the facility. Such configuration also allows to automatically provide the latest version of the data.

6.5.6. Design constraints

The facilities that will be officially acknowledged in EMSO such as the GranSea, in the shallow water of the south-west Sicily coast and the mooring in the Sicily Channel, will implement a data publication strategy compliant with the EMSO data management principles, together with the relative data and metadata specifications.

6.5.7. Relationships between sub-systems (internal interfaces)

SS-EMSO will have an internal interface with MDS to provide access to the data though the IT-IOOS web portal.

SS-EMSO will exchange information with SS-LNS where acoustic data, routine and QC procedure will be exchanged.

6.5.8. Overall architecture growth

SS-EMSO has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. All three facilities have the capability to accommodate extra instrumentation for atmospheric, air-sea interface and oceanographic observations. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. Furthermore, the setup and deployment of even high-cost, relatively large, and power-consuming instruments are feasible within this system. In respect to dataset and catalogue, the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM.

In the following paragraphs, upgrades to be realized within ITINERIS to every facility are described.

South Adriatic Sea

Within ITINERIS, the deployment of new sensors at the South Adriatic regional site is intended to integrate the observational capacity of biogeochemical and biological parameters (partial CO2, ocean sound and photosynthetically active radiation - PAR) at the ecosystem level and fill the gaps in EOV measurements, especially at the bottom and surface level. The development and installation of a land station will improve communication and real-time data transfer of open ocean observations through the transmission of satellite data. Several parameters that are already measured in DM will be transmitted as RT data via the ERDDAP server.

Western Mediterranean Sea

At the time of writing, the following oceanographic variables are included in the NRT data flow (see annex 1): water temperature and conductivity (6 m, 20 m, 36 m depth), dissolved oxygen and Chl-fluorescence (6 m depth). Additionally, meteorological variables stored in the same data files are the following: atmospheric pressure, horizontal wind speed, wind direction relative to the North, air temperature, relative humidity, short-wave radiation, long-wave radiation, photosynthetically active radiation, and rain accumulation. DM data may include the following variables: water temperature and conductivity (200 m, 500 m, 800 m depth), sea current velocity (100-400 m depth), acoustic pressure (40 m depth). NRT data are available at a 6-hour time interval. It has to be noted that during the project, it is expected to implement the acoustic data transmission of CTD data from two deep instruments. When the acoustic link is established, such data will enter the NRT data flow.

Western Ionian Sea

Within ITINERIS the installation of the new electro-optical cable will enhance the Western Ionian Sea real-time data transfer from the seafloor installations to the ground station extending the operation of the facility for the next 20 years.

For the expected new WIS datasets (see annex 1) the endpoint is not yet available, but a data flow from the sensors to the local publication system will be implemented by the existing data portal (www.moist.it) along with a new ERDAPP endpoint.

6.5.9. Internal failures analysis

South Adriatic Sea

At the South Adriatic Sea facility, although the observatory's on-board telemetry system is duplicated (satellite system and controller memory), there is no way to recover from temporary failures of these systems due to external environmental conditions or harsh events or an interruption of provider services. If no instrument fails, only instruments with an internal data logger and battery operation can temporarily continue data recording if the connection to the on-board controller fails or the on-board controller itself fails. If the on-board controller or satellite system does not fail, the data is regularly uploaded to the Google Cloud platform, which provides a second copy and offers a high level of intermediate security between data collection and subsequent processing and storage in the OGS database.

Due to the remoteness of the site, transmission can only take place via a satellite connection; in the unlikely event of a temporary outage of the service provider, there is no way to ensure NRT data flow unless a second satellite modem from another provider is installed. However, under normal conditions, bi-directional communication allows technicians to work on the remote system if required.

As for the data from stand-alone instruments for all four production segments, there is unfortunately no way to recover the data in case of a system failure.

Western Mediterranean Sea

The process for long-term archiving and back-upping of raw data collected by the Western Mediterranean research facility is based on a multi-level approach. As a general rule of thumb, a copy of all raw data is kept in the internal memory of the instrument or in the memory of the onboard data logger until all the data is secured through multiple and distributed copying. Data received through telemetry are stored three times at the time of receipt: the original data file is kept on the computer connected to the receiving equipment, a first copy is made to a Network-attached storage (NAS) attached to the aforementioned computer, configured with RAID level 1, password protected and with a 256-bit AES volume encryption, and a second copy is made on another computer performing the subsequent data processing phase. Using such configuration there is no latency in the copy of the data.

All three mentioned devices and the telemetry equipment (i.e., satellite modem) are powered through an Uninterruptible Power Supply (UPS) system that protects against power-related issues by providing temporary power during outages and regulating voltage to protect such sensitive equipment. The same UPS also powers and protects the computers running the portal of the research facility and the ERDDAP server, respectively. The overall network of computers can be accessed, managed and controlled remotely using TeamViewer.

Although the telemetry system onboard the observatory is duplicated (satellite and standard LTE system), there is no way to recover from any temporary failure of such systems depending on external environmental harsh conditions or events, nor to interruption of services of the providers (i.e., Iridium Communications Inc., Collecte Localisation Satellites, and Wind Tre S.p.A.).

Western Ionian Sea

The WIS acquisition system is located at the ground station which also provides power supply for all devices. A dedicated UPS system guarantees the continuity of the supply and data acquisition process in case of a main supply short failure. Data communication is guaranteed by one dedicated fibre of the submarine cable and with another one being available as a backup. Data are stored on local hard drives (in RAID1 configuration) and frequently remotely backed up in geographically different locations at the Portopalo data centre in Sicily and Rome servers to minimise data missing events.

6.6. Sub-system Euro-Argo

6.6.1. *Purpose of the sub-system*

Euro-Argo ERIC aims to develop a long-term European contribution to understanding the oceans, their role in the climate system and their health. It manages and coordinates European contributions to support 25% of the international global Argo mission, which is the largest in situ ocean observation system in the world-based on a permanent network of multiparametric profiling buoys (Argo float) that provide a quantitative description of the evolution of ocean state.

6.6.2. *Sub-system architecture*

General information about the Euro-Argo Eric

At European level, the Eıro-Argo Eric processes and distributes the data through a network involving different actors:

- Coriolis Data Assembly Centers (DACs): the data Centres who collect, qualify, process and distribute the float data for which they are responsible. Data are distributed to PIs and the GTS within 24 hours of the float surfacing. They also send the data to the Global Data Centers.
- Global Data Centre (GDAC): the distribution point of Argo data distribution on Internet. It is located in Coriolis/IFREMER/France.
- Delayed Mode Operators: they carry out delayed mode QC and return data to DACs and GDACs within a year from the observations.
- Three Regional Centres led by Euro-Argo partners, check the consistency of the Argo data for a specified geographical area: Atlantic ARC (A-ARC), Mediterranean and Black Seas ARC (Med-ARC) for which Italy (OGS) has taken the lead in establishing the MED-ARC, Southern Ocean ARC (SO-ARC). They also foster the collection of recent CTD data for delayed mode purposes, organize delayed mode QC of floats, provide specific Argo products.
- Argo Information Centre (AIC): located in Toulouse, France, responsible for providing information on the Argo program.
- Argo long term data archive: data Centre located in NODC(/USA) in charge of ensuring the long-term archive of all the Argo data.
- European Argo long term data archive in EMODnet/SeaDataNet: managed by Coriolis country by country.

Coordination, SS-components and Production Units

The SS-Euro-Argo in IT-IOOS is coordinated by OGS and is structured in 2 logical/functional components and an external interface:

- Euro-Argo-OGS. This component contains 3 PUs (PU OGS float, PU OGS drifter, and PU OGS glider) that manage respectively the floats, the drifters and the gliders belonging to OGS;
- Euro-Argo-CNR-ISMAR-RM. This component contains one PU (PU ISMAR-RM) which manages the floats belonging to CNR;
- Euro-Argo-Coriolis-GDAC, which is an external interface managed by IFREMER which manages and processes all Euro-Argo agreed internationally data.

The list of EOVs, EBVs and ECVs produced by SS-Euro-Argo, their datasets and related present and planned access points are listed in annex 1 at the end of this document. The list will be updated if any variation will occur.

The diagram in Figure 5 resumes the logical/functional architecture of the Euro-Argo sub-system.

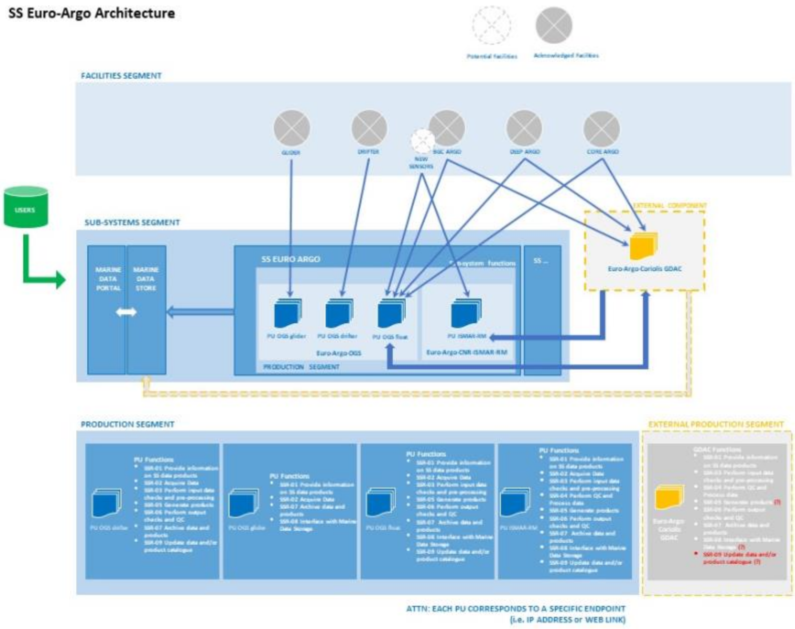

Figure 5. Logical and functional architecture of SS-Euro-Argo

6.6.3. *Local physical architecture and functional modules*

The local physical architecture of each Euro-Argo facilities is described hereafter with details on the data flow from the data acquisition to the data dissemination. The compliance of each SS facility and PU with the sub-system requirements is reported in annex 2.

Role of SS, components and production units and data flow

Briefly, the role and functions of the SS, components and PUs are described below. The data flow from the data acquisition to the data dissemination and the correspondent description of the subsystem requirements are grouped for homogeneous PUs/facilities and instrumentation of the SS.

Euro-Argo Sub-System

The SS-Euro-Argo will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-Euro-Argo into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the ITINERIS deliverable D5.11. The update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (SSR-09)

Euro-Argo OGS component

It will provide information on data products (SSR-01), on data acquisition (SSR-02), archiving of data and products (SSR-07) for data provided by float, glider and drifter facilities. It will interface with the Marine Data Store (SSR-08).

PU OGS float

PU OGS Float will manage the BioGeoChemical-Argo floats embedded with new sensors, Deep Argo and Core Argo. For all observed variables and products not yet included in the GDAC data stream, the PU OGS float will process the data and provide products satisfying for these variables, the additional SSRs 3, 4, 5, 6.

Data flow for Floats (Core, BGC, ONE?)

Both the Euro-Argo OGS components and the PU OGS float are involved in the data flow of OGS floats. Within ITINERIS, OGS float data will be originated by three different type of floats (Core, BGS and ONE) Here below the data flow (and SSRs) of data not included in the catalogue of Euro-Argo ERIC are described in details.

SSR-01: Provide information on SS data products

The information about data and products managed by OGS floats will be provided by the Euro-Argo OGS component.

SSR-02: Acquire data

Data are acquired from the different satellite communication systems (ARGOS or IRIDIUM) that send automatically on scheduling float data to ground station (CLS), and OGS component receives an email which contains all the data packets relating to all the floats that have been transmitted.

SSR-03: Perform input data checks and Pre-processing

The emails received from the satellite communication systems are processed: the text part is removed and the data packets are saved on the server Cayman. The files are then checked and the data packets are inserted into the file that belongs to each individual instrument. A program is started that updates all the data relating to the floats present such as last transmission, last position, number of profiles and status (active/dead). After 12 months from deployment a high level of QC is performed.

SSR-04: Perform QC and Process data

N/A

SSR-05: Generate products; SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

All data and products will be archive in NETcdf format. Euro-Argo data and products included in the Euro-Argo catalogue will be stored in Euro-Argo-Coriolis GDAC server. Data and products not included in the Euro-Argo catalogue, will be stored by the Euro-Argo OGS component in the OGS server in Trieste. The data back-up will be carried out on scheduling. Data will be stored for 10 years.

SSR-08: Interface with Marine Data Store

Two different end points will be available for the OGS Argo data and products:

- Euro-Argo-Coriolis GDAC server. Here, all data from CORE-ARGO Float will be stored. MDS will create an external interface.
- OGS data server (ERDDAP to be created). Here, all data and products from BGS-ARGO sensors will be stored. MDS will create an interface with this internal component.

