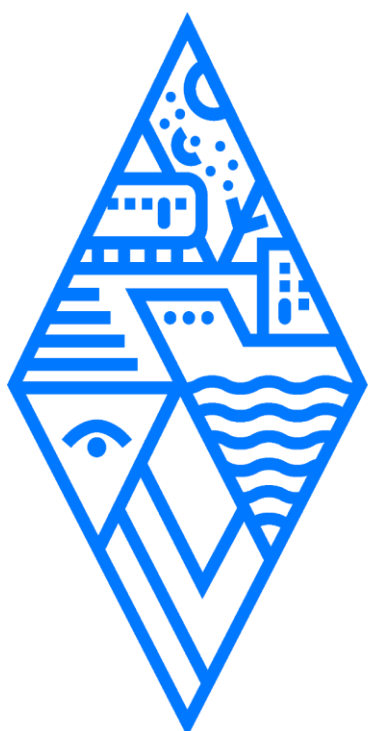




POLARIN

POLAR
RESEARCH
INFRASTRUCTURE
NETWORK

Deliverable 4.6

General guidance materials on FAIR data principles

V1, September 2025

www.eu-polarin.eu



**Funded by
the European Union**

POLARIN: POLAR RESEARCH INFRASTRUCTURE NETWORK

Funding programme: Horizon Europe

Grant Agreement No.: 101130949

Project Start Date: 01/03/2024

Duration: 60 months

Coordinator: Alfred Wegener Institute, Germany

Document information	
Work Package	WP4 Improvement of data services and customised data products
Deliverable No	D4.6
Deliverable title	General guidance materials on FAIR data principles
Version	V1
Dissemination level	☒ PU - Public ☐ PP - Restricted to programme partners ☐ RE - Restricted to a group specified by the consortium ☐ CO - Confidential, only for members of the consortium
Lead Beneficiary	ETT
Lead author	Antonio Novellino
Contributors	AU, CNR
Contributing authors	Rachele Bordoni, Elmer Topp-Jorgensen, Jonas K. Romer, Vito Vitale
Due date	31/05/2025
Delivery date	12/09/2025

Document history	
Creation Date	01/03/2025
Version	V1
Version Date	12/09/2025
Status	☐ Draft ☒ WP lead approved ☒ Coordinator approved ☒ Executive Board approved
Status date	12/09/2025

TABLE OF CONTENTS

1. SUMMARY	4
2. Open Science	5
2.1 Open Science under Horizon Europe	7
2.2 Open Access to Publications	8
2.3 Open Science in POLARIN	9
3. FAIR Principles and FAIR Guidelines	11
3.1 FAIR (Data) implementation in European Marine Research Projects.....	14
4. References.....	16

1. SUMMARY

The polar regions are experiencing rapid, large-scale changes with global implications. Accelerated ice loss, shifting oceanic and terrestrial systems, and complex interactions with ecosystems and climate patterns underscore the urgency of robust, high-quality scientific data. Yet, the Arctic and Antarctic's remoteness, extreme environmental conditions, and limited infrastructure significantly challenge consistent data collection and sharing.

In this framework, the Horizon Europe POLARIN project is grouping and coordinating 64 world-class polar research infrastructures (RIs) across both hemispheres, contributing in-situ and long-term monitoring datasets across land, sea, and ice domains. POLARIN is committed to open science, excellence in research, and high societal impact and its core services include:

- challenge-driven Transnational Access (TA);
- enhanced data services and products;
- virtual Access (VA) to data and tools;
- training and support for users and infrastructure providers.

At the heart of POLARIN's data policy is a commitment to FAIR principles ensuring that all digital research outputs are Findable, Accessible, Interoperable and Reusable.

This document, Guidance Materials for FAIR Principles, offers a broader and more descriptive discussion of Open Science and the FAIR principles (Findable, Accessible, Interoperable, Reusable), which are conceptually interconnected. In addition to give an overall explanation on FAIR principles and Open Science in accessible terms, this document provides a practical introduction to implementing both in the context of the project. It highlights concrete steps, tools and best practices to support researchers in aligning with these standards.

While this document offers a practical guidance, partners have to:

- refer to the DMP for specific policies on formats, vocabularies, metadata and licensing;
- contact the data management team (WP4) when uncertainties arise;
- use FAIR assessment tools before sharing data;
- ensure that data remains open "as much as possible, as restricted as necessary".

2. Open Science

Reliable science is inherently a collaborative endeavor and Open Science aims to ensure that every stage of the research process is conducted with maximum transparency, inclusivity and accessibility. As outlined in the UNESCO Recommendation on Open Science (2021), this international movement is dedicated to making scientific knowledge openly available, accessible and reusable by all, while actively fostering collaboration for the collective advancement of science and society.

At its core, Open Science functions as a guiding paradigm underpinned by transparent research practices, the use of open-access tools and the promotion of collaborative approaches and innovative social partnerships. It also relies on the development of robust infrastructures and the strengthening of research capacities across communities. This framework supports the free and wide dissemination of knowledge, ensuring that research results can be accessed and reused efficiently. It advocates for open access not only to scientific publications but also to data and promotes engagement in open peer review processes. Importantly, Open Science encourages meaningful dialogue with civil society and the integration of diverse knowledge systems.



Image adapted from UNESCO, Digital Library. *Towards a UNESCO recommendation on open science: building a global consensus on open science*

Figure 1. Open Science loop – adapted from UNESCO Recommendation on Open Science, 2021

By promoting new methods of working and establishing incentive structures that reward openness and cooperation, Open Science fosters a more inclusive, trustworthy and sustainable scientific ecosystem: one that can better respond to global challenges through shared knowledge and collective innovation.