PU OGS glider

All glider data will be entirely managed and processed by the PU OGS glider. Specifically, this PU will acquire, process, store and provide the data collected by ocean gliders, satisfying SSR 3, 4, 5, 6, 7 for glider data.

Glider data both in NRT and DM are/will be stored at OGS data servers. In NRT, the data are provided also to Ocean Glider program (https://www.oceangliders.org/) and Ocean OPS (International Center of Excellence for Coordination and Monitoring of Meteo-Oceanographic Observing Systems, https://www.ocean-ops.org/DBCP/). Delayed Mode data are tagged with a DOI and stored in the OGS NODC and in EMODnet.
The data back-up will be carried out on scheduling. Data will be stored for 10 years.
Data stored in NODC located at OGS and in EMODnet will be interfaces with MDS with a new ERDDAP data server, specific for glider data that will be set up at OGS within the ITINERIS project. The interface with SS-MDS will occur through this new ERDDAP data server.

PU OGS drifter

The PU OGS drifter will manage all data collected by the OGS drifters, it will process the data and prepare for storing. This PU will satisfy SSR 3, 4, 5, 6 for drifter data.
Data from drifter will be transmitted in delayed mode to OGS server from different satellite communication system. The drifter data after the processing and products generation a DOI will be released and the data set will be provided to NODC that will incorporate them into EMODnet. Data will be stored in the OGS data archive. Data stored in NODC located at OGS and in EMODnet will be interfaces with MDS with a new ERDDAP data server, specific for drifter data that will be set up at OGS within the ITINERIS project. The interface with SS-MDS will occur through this new ERDDAP data server.

Euro-Argo CNR-ISMAR-RM component

It will provide information on data products acquired by the *PU ISMAR-RM* (SSR-01), on data acquisition (SSR-02), archiving of data and products (SSR-07) for data provided by CNR floats. It will interface with the Marine Data Store (SSR-08).

PU ISMAR-RM

PU ISMAR-RM will manage BioGeoChemical-Argo floats. For all observed variables and products not yet included in the GDAC data stream, the *PU ISMAR-RM* will process the data and provide products satisfying for these variables, the additional SSRs 3, 4, 5, 6.

Data flow for Floats (Core, BGC, ONE?)

The PU ISMAR-RM will be the production unit of BGC Argo floats. Here below the data flow (and SSRs) of data not included in the catalogue of Euro-Argo ERIC are described in details.

SSR-02: Acquire data

Data are acquired from the different satellite communication systems (ARGOS or IRIDIUM) that automatically send float data to ground station (CLS), and CNR-IMSAR component receives an email which contains all the data packets relating to all the floats that have been transmitted. The PU ISMAR-RM automatically acquires the data from the CLS ground segment on scheduling. In addition, data already pre-processed by the GDAC are acquired automatically and used to develop new products. Check of file availability is performed before file download.

SSR-03: Perform input data checks and Pre-processing

Downloaded files are checked for file availability and integrity. In case of missing files or errors in downloading process, an alarm is sent to the PU operator. The received data are processed removing the text part and the data packets are saved on the server. The files are then checked and the data packets are inserted into the file that belongs to each individual instrument. A program is started to update all the data related to the present floats such as last transmission, last position, number of profiles and status (active/dead). Parameter conversion and a position control is carried out and quality flag is inserted in case of failure in the position values. For products downloaded from GDAC only the check of data transmission and file integrity is performed.

SSR-04: Perform QC and Process data

Data not included in GDAC, The PU ISMAR-RM performs an automatic QC based on the test range of measured values. Quality flag are assigned in case of outliers. Information on depth are added from the GDAC processed data. Finally, the data are transformed from raw counts to scientific units. Statistics on processed data, missing values and sensor failures are computed and made available to the MDP.

SSR-05: Generate products;

For product not included in the Euro-Argo catalogue, many products will be produced. Specifically, the PU ISMAR-RM float will produce new products (see list in the annex1) using internationally agreed algorithms. For example, density will be computed from temperature and salinity and mixed layer depth will be derived. PAR will be derived from hyperspectral measures. The backscattering will be converted in POC. For every product a test range will be performed and a quality flag will be assigned. All the products will be converted in NETcdf format and the variables will be named following the Climate Format.

SSR-06: Perform output check and QC

The output files will be quality checked before to be stored in the archive and provided to the MDS. Before the integration of the single product into the IT-IOOS a verification of the use of the standard nomenclature and the correctness of the data file format will be carried out and evidence of the compliance will be notified to the MDS. After integration a quality check will be carried out automatically to all the files to verify the completeness of metadata information and correctness of the data file format, presence of quality control information, quality check on file integrity and completeness on data transferred to the archive. In case of errors an alarm will be sent to the PU operator, MDP and the coordinator of the SS-Euro-Argo.

SSR-07: Archive data and products

All data and products from BGC-ARGO will be archived in NETcdf format in the Marine Data Store node located in ISMAR Rome as fully redundant to ensure back-up on scheduling. Data will be stored for 10 years.

SSR-08: Interface with Marine Data Store

The Marine data Store node located in ISMAR-Rome (already existing THREDS) will be interfaced with MDS central system located in Naples.

6.6.4. Computing and archiving facilities

Euro-Argo-OGS component

Float, drifter and glider data acquired by the PU OGS Float, PU OGS Drifter and PU OGS Glider will be managed and stored in a local cluster located in Trieste. Within ITINERIS project, the OGS storage capacity has been expanded with probably a storage Hitachi Vantara Virtual Storage Platform model VSP E590 S/N 611703 with at least 53 Terabyte. Further new disks (10 Tb SASNL 7.2K technology, and volume (configuration RAID6 (6D+2P)) have been acquired to guarantee the long-term sustainability of the data storage of SS-Euro-Argo.

Euro-Argo-CNR-ISMAR-RM component

Float data acquired by the PU ISMAR-RM will be managed and stored in a local cluster located in Rome. It consists of a VMware cluster, composed of 16 blades in 2 Dell PowerEdge MX7000 chassis, each blade with 2 Intel Xeon Silver 4216 CPUs with 16 cores and 32 threads and 256 GB of RAM. 2 of the 16 blades are equipped with 1 nVidia T4 GPU.
The storage part is composed of the following systems:
- Dell Complellent SC5020 SAN dedicated to the VMware cluster (13TB)
- Dell PowerStore 500T SAN dedicated to the VMware cluster (14TB)
- 2 high-capacity NAS (180TB, 250TB)
- Dell PowerScale/Isilon high-performance NAS with 1.2PB of data capacity.

All SAN storage are connected with a redundant 32Gb/s FC connection and all NAS storage are connected with a redundant 10Gb/s Ethernet connection.

This computing and storage capacity will guarantee the storage and the management of data for the next 10 years.

6.6.5. *Physical interfaces of the sub-system (external interfaces)*

The IT-IOOS SS-Euro-Argo has many external interfaces where Euro-Argo data will be exchanged as inputs and outputs. the Euro-Argo ERIC connected with the central Euro-Argo ERDDAP portal to provide access to Italian data and tools for data management.

Euro-Argo-Coriolis GDAC external interface

The Euro-Argo-Coriolis GDAC component that will be an external interface to the IT-IOOS system. MDS will interface with Euro-Argo ERDDAP portal (and FTP) to acquire Italian data and utilize tools for data management. The data and metadata acquired from GDAC will have a common and standardized format following the international Argo data protocols (Argo Data Management Team (2022), Argo user's manual. https://doi.org/10.13155/29825). The data management tools will be utilized through a GDAC central website to process, query, retrieve, plot and compare the profiling float data dynamically.

EMODnet external interface

EMODnet will be an external interface to the IT-IOOS system.

6.6.6. *Design constraints*

The Coriolis data server is fed automatically by all DACs with the latest version of their float profiles and float metadata. A security system is set up on the server to ensure that only the Data center has the necessary privileges to create/modify a NetCDF file for a given profiling float. This ensures that users access a unique source of data. Specific variables will be distributed by each PU.
OGS float component was created in response to the need for data management of sensors that are not yet managed by the Argo international program and thus by the Coriolis-GDAC component and to be able to develop products generated by it and make them FAIR. Glider data storage will be expanded with the real time data servers. Drifter data will be maintained in delay time.
The need of a local CNR-ISMAR-RM component derives from the following points: 1) Argo floats equipped with the newest sensors provide measured variables that are not ingested and thus distributed by the Coriolis-GDAC component; 2) making FAIR a number of ECVs, EOVs, EBVs and other products derived from measured Argo variables whose assessment and distribution are not pursued by the Argo international program and thus by the Coriolis-GDAC component.

6.6.7. *Relationships between sub-systems (internal interfaces)*

SS-Euro-Argo, specifically the two internal components (Euro-Argo CNR-ISMAR-RM and Euro-Argo OGS), will have an internal interface with MDS to provide access to the data though the IT-IOOS web portal.

SS-Euro-Argo will exchange information with SS-LNS where acoustic data, routine and QC procedure will be exchanged.

6.6.8. *Overall architecture growth*

SS-Euro-Argo has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. All facilities and instrumentation, like floats, gliders and drifters have the capability to accommodate extra sensors. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. In respect to dataset and catalogue, the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM.

Within ITINERIS project, the CNR-ISMAR-RM and OGS will be enriched with new measured Essential Variables acquired by new sensors and floats as well as by a number of derived Essential Variables and products (see annex 1 for more details).
The CNR-ISMAR-RM and OGS physical architecture will be improved and expanded to ensure an adequate storage capacity and transmission efficiency for the new products/variables that will be supplied. More details about the upgrades of the storage capacity are reported in section 6.6.4

6.6.1. *Internal failures analysis*

CNR-ISMAR-RM and OGS component will perform the following analysis of failures:

1. Local data store temporal failure;
2. Breaking of measurement sensors, transmission system, and/or floats;
3. Floats end of life caused by battery depletion.

The info, obtained during the QC of the SSR-03, will be transmitted to MDS.

6.7. Sub-system EUROFLEETS

6.7.1. *Purpose of the sub-system*

The EUROFLEETS sub-system within the Italian Integrated Ocean Observing System (IT-IOOS) is designed to facilitate integration and collaboration with the EUROFLEETS infrastructure. The primary purpose is to establish a robust framework for oceanographic research vessel coordination, ensuring interoperability with the broader EUROFLEETS initiative.

6.7.2. *Sub-system architecture*

Coordination, SS-components and Production Units

The EUROFLEET sub-system (SS-EUROFLEET) is coordinated by CNR-ISMAR and structured in three functional components and 4 production units:

- GAIA BLU R/V component is the acquisition system on board of the Gaia Blu research vessel. Several systems are connected to a local infrastructure, which collects NRT data and transmits them through internet.
- EUROFLEET-ISMAR component is the ground facility, located in Napoli (ISMAR-NA) where NRT are received, processed, stored and subsequently disseminated, through web interfaces and as data products.
- EUROFLEET-PRODUCTION component which manage the data of the single production units. This component collects four production units belonging to ISMAR, namely *PU ISMAR VE*, *PU ISMAR RM*, *PU ISMAR TS*, *PU ISMAR BO* and manage the following facilities: Salinometer, Ferrybox, Hyperspectral Radiometers, ROV, and Spectral absorption sensor.

The list of EOVs, EBVs and ECVs produced by SS-EUROFLEET, their datasets and related present and planned access points are listed in annex 1 at the end of this document.

The diagram in Figure 6 resumes the logical/functional architecture of the EUROFLEET sub-system.

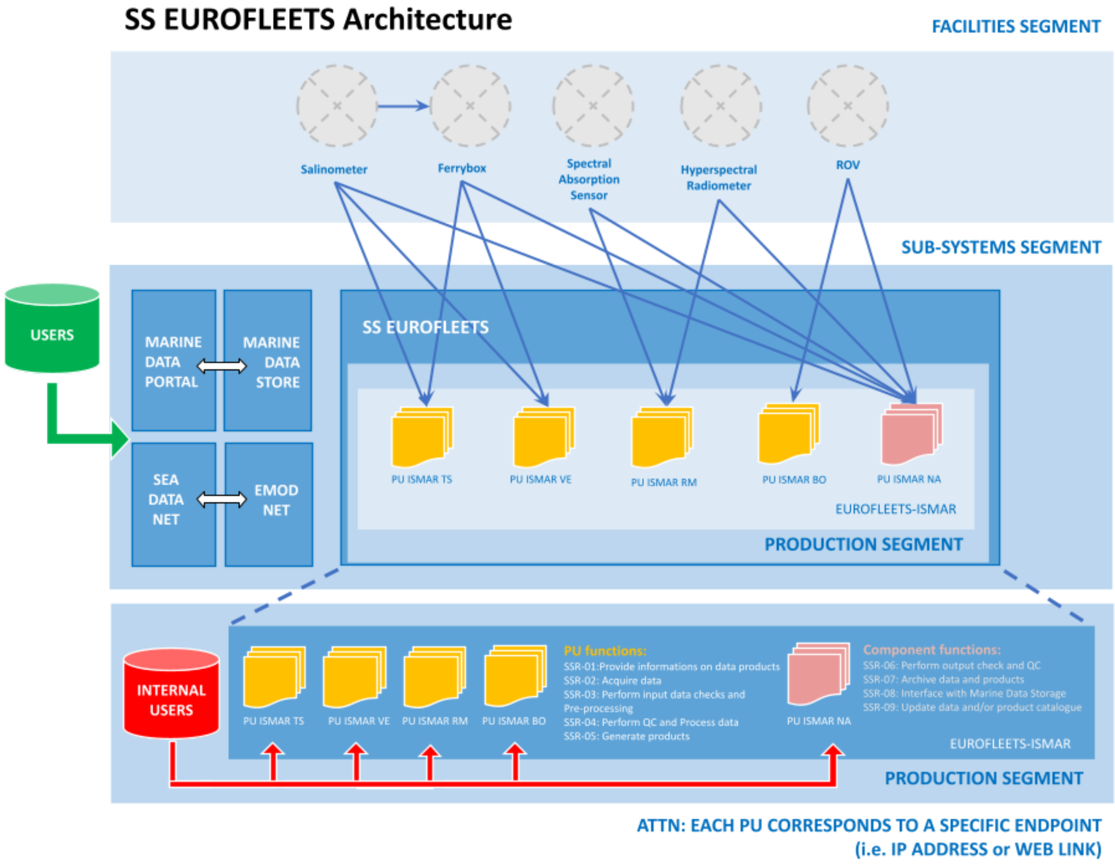

Figure 6. Logical and functional architecture of SS-EUROFLEET

This includes data acquisition, communication protocols, collaborative research planning, and real-time data sharing. The architecture is designed to accommodate the specific requirements and protocols set forth by EUROFLEETS while aligning with the overarching objectives of ITINERIS.

Facilities of the SS-EUROFLEET

The SS-EUROFLEET owns 5 different facilities: Ferrybox, Spectral absorption sensor Hyperspectral Radiometers, ROV, and Salinometer. Below, a brief description of each facility:

Ferrybox:

The Automatic Data Acquisition System (Ferrybox) is a comprehensive solution designed for continuous data acquisition while the RV Gaia Blu is underway. It includes termosalinometer, remote temperature probe, dissolved oxygen sensor, flow-through pH Sensor, flow-through pCO2 Sensor, chlorophyll turbidity (chle turb) sensor and dedicated water lines for optical measurements, additional sensors or seawater sample collection. Ferrybox is jointly managed by PUs ISMAR-TS, ISMAR-VE and ISMAR-NA.

Spectral Absorption Sensor

Spectral Absorption Sensor performs underway measurements of inherent optical properties in marine environments to assess water quality, phytoplankton distribution, and the overall health of marine ecosystems. The system is expected to operate primarily autonomously, enabling measurements in the absence of an operator and transmitting data in near real-time. Spectral Absorption Sensor is jointly managed by the PUs ISMAR-RM and ISMAR-NA.

Hyperspectral Radiometers

Hyperspectral Radiometers provide continuous hyperspectral radiometric measurements above water across a spectral range from 350 to 900 nm. The system is expected to operate primarily autonomously, enabling measurements in the absence of an operator and transmitting data in near real-time. The Hyperspectral Radiometers is jointly managed by PUs ISMAR-RM and ISMAR-NA.

Remotely Operated Vehicle (ROV)
A Remotely Operated Vehicle (ROV) is an underwater robot controlled from Gaia Blu and used to perform tasks in deep underwater environments. It is equipped with cameras, sensors, tools and two manipulators to acquire data and collect geological and biological samples. ROV is jointly managed by the PUs ISMAR-BO and ISMAR-NA.

Salinometer
The high-precision semi-portable salinometer is a designed for offline use and is utilized to enhance data quality and accuracy of the underway salinity readings and to verify salinity data from the Ferrybox system. Salinometer is jointly managed by PUs ISMAR-TS, ISMAR-VE and ISMAR-NA.

6.7.3. Local physical architecture and functional modules

The local physical architecture of each EUROFLEET facilities is described hereafter with details on the data flow from the data acquisition to the data dissemination.

Role of SS, components and production units and data flow

Briefly, the role and functions of the SS, components and PUs are described below. The data flow from the data acquisition to the data dissemination and the correspondent description of the subsystem requirements are grouped for homogeneous PUs/facilities and instrumentation of the SS.

EUROFLEET-ISMAR Component

The component's primary function is to archive data (**SSR-07**) and interface with the Marine Data Store (**SSR-08**) preparing data formats and data and/or products retrieval according to specifications and vocabularies decided by the EUROFLEET sub-system. The EUROFLEET-ISMAR component is also performing output checks and quality control procedures (**SSR-06**) on output files according to the same specifications and vocabularies before being transfer to the EUROFLEET sub-system. The component is also responsible for updating data and/or product catalogue (**SSR-09**) that are provided and made accessible on a planned schedule to the EUROFLEET sub-system first and eventually to the Marine Data Portal. All other functions are relegated to the production unit level for the specific type of facility detailed below.

PUs

The PUs of SS-EUROFLEET (PU-ISMAR-VE, PU ISMAR-BO, PU ISMAR-NA, PU ISMAR-RM) will manage the data flow from the acquisition (SSR-02) to the pre-processing (SSR-03) and quality check procedures (SSR-04). Specifically:

SSR-02: Acquire data

The sensors and associated CTD sensor installed in the Science Store SB onboard R/V Gaia Blu are part of the network hub that will allow the communication of the instrument acquisition units with the ship data center for data storage and data upload to main land servers. For some specific instrumentation, like Ferrybox, the data gathered by the different sensors are automatically integrated with the ship navigation system and eventually with ship's meteorological parameters and delivered in real time to the ship data centre.

SSR-03: Perform input data checks and Pre-processing

Most of the sensors and instrumentation present on board provides binary data that need pre-processing before ingestion into the ship data center. This will be automatically performed by local dedicated computer. Afterwords the acquired and processed data will be transferred to the ship data center.

SSR-04: Perform QC and Process data

Within the PU, data are processed and subjected to QC procedures. Each PU will perform QC and processing depending on the measured variable and how it is acquired (NRT, DM, and sample base). All information on how the QC and processing was performed will be stored in a file containing information on provenance.

The SS-EUROFLEET will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-EUROFLEET into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the ITINERIS deliverable D5.11. The update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (SSR-09)

6.7.4. Computing and archiving facilities

The computing and archiving facility that hosts SS-EUROFLEET data and products is MDS, located at CNR-ISMAR Napoli. Its storage capacity is 1000 TB. The archive will ensure real-time accessibility of data and maintenance of data for at least 10 years and backup of online data, including any previous versions.

6.7.5. Physical interfaces of the sub-system (external interfaces)

The physical interfaces of the EUROFLEETS sub-system within IT-IOOS will be designed to facilitate interaction with the Marine Data Store, a key component of ITINERIS. These interfaces will facilitate the exchange of research plans, observational data, and metadata, fostering a collaborative environment for oceanographic research activities.

6.7.6. Design constraints

While the EUROFLEETS sub-system is designed to maximize interoperability, certain design constraints are acknowledged. These may include technological limitations, divergent vessel capabilities within the EUROFLEETS consortium, and specific data format requirements. Efforts are directed towards mitigating these constraints through ongoing collaboration and adherence to EUROFLEETS standards.

6.7.7. Relationships between sub-systems (internal interfaces)

SS-EUROFLEET will have an internal interface with MDS to provide access to the data though the IT-IOOS web portal.

SS-EUROFLEET will exchange information with SS-LNS where acoustic data, routine and QC procedure will be exchanged.

6.7.8. Overall architecture growth

Anticipating the dynamic nature of oceanographic research, the EUROFLEETS sub-system is designed to accommodate growth and evolution. The architecture incorporates scalability features, allowing for the integration of new technologies, methodologies, adjustments in data size, additions of variables, sensors, and other elements as EUROFLEETS advances.

6.7.9. Internal failures analysis

Conducting a thorough examination of potential internal failures, the EUROFLEETS sub-system prioritizes resilience. It will incorporate redundancy measures, failover protocols, and continuous monitoring mechanisms to effectively mitigate the impact of unforeseen issues.

6.8. Sub-system ICOS

6.8.1. *Purpose of the sub-system*

ICOS mission is to produce standardized, high-precision and long-term observations and facilitate research to understand the carbon cycle and to provide necessary information on greenhouse gases. ICOS Ocean Thematic Center (OTC), currently coordinates 22 ocean stations in seven countries, monitoring carbon uptake and fluxes in the North Atlantic and the Nordic, Baltic and Mediterranean Seas. ICOS data helps to give an account of the Earth system and its response to climate change and other environmental challenges.

6.8.2. *Sub-system architecture*

Coordination, SS-components and Production Units

The ICOS RI Eric is coordinated at national level by CNR. Within IT-IOOS, the SS-ICOS is composed by 3 components:

- ICOS-OGS. This component contains one PU (OGS) which uses data from the facilities E2M3A and Miramare
- ICOS-CNR. This component contains two PUs, ISMAR-TS and CNR-IAS which uses data from the facilities PALOMA and W1M3A, respectively
- ICOS-ENEA. This component contains one PU - ENEA which used data from the facility Lampadusa

The diagram in Figure 7 resumes the logical/functional architecture of the ICOS sub-system.

The list of EOVs, EBVs and ECVs produced by SS-ICOS, their datasets and related present and planned access points are listed in annex 1 at the end of this document.

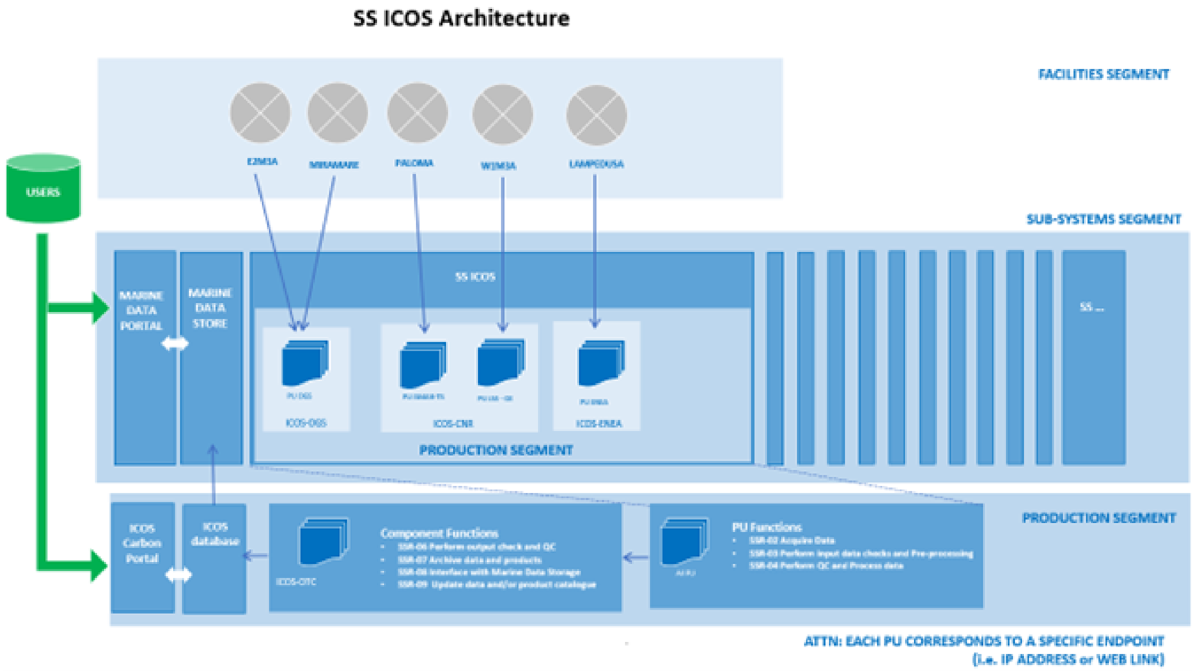

Figure 7. Logical and functional architecture of SS-ICOS

6.8.3. Local physical architecture and functional modules

The local physical architecture of each ICOS facilities is described hereafter with details on the data flow from the data acquisition to the data dissemination. The compliance of each SS facility and PU with the sub-system requirements is reported in annex 2.

Within SS-ICOS, all 5 PUs actively contribute to Sub-System Requirements. They all contribute separately to acquiring data (SSR-02), performing input data checks and pre-processing (SSR-03), perform QC and process data (SSR-04), checking output data (SSR-06), archiving data and products (SSR-07) and provide the interface with Marine Data Storage and ICOS Carbon Portal (SSR-08).

All PUs have a similar data flow and satisfies the SSRs in a similar manner. Details of each requirement and data flow are given in the description below.

SSR-02: Acquire data

ICOS stations acquire data both in NRT and in Delayed Mode. NRT data collected from the ocean surface are transferred to a dedicated hosting service (CLS – ground stations) that store them in its processing center through the different communication systems (i.e., satellite or GSM modem). DM data are downloaded from the instruments during periodic maintenance.
Specıfically, data from Miramare and E2M3A are transferred to the ground station both in NRT and DM.
Data from W1M3A are transferred both in NRT and in DM to the ground segment of W1M3A observatory which is equipped for receiving data (antenna and modem of the IRIDIUM global communication system) and for storing, processing and distributing the received data files. Data flow from the instrument to shore in Delayed Mode.
For PALOMA, (PU-ISMAR-TS) data are both transmitted in NRT to the server dedicated at the PALOMA station in Trieste and periodically downloaded directly from the instruments.
Data from the Lampedusa station are transmitted to the ground segment of the Lampedusa Oceanographic Observatory which is equipped with a NRT data transmission system, delivering data daily to the ENEA data center.

SSR-03: Perform input data checks and Pre-processing

Within the PU, once the data have been transferred, they are checked and pre-processed. These first steps consist of an initial gross check of the data, visual inspection, and graphing of the data to assess the reliability of the measurements. These operations usually are performed by humans, only few operations can be executed by scripts or algorithms. Important at this stage is the creation of the metadata of the dataset, which also includes the method or instrumentation by which it was acquired, and the labelling of the variables and units of measurement by controlled vocabularies. All information on how the data check and pre-processing was performed should be stored in a file containing information on provenance.

SSR-04: Perform QC and Process data

Data post processing according to ICOS protocols can be performed only after a "post deployment calibration" of the pCO2 sensor that is performed once a year and allows the QC data release. All data produced by stations in the ICOS Ocean Thematic Center (OTC) network are processed using a central online tool called QuinCe. Station PIs upload raw data files to QuinCe directly from their instruments, where the measurements are extracted and processed using algorithms approved by the scientific community. QuinCe perform automatic quality control routines to find the most obvious problems in uploaded data (e.g. values outside reasonable ranges, spike detection etc.) and flag these. The station PI review the results, along with the rest of the data, to complete the quality control process and prepare the data for publication at the Carbon Portal. Once the PI has completed quality control, an OTC expert will perform a final check on the data. Once this is complete, the data are published at the Carbon Portal, as well as being passed on to other projects and data centers as appropriate. Yearly ICOS releases integrated data products (level 3 data) and maps for the European region, thanks to the integration of satellite data and model outputs

SSR-05: Generate products

N/A

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

Every PU archives the acquired data and products in different server and computing stores. Specifically, for OGS (stations of Miramare and E2M3A), the dedicated hosting service to archive data is located at the OGS-NODC data center, where the received data are sent to two different facilities. The first is the ICOS data center, in this case the data are formatted and completed by metadata as required by the ICOS protocol. The second is an ERDDAP server located at the OGS-NODC data center, this mode of data access is free and accessible to anyone who requests it.

Data from W1M3A station are provided in netCDF data files stored in an ERDDAP data server at the following address http://erddap.w1m3a.cnr.it/erddap/index.html. Data will be available in delayed mode after the recovery of the deployed instruments. The following name format will be used to publish the data on the ERDDAP server: OS_ITFOSW1M3A_yyyymmdd_yyyymmdd.nc. "OS" stands for OceanSITES, "ITFOSW1M3A" is the code of the research facility, "yyyymmdd" is the basic format of the ISO 8601 date format, the first set for the starting date of the deployment and the second set for the last day of deployment. A single data file in netCDF format contains all data successfully received for one each single deployment. Data from the Western Mediterranean research facility is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). The generation and the archiving of such data products is an automatic process (SSR-05, SSR-07), provided that an operator upload manually the original data files.

Parameters of the W1M3A station will be provided in UNIDATA netCDF files compliant with the netCDF Oceansites, Climate and Forecast (CF) metadata conventions. Metadata stored in the netCDF data files are fully compliant with the L22 vocabulary (https://vocab.nerc.ac.uk/collection/L22/current/) managed by the combined SeaDataNet and MarineXML Vocabulary Content Governance Group (SeaVoX). The usage of such sustained and standardised and hierarchically-organised vocabulary guarantees the scalability of the data management process.

Data from Lampedusa station are archived daily to the ENEA data center located following the NETcdf format.

SSR-08: Interface with Marine Data Store

The data are made available through the Surface Ocean CO_2 Atlas, that brings together all publicly available surface water data from the global oceans and seas. The data provided by the Italian stations will be made available to the ITINERIS data center and data portal through an interface with ICOS Carbon Portal (https://www.icos-cp.eu/data-services) which stores the final quality checked data and products.

Data will be made available and distributed with an open access through the Lampedusa web site and through the ICOS Carbon Portal.

The data acquired with the implementation of the ICOS stations with additional sensors with respect to the core ICOS data will be made available directly through the OGS, CNR and ENEA Data centers to the ITINERIS data center and data portal.
For OGS those data are available through an ERDDAP server, located at the OGS-NODC data center and through the NODC GeoPortal, those mode of data access are free and accessible to anyone who requests it.

The SS-ICOS will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-ICOS into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the ITINERIS deliverable D5.11. The update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (**SSR-09**)

6.8.4. Computing and archiving facilities

The computing and archiving facility that hosts SS-DANUBIUS data and products is MDS, located at CNR-ISMAR Napoli. Its storage capacity is 750 TB. The archive will ensure real-time accessibility of data and maintenance of data for at least 10 years and backup of online data, including any previous versions.

6.8.5. Physical interfaces of the sub-system (external interfaces)

The IT-IOOS SS-ICOS has as external interface the ICOS ERIC connected with the central ICOS Carbon Portal to provide access to Italian data.

To avoid data duplication and mismatches, ICOS ERIC provides a single ERDDAP endpoint for end-users through a distributed/federated network of ERDDAP servers, referencing datasets served from the ERDDAP server of each facility. Thus, the unique data repository where the most up-to-date data files from each ICOS facility are stored is the ERDDAP server of the facility. Such configuration also allows to automatically provide the latest version of the data.

6.8.6. Design constraints

ICOS data flow to Ocean Thematic Center, where the quality control is performed, and finally are made freely available through the ICOC Carbon Portal.

6.8.7. Relationships between sub-systems (internal interfaces)

All data that is part of the ICOS network are released and made publicly and freely available. Yearly ICOS releases integrated data products (level 3 data) and maps for the European region, thanks to the integration of satellite data and model outputs

6.8.8. Overall architecture growth

SS-ICOS has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. All five facilities have the capability to accommodate extra instrumentation for atmospheric, air-sea interface and oceanographic observations. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. Furthermore, the setup and deployment of even high-cost, relatively large, and power-consuming instruments are feasible within this system.

In respect to dataset and catalogue, the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM.

6.8.9. Internal failures analysis

It is managed at European level

The process for long-term archiving and back-upping of raw data collected by the IT-FOS-W1M3A station is based on a multi-level approach. As a general rule of thumb, a copy of all raw data is kept in the internal memory of the instrument or in the memory of the onboard data logger until all the data is secured through multiple and distributed copying. Data recovered from the deployed instruments are stored three times at the time of availability: the original data file is kept on the computer connected to the recovered instrument, a first copy is made to a Network-attached storage (NAS) attached to the aforementioned computer, configured with RAID level 1, password protected and with a 256-bit AES volume encryption, and a second copy is made on another computer performing the subsequent data processing phase. Using such configuration there is no latency in the copy of the data.

All three mentioned devices are powered through an Uninterruptible Power Supply (UPS) system that protects against power-related issues by providing temporary power during outages and regulating voltage to protect such sensitive equipment. The same UPS also powers and protects the computers running the portal of the research facility and the ERDDAP server, respectively. The overall network of computers can be accessed, managed and controlled remotely using Teamviewer.

There is no way to recover from any instrument failure.

6.9. Sub-system JERICO

6.9.1. *Purpose of the sub-system*

The objective of JERICO is to formally design, implement and operate a pan-European Research Infrastructure, dedicated to set up a system of observation and related services for European coastal seas, thus empowering European research excellence and expertise for the benefit of society. JERICO aims at closing the critical gap of not having an integrated European Research Infrastructure addressing the complexity of marine coastal systems.

The JERICO sub-system aims at being the future coastal component of the national ocean observing effort, as part of the IT-IOOS. Through multidisciplinary observations, the JERICO sub-system will thus improve the knowledge on how national coastal marine systems respond to global and local drivers.

6.9.2. *Sub-system architecture*

Coordination, SS-components and Production Units

The JERICO sub-system (SS-JERICO) is coordinated by CNR-ISMAR and structured in three logical/functional components, namely two internal components, i.e. "JERICO-ISMAR" and "JERICO-OGS", and one external component named "EUROGOOS HFR NODE". Below, details of components and production units are reported:

- JERICO-ISMAR. This component collects all five production units belonging to ISMAR, namely *PU ISMAR SP, PU ISMAR RM, PU ISMAR TS, PU ISMAR BO and PU ISMAR VE* and manage the following facilities: AAOT, the moorings in the Corsica and Sicilian Channels, PALOMA, S1-GB elastic beacon, Surface Velocity profilers (SVP), HF radars, drifters, GUARD-one imaging systems and a smart observatory facility.
- JERICO-OGS. This component consists of one Production Unit (PU) *PU OGS - GoT*, supported by an existing and new planned facility, namely MAMBO–Miramare, MAMBO2, MAMBO3, MAMBO4, DWRG1, DWRG2, DWRG3, Isonzo, Bevazzana, and GoT HF-Radar. These observational platforms are part of the Gulf of Trieste (GoT) Observing System and operate in near real-time.
- EUROGOOS HFR NODE. This component is made up of a single production unit, the *PU HFR JERICO*, which operates on the HF radars belonging to the Italian JERICO partners, i.e. CNR-ISMAR and OGS.

The list of EOVs, EBVs and ECVs produced by SS-JERICO, their datasets and related present and planned access points are listed in annex 1 at the end of this document. This is a living list that will get updated with also the new data that will be acquired by the planned ITINERIS equipment under the JERICO-RI framework.

The diagram in Figure 8 resumes the logical/functional architecture of the JERICO sub-system.

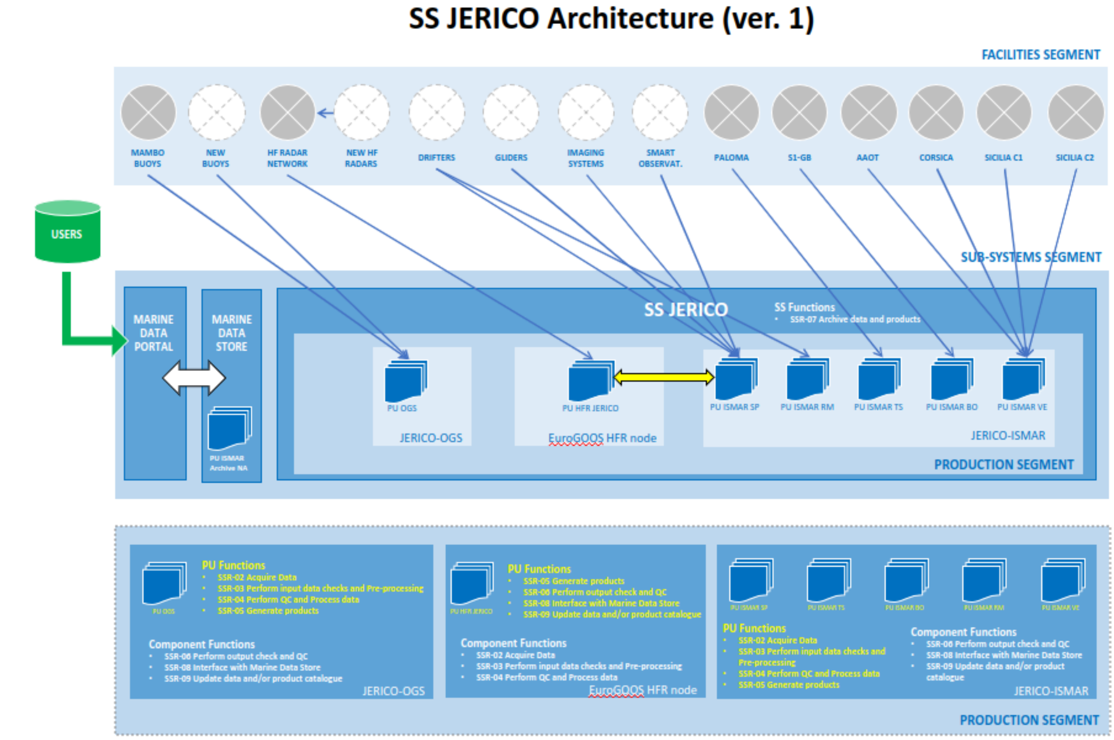

Figure 8. Logical and functional architecture of SS-JERICO

6.9.3. Local physical architecture and functional modules

The three above-cited components, trough the Marine Data Store, provide the variables listed above, together with the metadata information, to the ITINERIS Marine Domain Management Team and to the Marine Data Portal. The corresponding data are acquired by each of the component's PUs.

In the following paragraphs the detailed description of each JERICO SS Component is provided together with their production unit constituents.

The JERICO-ISMAR Component

The JERICO-ISMAR component collects all five production units belonging to ISMAR, namely PU ISMAR SP, PU ISMAR RM, PU ISMAR TS, PU ISMAR BO and PU ISMAR VE. The component's primary function is to interface with the Marine Data Store (**SSR-08**) preparing data formats and data and/or products retrieval according to specifications and vocabularies decided by the JERICO sub-system. The JERICO-ISMAR component is also performing output checks and quality control procedures (**SSR-06**) on output files according to the same specifications and vocabularies before being transfer to the JERICO sub-system. The component is also responsible for updating data and/or product catalogue (**SSR-09**) that are provided and made accessible on a planned schedule to the JERICO sub-system first and eventually to the Marine Data Portal. All other functions are relegated to the production unit level for the specific type of facility detailed below.

PU ISMAR SP

The JERICO production unit ISMAR SP is currently the endpoint of the European HFR Node for the near-real time data for HF radars (see EUROGOOS HFR NODE component description below and the yellow arrow in the diagram in Figure 8) and will be the end point of several facilities acquired during the ITINERIS project and belonging to CNR-ISMAR for both near real time and delayed mode data. These facilities will include different type of drifters (CARTHE and MARTA), three gliders, various GUARD-one imaging systems and a smart observatory facility that will be deployed nearby the Tino Island. The unit currently acquires HF radar raw data automatically and will acquire raw data from the above-cited facilities always automatically (**SSR-02**). Input raw data are and will be checked and pre-processed with automatic systems (**SSR-03**) and subsequently processed, quality-controlled and provided, with the corresponding metadata, in standards and formats recognized at the international level (**SSR-04**). When possible additional products (like for example diffusivity and heat content estimates) will be generated always using standards and formats recognized at the international level (**SSR-05**). Drifters' data and those coming from the profilers from the PU ISMAR RM will be harmonized at the component level.

PU ISMAR RM

The JERICO production unit ISMAR RM will be the endpoint of the Surface Velocity Profilers (SVP), for both the normal SVP and the new augmented SVP-bio ones. SVP raw data will be automatically acquired (**SSR-02**), checked and pre-processed with automatic systems (**SSR-03**), processed, quality-controlled and provided in standards and formats recognized at the international level (**SSR-04**). SVP data and those coming from other Lagrangian instruments from the PU ISMAR SP will be harmonized at the component level.

PU ISMAR TS

The JERICO production unit ISMAR TS is the endpoint of the data acquired on the PALOMA elastic beacon. Some PALOMA data are transmitted in near-real time to a dedicated server, some other data are periodically downloaded from the instruments. PALOMA data will be acquired (**SSR-02**), checked and pre-processed with automatic systems (**SSR-03**), processed, quality-controlled and provided in standards and formats recognized at the international level (**SSR-04**). When possible additional products (like for example heat content estimates) will be generated always using standards and formats recognized at the international level (**SSR-05**). PALOMA data and those coming from other fixed platforms from the PU ISMAR BO and PU ISMAR VE will be harmonized at the component level.

PU ISMAR BO

The JERICO production unit ISMAR BO is the endpoint of the data acquired on the S1-GB elastic beacon. S1-GB data will be acquired (**SSR-02**), checked and pre-processed with automatic systems (**SSR-03**), processed, quality-controlled and provided in standards and formats recognized at the international level (**SSR-04**). When possible additional products (like for example heat content estimates) will be generated always using standards and formats recognized at the international level (**SSR-05**). S1-GB data and those coming from other fixed platforms from the PU ISMAR TS and PU ISMAR VE will be harmonized at the component level.

PU ISMAR VE

The JERICO production unit ISMAR VE is the endpoint of the data acquired on the Acqua Alta Oceanographic Tower (AAOT), the mooring in the Corsica and Sicilian Channels. Acqua Alta and mooring data will be acquired (**SSR-02**), checked and pre-processed with automatic systems (**SSR-03**), processed, quality-controlled and provided in standards and formats recognized at the international level (**SSR-04**). When possible additional products (like for example heat content estimates) will be generated always using standards and formats recognized at the international level (**SSR-05**). Acqua Alta and mooring data will be harmonized with those coming from other fixed platforms from the PU ISMAR TS and PU ISMAR BO at the component level.

The JERICO-OGS Component

The JERICO-OGS component consists of a Production Unit (PU) PU OGS - GoT, supported by an existing and new planned facility.
The OGS - GoT Production Unit can provide SSR from 2 to 6. About SSR from 7 to 9, they will be implemented according to IT-IOOS requirements and will be made available at OGS-NODC end point in communication with the Marine Data Store. In addition to data from facilities, there are also products, such as Numerical Model Datasets.

The PU OGS - GoT is the end point of a network of observing systems composed by the following platforms: MAMBO–Miramare, MAMBO2, MAMBO3, MAMBO4, DWRG1, DWRG2, DWRG3, Isonzo, Bevazzana, and GoT HF-Radar. These observational platforms are part of the Gulf of Trieste (GoT) Observing System.
These facilities, operate in near real-time. Data collected by platforms are sent, through GSM Modem, to the National Oceanographic Data Center (NODC) at OGS, where they are stored and made available through an ERDDAPP Server. Access to the ERDDAPP server and thus to the data is free and accessible to anyone who requests it.
The only exception is the GoT HF-Radar. This facility is part of the Northern Adriatic High Frequency Radar Network (HFR-NAdr), which is endpoint of the European HFR node for near-real-time HF radar data. In this case data are collected on a virtual machine at CINECA. On the same machine, data is processed and is executed QC. Then data is sent to the European HFR node.
Access to the HFR-NAdr data through the European HFR node facility is free and accessible to anyone who requests it.
As mentioned, above, the PU OGS - GoT, through the NODC facilities can provide the following services:

- SSR-02 Acquire Data
- SSR-03 Perform Input Data Checks and Pre-processing
- SSR-04 Perform QC and Process Data
- SSR-05 Product Generation
- SSR-06 Archive data and Product

As part of the ITINERIS project, it is planned to expand the observational capabilities of the JERICO-OGS component by extending measurements capabilities of MAMBO – Miramare wtih an ADCP and expanding the coverage of the HFR-NAdr Network through the installation of a new HF-Radar system.
The data collected by these new facilities, will also be processed according to similar data flow

The EuroGOOS HFR Node Component

The EuroGOOS High Frequency Radar (HFR) Node is a thematic component within the JERICO SS dedicated to the processing chain of High Frequency Radar data. The component capitalizes on the practice of most of the HF radar operators and data providers in Europe which are organized under the EuroGOOS HFR Task Team framework and provide data to a data management competence center named as the "European HFR Node". The Node was established by three JERICO partners (AZTI, CNR-ISMAR and SOCIB) in 2018 and it has been operational since April 2019 for delivering data to the Copernicus Marine Service (CMEMS-INSTAC), SeaDataNet (SDN/SDC) and EMODnet Physics. The Node also distributes tools and support for standardization to the HFR providers, as well as standardized near-real time and delayed-mode HFR data on surface marine currents. In particular, the Node implements the functions of data acquisition and harvesting, quality control, validation/assessment, near-real time data delivery and historical data distribution with different reprocessing levels. All ITINERIS HF radar facilities will be connected to the Node.

Within the JERICO SS, the EuroGOOS HFR Node component is made up of a single production unit, the PU HFR JERICO, which operates on the HF radars belonging to the Italian JERICO partners, i.e. CNR-ISMAR and OGS. The logic of this JERICO SS component is different from the two previous ones as lower-level functionalities are performed this time at the component level. The node is indeed used for the near-real time HF radar marine current data collection (**SSR-02 and SSR-03**), for the application of the standard QC model and the conversion of the quality-controlled data files into the European standard data and metadata model (**SSR-04**).

CNR-ISMAR is actively involved into the Node operations as it is in charge for the above-described functionalities for near-real time HF radar data. All other functionalities are relegated at the unit level as described below.

PU HFR JERICO

The production unit HFR JERICO is strongly linked to the production unit ISMAR SP (yellow arrow in the diagram in Figure 8). It will manage new product generation like coastal marine current climatology in the Northern Adriatic, Tyrrhenian and Ligurian seas (**SSR-05**) and perform output check and quality control (**SSR-06**). It will also coordinate with the Marine Data Store (**SSR-08**) both for HFR data and data catalogue update (**SSR-09**).

6.9.4. Computing and archiving facilities

The computing and archiving facility that hosts SS-SIOS data and products is the Data center, located at Bologna at the CNR-ISP. Its storage and capacity is 750 TB. The archive will ensure real-time accessibility of data and maintenance of data for at least 10 years and backup of online data, including any previous versions.

6.9.5. Physical interfaces of the sub-system (external interfaces)

SS-JERICO will have the following external interfaces:

- JERICO RI
- HFR network
- COPERNICUS Marine

6.9.6. Design constraints

N/A

6.9.7. Relationships between sub-systems (internal interfaces)

SS-JERICO will have an internal interface with MDS to provide access to the data though the IT-IOOS web portal.

SS-JERICO will exchange information with SS-LNS where acoustic data, routine and QC procedure will be exchanged.

SS-JERICO will be have an interfaced with SS-eLTER and SS-DANUBIUS

6.9.8. Overall architecture growth

SS-JERICO has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. All facilities, gliders and buoys have the capability to accommodate extra instrumentation for air-sea interface and oceanographic observations. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. In respect to dataset and catalogue, the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM.

6.9.9. Internal failures analysis

N/A

6.10.Sub-system Laura Bassi

6.10.1. Purpose of the sub-system

The primary purpose of the Laura Bassi sub-system within the Italian Integrated Observing System (IT-IOOS) is to enhance and broaden digital integration, with a specific focus on scientific equipment, automation, and data harvesting. The aim is to optimize the collection and storage of information from all onboard equipment.

At present within the Laura Bassi-SS NRT-data are provided by the navigation facility and the weather station, which will be described in the following paragraphs. It is to notice that our system is undergoing constant upgrades to enhance its capabilities.

6.10.2. Sub-system architecture

General information about the Laura Bassi R/V

The Laura Bassi is equipped with a series of advanced hardware and infrastructure. This setup is crafted to support collaborative research initiatives.

The Kongsberg Seapath 380 is developed specifically for hydrographic surveying, catering to the need for utmost precision in measurements of heading, position, roll, pitch, heave and timing. This product combines state-of-the-art inertial technology and processing algorithms, complemented by dual frequency GPS, GLONASS, Galileo and Beidou satellite signals. The Seapath navigation algorithms integrate the RTK (Rea-time Kinematics) GNSS data with the inertial sensor data. This unique integration grants the Seapath 380 distinct advantages over standalone RTK products. The precision in roll, pitch, and heading measurements allows for referencing the RTK antenna position to any point on the vessel, ensuring accuracy in position and velocity where needed. All the data from Seapath have the same time stamp and the output is in real-time. Sub-decimeter positional accuracy can be attained by acquiring satellite orbit and clock data from the internet, coupled with post-processing of both satellite and IMU (Inertial Measurement Unit) data.

Airmar's ultrasonic WeatherStation® WX200 delivers precise and dependable weather data through a robust, compact, and maintenance-free design, devoid of any moving parts. The instrument measures wind speed, wind direction, air temperature, relative humidity and barometric pressure. The system calculates the dynamic true wind speed and direction by considering the apparent wind, the vehicle's speed, and its heading. The internal 10 Hz GPS and three-axis electronic compass are instrumental in providing essential data, including heading, position, speed-over-ground, and course-over-ground. These functionalities are crucial for the accurate processing of dynamic true wind data, especially when the vessel is in navigation.

Coordination, SS-components and Production Units

The Laura Bassi SS is coordinated by OGS and is composed by two logical components: LAURABASSI-RV and LAURABASSI-OGS.

- The **LAURABASSI-RV component** is the acquisition system on board of the Laura Bassi polar research vessel. Several systems are connected to a local infrastructure, which collects NRT data and transmits them through internet.
- The **LAURABASSI-OGS component** is the ground facility, located in Sgonico (Trieste, Italy), where NRT are received, processed, stored and subsequently disseminated, through web interfaces and as data products.

The list of EOVs, EBVs and ECVs produced by SS-Laura Bassi, their datasets and related present and planned access points are listed in annex 1 at the end of this document. The list will be updated if any variation will occur.

Facilities of the SS-Laura Bassi

The SS-Laura Bassi owns 9 different kind facilities and new ones will be implemented within ITINERIS project. Here below a brief description of them is reported:

The instruments currently active are:

- **Kongsberg SeaPath GPS.** Seapath is a system providing navigation information about heading, position, roll, pitch, heave and timing. It combines state-of-the-art inertial technology and processing algorithms, complemented by dual frequency GPS, GLONASS, Galileo and Beidou satellite signals.
- **Airmar Weather Station**. The weather station provides meteorological data. Specifically, the instrument measures wind speed, wind direction, air temperature, relative humidity and barometric pressure. The system calculates the dynamic true wind speed and direction by considering the apparent wind, the vehicle's speed, and its heading.

The instruments that new facilities that will be implemented are:

- **Kongsberg HUGIN Autonomous Underwater Vehicle (AUV).** The AUV system is a multi-role vehicle capable of collecting high resolution data for commercial, scientific and defense applications. It is equipped with a multibeam echosounder, a side-scan sonar, and mounts sensors for measuring carbon dioxide, nitrates, turbidity and oxygen. In the near future, HUGIN AUV will be operative onboard the Laura Bassi SS and additional instrumentation will be implemented.

6.10.3. Local physical architecture and functional modules

Laura Bassi RV component

SSR-02: Acquire data

The Laura Bassi R/V component manages the acquisition of all instrumetns installed on the ship vessel. The LAURABASSI-RV system is composed by Linux servers interconnected through the LAN. Data acquisition is performed by Python scripts, launched periodically by "cron". The measures from the SeaPath and the Airmar weather station are broadcasted through UDP messages.

All the data from Seapath have the same time stamp and the output is in real-time. Sub-decimeter positional accuracy can be attained by acquiring satellite orbit and clock data from the internet, coupled with post-processing of both satellite and IMU (Inertial Measurement Unit) data.

All the processed data are collected and put into a structured JSON file, which provides a NRT snapshot, every minute. The JSON data are sent to LAURABASSI-OGS by means of satellite internet connection (provided on VSat C-band, Ku-band or Starlink). The receiver is a web service operative at the OGS headquarters and accessible through a Rest API developed in PHP. The data flow from the PU Receiver to the PU Storage.

Laura Bassi OGS component

This component will manage the processing of the all data collected and transmitted by the Laura Bassi R/V component.

SSR-03: Perform input data checks and Pre-processing

The data are pre-processed, validated, and stored in a PostgreSQL database.

SSR-04: Perform QC and Process data

The data are pre-processed, validated.

SSR-05: Generate products

N/A

SSR-06: Perform output check and QC

N/A

SSR-07: Archive data and products

The data are stored in a PostgreSQL database.

SSR-08: Interface with Marine Data Store

Data are made available by the PU Access, both as web user interface and as web services for machine-to-machine access. The data can be retrieved by the Marine Data Store through queries to a Rest API, returning a result in JSON format, based on a coordinated dictionary for metadata.

The SS-Laura Bassi will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-Laura Bassi into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the ITINERIS deliverable D5.11. The update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (SSR-09)

6.10.4. Computing and archiving facilities

Laura Bassi will use the computing and archiving facilities located at OGS. The size of these facilities, thanks also to the upgrades that will be done during ITINERIS, will guarantee 10 years of usage.

6.10.5. Physical interfaces of the sub-system (external interfaces)

N/A

6.10.6. Design constraints

N/A

6.10.7. Relationships between sub-systems (internal interfaces)

The physical interfaces of the Laura Bassi sub-system within IT-IOOS will be specifically designed to be interfaced with the Marine Data Store. This interface will streamline the exchange of research plans, observational data, and metadata, promoting a collaborative environment for oceanographic research activities.

6.10.8. Overall architecture growth

The Laura Bassi SS is easily expandable and scalable to incorporate additional sensors and variables. The integration with the technical data harvested by the vessel automation system will be implemented during the next year. This system which has been implemented onboard but currently operates a “closed” system and is available exclusively in consultation mode both within the LAURABASSI-RV and the LAURABASSI-OGS components. Once the data will be briefly analysed, they will be transfer to the ITINERIS data facility. The above system constitutes the informatica backbone of the scientific instruments and systems. Scientific data retrieved by onboard facilities during the upcoming campaigns will be made available, when possible, to contribute seamlessly to their NRT and DT transmission. Soon the Kongsberg HUGIN Autonomous Underwater

Vehicle (AUV) will be operative onboard the Laura Bassi SS. The design of the Kongsberg HUGIN AUV follows a modular open system approach. Different sensors can be installed on the AUV, depending on the operations and data requirements. Even though the AUV system is not part of the ITINERIS project, the system for deployment and underwater communication that will allows the data retrieving during immersion is one of the goals of the project.

6.10.9. Internal failures analysis

The NRT acquisition system on board of R/V Laura Bassi is remotely monitored and controlled, through several links, like a point-to-point WireGuard VPN, local machines connected by AnyDesk remote desktop, and the automatic transmission of status and parameters. The overall connectivity of the R/V is constantly checked and monitored and has an automatic fail-over management, switching between different satellite providers. The processing and deployment system on the ground has a high degree of resilience, being based on multiple servers, with periodic backup on NAS units and a high-profile internet connectivity (OGS scientific institution is connected to the GARR backbone).

6.11. Sub-system LNS

6.11.1. Purpose of the sub-system

The LNS is building and will soon deploy and operate a subsea node (junction box) for the Italian Integrated Observing System, capable to provide high reliability/high-availability/high-speed data link and uninterruptible power supply to multiparametric deep-sea observatories, with demanding time-synchronisation capabilities and data payloads and transfer rates, such as acoustic sensors.
In this aim LNS has designed a sub-system to acquire, store and expose though the ITINERIS Marine data portal and Project data portal the Ocean Sound EOV measured from acoustic data.
The system will be first implemented using the available LNS acoustic data resources (streamed or stored) and therefore extended to resources from partners RI's.

In recent years, indeed, the attention of the marine science community to the ocean sound has dramatically increased, focusing mainly on impact of anthropogenic noise in marine habitat.
Recommendation on ocean sound measurement strategies and its implementation plans have been provided in the context of the Marine strategy Framework directive and several initiatives guided by JPI oceans (https://www.jpi-oceans.eu/en).

In the context of the ITINERIS project a subsystem for Ocean Sound data collection, analysis and distribution is under development, with primary aim at providing interoperability of data and analysis tools among the RI's providing ocean sound data.

Ocean Sound is an information highly variable in space and time; a detailed implementation plan for acoustic measurement should considered: measurement type (pressure or particle motion) detector location, directivity and sensitivity, sampling and resolution, duration of single recording and interval of acquisition.
Complete recording of the soundscape of a marine area requires a sizable effort in terms of data streaming/storing capabilities, considering a typical data transmission rate of 6.2Mbps for a single hydrophone sampling data at 24bits/192 kHz to detect the majority of sound sources (see Figure 9).

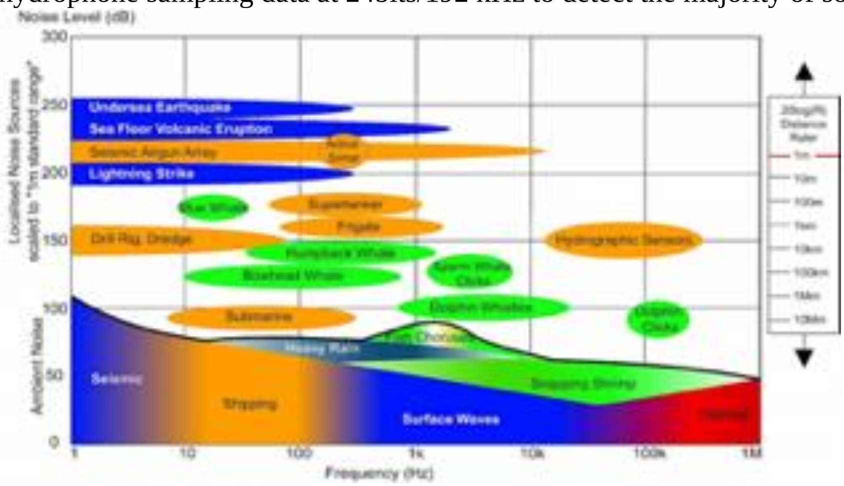

Figure 9. Complete recording of the soundscape of a marine area.

Shared tools and algorithms for data analysis is also a critical point, to provide to users and stakeholders interoperable and clearly interpretable data.

It is also worth mentioning that acoustic marine data are sensible data, therefore open distribution of raw data is not foreseen in the project.

For the ITINERIS implementation of the Italian IT-IOOS the selected variable for Ocean Sound Analysis is the Sound Pressure Level (SPL) measured in 1/3 octave bands.

6.11.2. Sub-system architecture

A schematic of the design of the Ocean Sound Sub System is reported in figure below (Figure 10Figure 1).

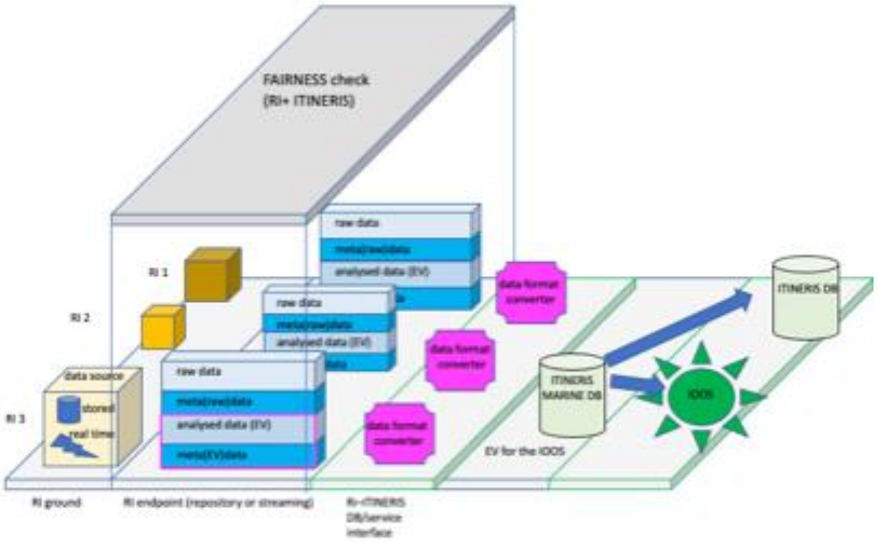

Figure 10. Logical and functional architecture of SS-LNS.

Each RI will be responsible to ensure the quality of raw data and the availability of a subset of raw data for checks and reproducibility.
At the ITINERIS Project level FAIRness will be ensured through a shared protocol for data analysis and data quality check. Eventually every RI (or data provider) will provide Ocean SPL data in a proper format and through an identified and certified end-point to the Marine database.

6.11.3. Local physical architecture and functional modules

At the physical level, the Ocean Sound sub-system (specifically the RI-LNS) is be equipped with the necessary hardware and infrastructure to support collaborative analysis and data-sharing.

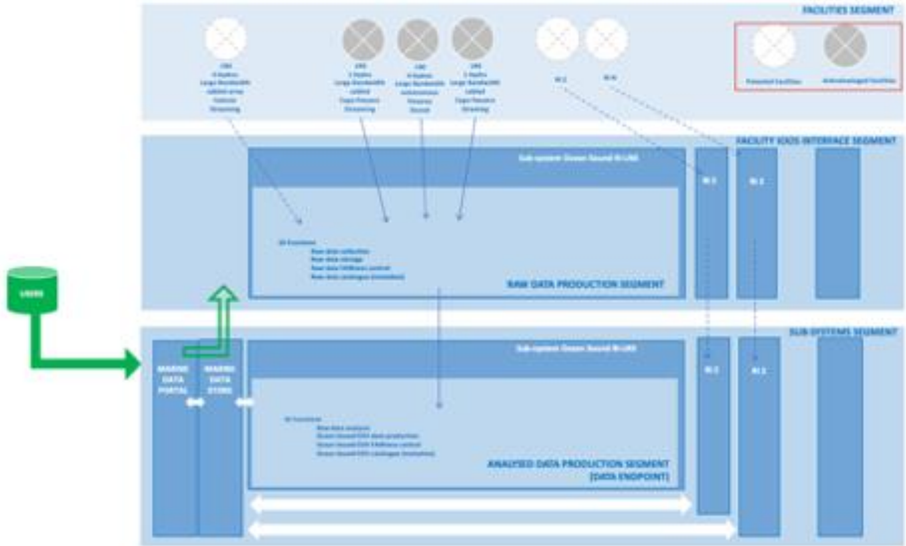

Figure 11. Architecture design of SS-LNS

LNS-RI has already in place data acquisition, transport, processing/analysis and (redundant) storage systems for raw ocean sound data both for cabled and for autonomous acoustic marine sensors.

High availability/high reliability/high speed networking connection from RI to the Marine Data Store is ensured, as LNS-RI is part of the GARR-X network

6.11.4. Physical interfaces of the sub-system

The physical interfaces between the IT-IOOS Ocean Sound Sub System and RI endpoints are designed to allow continuous (real-time, when possible) update and retrieving of analyses data.

Ack/update messages are exchanged between the Marine data portal and RI endpoint after production of new data at the RI facility.

The actual physical data endpoint of the SPL data is located at LNS, reachable at 172.16.14.110 (with access restricted to sigle users, and open to the Marine data Portal services).

6.11.5. Design constraints

According to (and extending) best practices suggested by GOOS and Emodnet on acoustic metadata formats, the following information should be associated to raw data by the data provider (RI):

Category	Subcategory	Unit	Notes
Recording time	recording start time recording duration	UTC s	
File specs	data-format n. of channnels sampling rate	Hz	n. of channels or tracks in the recording

	quantization	bits	
Sensor	name manufacturer part number serial number sensitivity: average value sensitivity curve	 dB re V/uPa dB re V/uPa	This is an array in case of multi-track recording
Readout-Electronics	preamplifier gain amplifier gain total gain ADC_V_FullScale ADC n of bits filter low cutoff filter high cutoff	dB dB dB V Hz Hz	This is an array in case of multi-track recording
Mooring	mooring name mooring type seabed depth Easting Northing WGS84 zone date of installation date of recovery mooring nickname(s)	 m m m UTC UTC	
Mooring property	Institute Project Contact person Contact email Data licence		Name, country Name, funding source
Data property	Contact Person Contact email Data Licence		

Raw data quality:
Sensor calibration must be known and reported. Data from uncalibrated sensors should be clearly indicated and can be used for a limited subset of analyses (such as noise trends).

Data harvesting strategy:
it is recommended to record at minimum 5 minutes of data every hour. This permits the study of many indicators and this is in line with those described by the MSFD

Data analysis:
SPL values should be calculated for octave thirds comprising 63 and 125 Hz.
When possible supplemented with those for 250 Hz, 500 Hz, 1 Hz, 2 Hz, 5 Hz, and 20 kHz.
For bio-acoustic, underwater communication and high energy physics studies further bands should be included (sampling should be typically 192 kHz).
Spectral resolution of 1Hz should be provided till 1 kHz (e.g. 2048 points FFT, 50% window overlap @2kHz)
Spectral resolution of 100 Hz should be provided for higher frequencies (2048 points FFT, 50% window overlap @192kHz)
The average and median values of SPL should be calculated together with the 25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles at 1, 2, and 5 minutes.
Algorithms and code for analysis will be shared over the project software sharing platform.
Automated or existing data analysis procedures and codes available in the RI portfolio must be cross-checked with the ITINERIS project official SPL analysis code to ensure data FAIRnness.

SPL data quality:
According to (and extending) best practices suggested by GOOS and Emodnet on acoustic metadata formats, the following information should be associated to SPL data.

Table 1. List of acoustic metadata formats and the necessary information to be associated to SPL data according to best practices suggested by GOOS and Emodnet.

Category	Subcategory	Unit	Notes
SPL data			This is an array in case of multi-track recording
	SPL Unit		dB re 1uPa
	Start time	UTC	
	Time window	s	
	FFT points	N	
	Window type		
	Window points		
	Overlap	%	
	F centre-band list	Hz	1/3 octave band central F
	F min-band list	Hz	1/3 octave band min F
	F max band list	Hz	1/3 octave band max F
	Number of calculated quantiles		average, median, 25, 75, 95, arb1, arb2, arbN
Inherited from raw data source file			
Raw data source	filename		
Recording time	recording start time	UTC	

	recording duration	s	
File specs	data-format n. of channels sampling rate quantization	 Hz bits	 n. of channels or tracks in the recording
Sensor	name manufacturer part number serial number sensitivity: average value sensitivity curve	 dB re V/uPa dB re V/uPa	This is an array in case of multi-track recording
Readout-Electronics	preamplifier gain amplifier gain total gain ADC_V_FullScale ADC n of bits filter low cutoff filter high cutoff	dB dB dB V Hz Hz	This is an array in case of multi-track recording
Mooring	mooring name mooring type seabed depth Easting Northing WGS84 zone date of installation date of recovery mooring nick-name(s)	 m m m UTC UTC	
Mooring Property	Institute Project Contact Person Contact email Data Licence		
Data Property	Contact Person		

	Contact email Data Licence		

6.11.6. Relationships between sub-systems

The Ocean Sound Sub System can be interconnected to other RI's / Sub Systems to retrieve acoustic data from other sensors in order to extend the time and spatial coverage of the IT-IOOS with respect to the Ocean Sound variable.

In terms of holistic approach in the study of the marine environment and impact of anthropogenic noise, the Sub System will benefit from interaction with Subsystem providing meteorological data, oceanographic data, accurate bathymetries and vessel and marine fauna distribution

6.11.7. Overall architecture growth

The proposed Ocean Sound Sub System is scalable in terms of:

- number and type of data sources that can be integrated;
- number of Ocean Sound Variables that can be measured and delivered.

Fostering implementation of more complex sound analysis (e.g. automatic detection of somniferous fauna or man-made sounds, or sound source tracking) the number of Ocean Sound variables can be either increased.

6.11.8. Internal failures analysis

Incident tracking is elaborated at the level of RI and should be acknowledged to the Marine data portal. This is relevant for streamed data (cabled observatories), while for stored data provision the main concern is related to raw data redudancy and conservation.

Streamed and stored data availability: LNS-RI provides already a space and time distributed network of acoustic marine sensors as seed of the IOOS. Stored dataset is also available. Networking: GARR provides excellent network performances between data raw data providers, data analysis sites (RI endpoints) and the portal. Failure rate of the network is negligible.

Storage: Raw data storage is demanding. LNS-RI benefits of well-equipped data center to host storage capacity available and to be increased with project funding. Raw data will be stored in RAID. RI entering in the Ocean Sound Subsystem should provide adequate storage resources. Very limited storage resources are needed to store the SPL in the Database, that can be duplicated in case of failure.

Computing resources: Computing resources can be provided as physical or virtual machines with default OS and codes for connection and analysis, providing fast replacement in case of failure.

Interfaces: Documentation for data exchange/access interfaces will be provided to minimize downtime at system bootstrap and in case of upgrades at the Marine portal or RI side.

6.12.Sub-system SIOS

6.12.1. Purpose of the sub-system

The SIOS sub-system aims to expand the research capacity to investigate the artic environment of the Svalbard region in a multidisciplinary way, through regular and up-to-date delivery of data and products produced by the SIOS-RI facilities. At national level, the mission of the SS-SIOS is linked to that of the Svalbard Integrated Arctic Earth Observing System (https://sios-svalbard.org/), which has developed and maintains a regional observing system for in-situ long-term measurements in and around Svalbard to address key Earth System Science questions related to Global Change.

6.12.2. Sub-system architecture

Coordination, SS-components and Production Units

The SS-SIOS is coordinated by CNR-ISP and composed by two PUs: CNR-ISP-ME and OGS.

- CNR-ISP-ME uses data from the facilities: MDI (boa and mooring), KIM, S1 and Coastal Buoy
- OGS uses data from the S1

The diagram in Figure 12 resumes the logical/functional architecture of the SS-SIOS.

The list of variables, within each component, with defined end points is provided reported in the annex 1. This list will be a living document that will also be updated with new data from planned ITINERIS equipment. Within SS-SIOS, the catalogue of SIOS-RI provides a common SIOS metadata catalogue for all facilities.

Facilities of the SS

The SS-SIOS regional facilities managed by Italian research organizations are three: **S1**, **KIM** and **MDI (mooring).**

S1:
S1 is jointly managed by **CNR-ISP** and **OGS** and consists of moorings and a surface buoy for data transmission via satellite; the facilities acquire measurements addressing mass properties, biogeochemical cycles, ecosystem function, especially concerning carbon sequestration dynamics and acidification processes in surface/deep waters;

KIM
KIM is managed by **CNR-ISP** and consists of moorings and a surface buoy for data transmission via satellite; the facilities acquire measurements addressing mass properties, biogeochemical cycles, especially concerning carbon sequestration dynamics and acidification processes in surface/deep waters;

MDI (mooring)
MDI (mooring) is managed by **CNR-ISP** and consists of moorings and a surface buoy for data transmission via satellite; the facilities acquire measurements addressing mass properties, biogeochemical cycles, ecosystem function, especially concerning carbon sequestration dynamics and acidification processes in surface/deep waters;

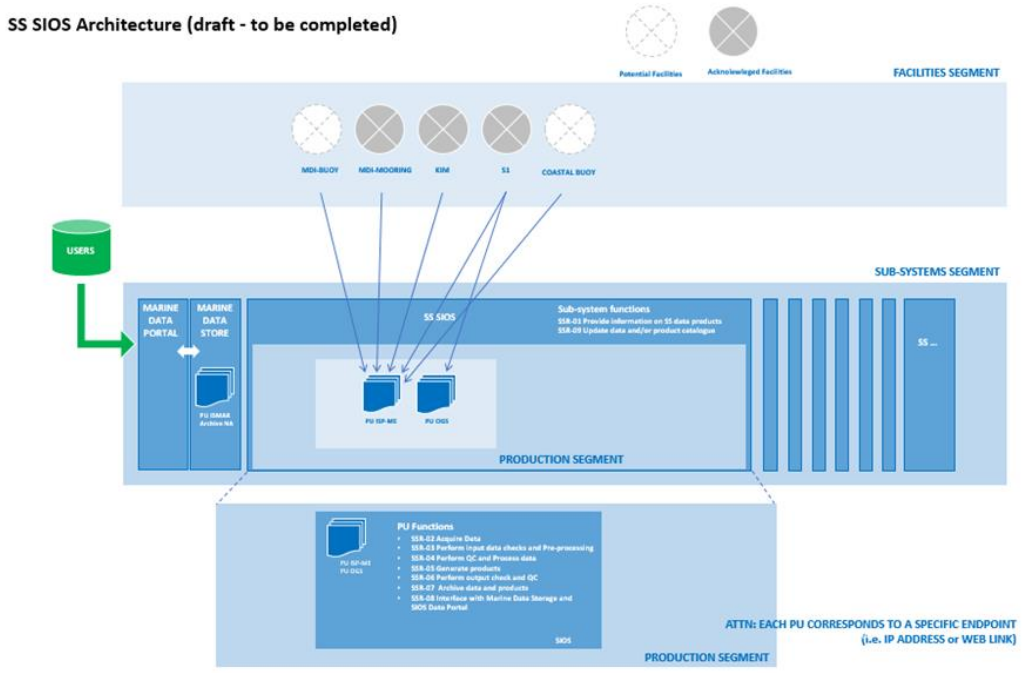

Figure 12. Logical and functional architecture of the SS-SIOS

The sub-system will provide the variable listed reported in the annex 1 accompanied by metadata to the ITINERIS Marine Domain Management Team and the Marine Data Portal.

6.12.3. Local physical architecture and functional modules

The local physical architecture of each SIOS facilities is described hereafter with details on the data flow from the data acquisition to the data dissemination. The compliance of each SS facility and PU with the sub-system requirements is reported in annex 2.

Within SS-SIOS, two PUs actively contribute to Sub-System Requirements. ISP-ME and OGS do contribute to Acquire Data (SSR-02), perform input data checks and pre-processing (SSR-03), perform QC and process data (SSR-04). For all data and products, the ISP-ME performs output check and QC (SSR-06), archive (SSR-07).

Specifically, ISP-ME PU is the end point of several facilities belonging to CNR-ISP. The facilities are: MDI (boa and mooring), KIM, S1 and Coastal Buoy. The PU manages these facilities and the data are the results of multiple measurements made by installed equipment and sensors (SSR-02, SSR-03, SSR-04). The PU will perform output check and QC (SSR-06), archive data and products (SSR-07) and provide the interface with Marine Data Storage and SIOS Data Portal (SSR-08) through the Italian Artic Data Center (IADC).

The OGS PU is the end point of S1 facility, which is managed together with the ISP-ME (SSR-02, SSR-03, SSR-04).

Details of each requirement and data flow are given in the description below.

SSR-02: Acquire data

Within the PU, raw data in delayed and real-time mode recorded by the automatic sensors in the facilities are acquired. Data in delayed mode are recorded and stored within the instruments, and after the instrument's retrieval, they are downloaded by the operator and enter the processing chain manually. Sensor management in the delayed mode configuration is independent for each instrument which is programmed before deployment. Each instrument has an independent power supply, and the sampling is the results of the best sampling strategy related to the scientific purpose and power consumption.

Sensors management (sampling, storage, transfer) in the real-time configuration is performed with a Logical Intelligent System Control (LISC) system, and data is transferred via Iridium satellite transmission.

SSR-03: Perform input data checks and Pre-processing

Within the PU, once the data have been transferred to the ground segment, they are checked and pre-processed. Pre-processing and check of inputs data of both RT and DM data follow the same procedures and it is done through dedicated computers. The routines developed for data processing and QC are performed using Matlab. These first steps consist of an initial gross check of the data, visual inspection, and graphing of the data to assess the reliability of the measurements. The metadata field declares information on the methods adopted for this procedure.

SSR-04: Perform QC and Process data

Within the PU, data are processed and subjected to QC procedures. Data control from the acquisition system to the receiving system and processing is done through dedicated computers. The routines developed for data processing and QC is performed using Matlab. For delayed mode data, the procedures refer to the SeaDataNet Manual. Regarding the processing and QC of Real-time data the procedures refer to "Recommendations for in-situ data Near Real Time Quality Control of GOOS", "Manual for Real-Time Quality Control of In-situ Temperature and Salinity Data – Ocean Optics - Dissolved Oxygen Observations - In-Situ Current Observations" of IOOS. The metadata field contains information of the methods adopted for this procedure.

SSR-05: Generate products

N/A

SSR-06: Perform output check and QC

Within the PU quality control operations (such as compliant test) of the output datafile to be stored into the archive are e performed to verify the completeness of metadata information and use of standard nomenclature, correctness of the data file format, presence of quality control information, quality check on file integrity and completeness on data transferred to the archive.

SSR-07: Archive data and products

Within the PU all data and products are archived in NetCDF format in the SIOS ERDAPP at https://data.iadc.cnr.it/erddap/index.html. EOVs and ECVs of the SIOS research facilities are provided in NetCDF files. The data published on the ERDDAP server use the following name format: FN_INSTR_DEPTHm_yyyymmdd-yyyymmdd_IADC_DM.nc. "FN" stands for "Facility Name" and the options are MDI, KIM, S1, BOA COSTIERA. "INSTR" stands for "INSTRUMENT", e.g. ADCP, CTD etc, "yyyymmdd" is the basic ISO8601 date format of the start and end of deployment, and "DM" stands for "Delivery Mode" which can be NRT (near real-time data) or D (Delayed). The data produced by SIOS-RI are released under the Creative Commons Attribution 4.0 International License (CC BY 4.0). Data sets are organized by instruments and deployment and are organized and stored locally in separate folders.

SSR-08: Interface with Marine Data Store

SS-SIOS will be interfaced with the Marine Data Store though the SIOS Data Portal, through the Italian Artic Data Center (IADC) with a single endpoint ERDDAP (https://data.iadc.cnr.it/erddap/index.html). Interface for data and products through to the Marine Data Store will be provided within the PU, enabling retrieval of data and products according to thI sub-system described procedure.

The SS-SIOS will maintain and keep constantly updated the data and products catalogues and make it accessible to the Marine Data Portal. Before the integration of the SS-SIOS into the IT-IOOS a complete catalogue of data and products will be provided to the MDS following the specifications indicated by MDS in the ITINERIS deliverable D5.11. The update of data and products catalogues will be planned and communicated in advance to the Marine Data Portal in order to highlight the updates to the users (**SSR-09**)

6.12.4. Computing and archiving facilities

The computing and archiving facility that hosts SS-SIOS data and products is the Data center, located at Bologna at the CNR-ISP. Its storage and capacity is 750 TB. The archive will ensure real-time accessibility of data and maintenance of data for at least 10 years and backup of online data, including any previous versions.

6.12.5. Physical interfaces of the sub-system (external interfaces)

The IT-IOOS SS-SIOS has as external interface the SIOS connected with the central SIOS ERDDAP portal to provide access to Italian data. SS-SIOS provides a single ERDDAP endpoint for end-users and this configuration is based on a simple URL (https://data.iadc.cnr.it/erddap/index.html) that leads to the desired dataset hosted by the remote ERDDAP server available from Italian Artic Data Center (IADC). The central SIOS-RI ERDDAP server then provides these datasets as if they were hosted locally.

SS-SIOS provides a single ERDDAP endpoint for both PU for end-users. Thus, the unique data repository where the most up-to-date data files of SIOS-RI facilities are stored is the ERDDAP data server of the SIOS-SS. This configuration also allows the content of the files to be updated, automatically providing users with the most recent versions of the data.

Interoperability is guaranteed as users can access the data transparently via the central ITINERIS hub. This optimizes the use of storage space and centralize the polar repository with the ITINERIS Marine Data Store and the international SIOS data portal (https://sios-svalbard.org/). The metadata will enable easy access and ensure the FAIRNESS principle.

6.12.6. Design constraints

SS-SIOS is designed to be as FAIR-compliant as possible. The specific data format with standardized structure and the accessible interface of the data repositories will limit constrains and ensure Interoperability and Accessibility with the subsystem.

6.12.7. Relationships between sub-systems (internal interfaces)

SS-SIOS will have an internal interface with MDS to provide access to the data though the IT-IOOS web portal.

SS-SIOS will exchange information with SS-LNS where acoustic data, routine and QC procedure will be exchanged.

6.12.8. Overall architecture growth

SS-SIOS has been designed as a modular system which allows to expand and add new modules, new facilities, and components without impacting the overall existing architecture. All facilities have the capability to accommodate extra instrumentation for atmospheric, air-sea interface and oceanographic observations. Additional instruments can be seamlessly integrated, either on a permanent or temporary basis, into the existing data acquisition and control system, allowing for the prompt availability of new measurements in near real-time. Furthermore, the setup and deployment of even high-cost, relatively large, and power-consuming instruments are feasible within this system. In respect to dataset and catalogue, EOVs and ECVs are provided by SIOS-RI in UNIDATA NetCDF files that comply with the NetCDF Oceansites, Climate and Forecast (CF) metadata conventions. Metadata stored in NetCDF data files are fully compliant with the L22 vocabulary (https://vocab.nerc.ac.uk/collection/L22/current/) managed by the SeaDataNet and MarineXML Vocabulary Content Governance Group (SeaVoX). The use of a defined, standardized and hierarchically organized vocabulary ensures scalability of the data management process. Any new parameter and/or metadata can be added to any data file, even in delayed mode, simply updating the NetCDF file.the system is easily expandable ad any new parameter and/or metadata can be added to any data file, even in DM.

In the following paragraphs, upgrades to be realized within ITINERIS are described.

With ITINERIS, new EOVs and ECVs acquired from SIOS-RI are planned to be implemented, but the data architecture of the data flow from acquisition to Marine Data Store will not change. The planned implementation with ITINERIS has already been considered in the data transfer system to the IADC ERDDAP interface.

6.12.9. Internal failures analysis

A control system able to keep track of problems during the different phases of data management will be developed. Specifically, it will track any problems that may occur during data acquisition, transmission, processing and storage. Different precautions and controls are provided depending on the type of system, NRT or D.

The NRT system is provided with a redundant power supply system consisting of both solar panel-powered batteries and power batteries. The system is equipped with an automatic position check within a specific boundary radius that generates an alarm if it is exceeded. The NRT system has a redundant communication and data transfer system.

The different sensors and instruments that acquire data in Delayed Mode are planned to have a power and logging autonomy that ensures it will continue measurements even if the usual maintenance interval is exceeded. The instruments are equipped with backup batteries that protect against power-related issues. There is no way to recover from any temporary failure of such systems depending on external environmental harsh conditions or events, nor to interruption of services of the providers (i.e., Iridium Communications Inc., Radio communication).

In both cases (NRT and DM) the data acquired are duplicated and stored in multiple backup archives to protect data from unforeseen issues.

6.13. GEOSCIENCES

6.13.1. Purpose of the sub-system

The GeoSciences RI aims to develop and strengthen the Italian scientific competencies of the Regional Geological Surveys concerning prioritized geological topics: geological and geo thematic mapping and modeling; landslides and sinkholes; monitoring and risk management; monitoring of georesources and territories. GEOSCIENCES is a coordination network among ISPRA (Institute for Environmental Protection and Research), the Geological Survey of Italy, and the Regional Geological Services (RGS), i.e. the technical offices within the Regions, Autonomous Provinces and Regional Environmental Agencies.

Within ITINERIS, GEOSCIENCES RI will be coupled with IT-IOOS. However, as this GEOSCIENCES RI is currently undergoing the process of construction, the modality of connections with IT-IOOS (as an internal sub-system or an external interface) is still under discussion.

6.13.2. Data integration

The data provided by SS-GEOSCIENCES will be open and shared across the cloud infrastructure, according to the FAIR principles and the INSPIRE standards.

Monitoring data collected by the RI will be integrated and harmonized to be used on Geodatabase and GIS platforms. Standards already available (e.g. SeaDataNet, WMO) will be used and other formats will be proposed and implemented. ISPRA itself will provide monitoring data collected by its own present and future infrastructures.

Data integration will also fully involve the Italian Information System – SIC which collects Italian Monitoring Programme for MSFD. Validation procedures will consolidate monitoring data according to planned will be evaluated with reference to criteria C1-C5) quality assurance processes. Validated data will be integrated with re-analysis modelling products. Process of integration, harmonization, validation has the objective to provide data sets to support the following issues:

-Assessment of GEnS for MSFD, characterization of pelagic and benthic habitats for the identification of new marine protected areas for the implementation of EU Biodiversity Strategy 2030

-Identification of suitable areas for the implementation of renewable energy plants

-Environmental Impact, Strategic Assessment procedures for coastal and marine infrastructures and for programmes and plans as MSP.

7. GOVERNANCE

A strong national governance structure is required to establish policy and provide oversight for all components of the IT-IOOS and to ensure strong integration among the regional, national, and global levels. A coordination and consensus with the Italian Oceanographic Commission (COI) will be established in order to include in the governance of IT-IOOS the priorities of Italian oceanographic community at national and international levels.

7.1. General structure

The governance of the IT-IOOS will be composed of three main boards with executive, management and advisory functions.The **Executive Board** will be the highest-level decision-making body responsible for overall strategic direction, policy development, and major decisions regarding the IOOS. In the first phase (short-term horizon) it will be composed by the representatives of the Italian Marine Institutions parts of ITINERIS (CNR, OGS, ISPRA, INGV, INFN). At a later stage (long-term horizon), the board could be expanded including all Italian Marine Institutions that will enter to participate to the IOOS. The number of the components that will take part to the Executive Board will consider relative dimension of each Research Institution, assigning 2 representatives to the larger institutions and one to the others.

The **Operational Commitee** will be in charge for day-to-day decision-making, implementation oversight, and coordination of IT-IOOS activities. It will be composed by 26 members who are two representatives of each Sub-System of IT-IOOS (11 RIs plus 1 Marine Data Store plus 1 Marine Data Portal).

The **Advisory Council** will provide expertise and recommendations to the Executive Board. It will include representatives from government agencies, academia, industry, and other relevant stakeholders.

7.2. Data Policy

ITINERIS has developed a Data Management Plan and adopted a specific Data Policy. Below the principles of the data policy are listed:

1. The IT-IOOS anticipates that Italian data will be accessible in a comprehensive, unrestricted, and transparent manner for all users, except in cases where limitations are necessary for ethical, cultural, or legal reasons. The IT-IOOS advocates for the principles of open and free sharing of scientific inquiries and data to be universally applicable to Italian data. Therefore, Italian data should strive to be "as open as possible, as closed as necessary," a concept also referred to as "ethically open."
2. The distribution and reuse of Italian data should be without charge.
3. Italian data are expected to be published promptly after collection, preferably in near real-time, unless restricted according to §1 or if an embargo is requested to meet reasonable ethical, cultural, legal, operational, or scholarly needs. Any imposed embargo should have a maximum time limit, accompanied by documented justifications for the embargoed status.
4. Italian data should, to the best extent feasible, adhere to the principles of being Findable, Accessible, Interoperable, and Reusable (FAIR), while considering §1.
5. Italian data must be accompanied by comprehensive metadata, complying with an appropriate international standard such as ISO-19115, DIF10, or DarwinCore. This metadata should contain sufficient details to comprehend, access, and replicate the dataset to the level of quality, accuracy, and precision specified. Additionally, it should include information to support proper attribution.

6. Italian data should be assigned a persistent and globally unique identifier (PID) to facilitate clear identification, attribution, data citation, provenance tracking, linkage with scientific results, and monitoring of data distribution and impact.
7. Italian data should be designated as reusable by attaching a rights waiver, public domain statement, or an internationally recognized data license. These licenses should be non-restrictive, allowing data reuse with no requirements more burdensome than acknowledging the data source, such as the Creative Commons Open Attribution License (CC-BY). Whenever feasible, the rights waiver or license should be assigned by the data owner or source, with these parties identified in accompanying metadata.
8. Users of Italian data should formally acknowledge data authors and sources, employing citation practices equivalent to those used for scientific publications.
9. Data centers and institutions hosting Italian data are expected to adhere to best practices in data curation and preservation for the long term, including the TRUST principles, and to publish data in a FAIR manner, subject to §1.
10. All science projects endorsed by IT-IOOS that anticipate generating data must develop a Data Management Plan, which will undergo review by the IT-IOOS International Project Office and/or the IT-IOOS Data Management Sub-Committee.
11. Data providers are responsible for implementing any necessary quality assurance and control measures to meet community standards, while data users are responsible for ensuring that the data they utilize is suitable for their intended purposes.
12. Data providers are responsible for ensuring that data submitted to data centers are suitable for publication in accordance with their specified license terms.
13. To ensure the long-term preservation and curation of data, particularly in alignment with §9, it is essential for science funders and research institutions to ensure sustained resources for data management beyond individual project lifespans. This includes provisions for hardware, software, and skilled personnel in data preservation and curation, as well as the necessary infrastructure for effective data service delivery.