A detailed taxonomy, developed by the FOSTER (Facilitating Open Science Training for European Research) project, provides a comprehensive map of the Open Science landscape. This taxonomy is instrumental for understanding the multifaceted components of Open Science and how they align with FAIR principles (Findable, Accessible, Interoperable, Reusable).

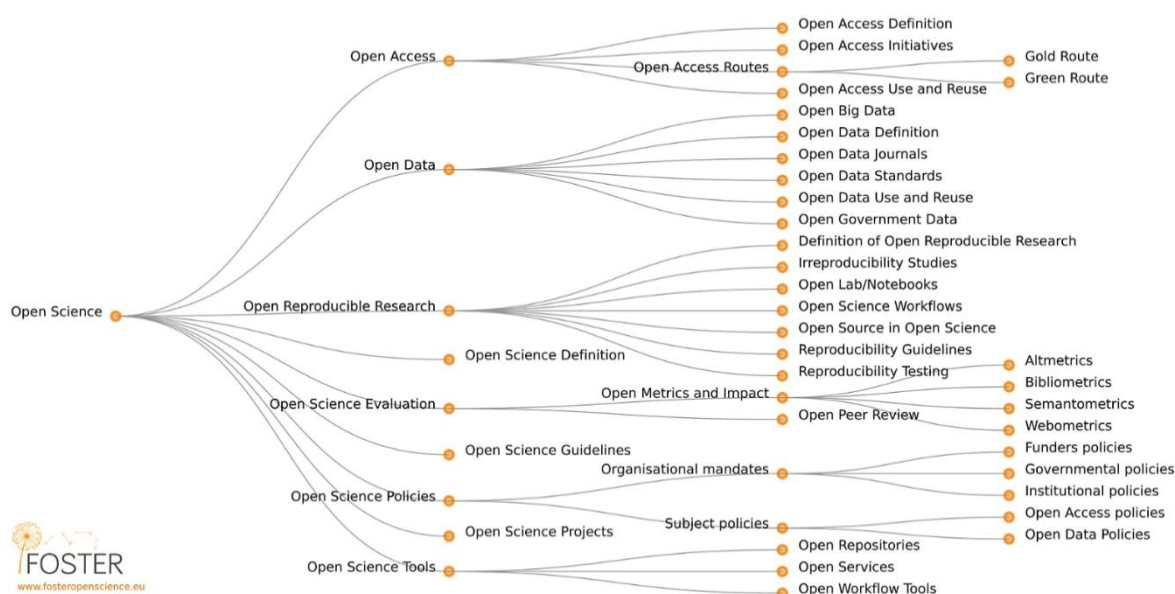


Figure 2. FOSTER - adapted from Knoth and Pontika, 2015¹.

FOSTER identifies the following key domains:

- Open Access: research outputs, such as publications, should be freely accessible, and researchers have to adopt the Gold/Green initiatives.
- Open Data: data availability and transparency is key for any open research activity. Open Data supports FAIR where (primary) data should be findable, available, citable and usable through trusted and well-documented repositories.
- Open Reproducible Research: encourages the adoption of resources like open lab notebooks, workflows and reproducibility testing, cornerstones for research integrity and replication.

¹ https://figshare.com/articles/figure/Open_Science_Taxonomy/1508606/3?file=2250225

- **Open Science Evaluation:** transparency in impact assessment. It supports the definition of key impact indicators, the use of clear metrics (e.g., altmetrics, webometrics) to ensure clear measures of research success.
- **Open Science Policies and Guidelines:** includes the clear description of the needs and settings for implementing compliance with FAIR principles. It includes the adoption of organizational mandates, funder requirements and governmental/institutional frameworks, the design and implementation of open science (data) management plan, the precise and clear reporting of outcomes;
- **Open Science Tools and Projects:** comprise digital infrastructure such as repositories, workflow tools, and services that enable FAIR-compliant data management and sharing.

2.1 Open Science under Horizon Europe

Open Science is both a foundational principle and a legal requirement of Horizon Europe. The program promotes Open Science as a strategy to increase transparency, collaboration, innovation and societal impact in publicly funded research across Europe.

Compliance with Open Science Policies

Projects funded by Horizon Europe must adhere to several policies:

- **Open Access to Publications:** all peer-reviewed scientific publications must be made openly accessible immediately. This can be achieved through publishing in Open Access journals (Gold OA), depositing manuscripts in repositories (Green OA), or using compliant platforms. Open Access increases visibility, citation impact, and public benefit;
- **Research Data Management and FAIR Principles:** data must be Findable, Accessible, Interoperable, and Reusable (FAIR). Projects must submit a Data Management Plan (DMP) detailing how data will be collected, stored, shared, and preserved, addressing metadata, legal and ethical compliance, and stewardship. Data that cannot be openly shared due to privacy or security concerns must still be FAIR to the extent possible, with access conditions clearly defined;
- **Open Research Outputs beyond publications:** researchers are encouraged to openly share software, models, protocols, and workflows using open licenses (e.g., Creative Commons, GPL) to clarify reuse rights and support reproducibility and innovation. This includes promoting open-source software, open peer review and active stakeholder engagement, including citizen science.

Transparency, reproducibility and inclusivity

Horizon Europe supports transparent research practices such as open peer review, preregistration of studies, and sharing of negative or inconclusive results to strengthen scientific reliability. The program values inclusive, responsible research involving citizens, industry, policymakers, and interdisciplinary teams, recognizing public engagement as key to fostering trust and societal relevance.

Support and enforcement

The European Commission provides extensive resources—templates for Data Management Plans, training, guidelines on repositories and licensing, and infrastructure like the EOSC to assist researchers in meeting Open Science requirements. Non-compliance with Open Access or FAIR data mandates may affect grant reporting and future funding, highlighting the importance of integrating Open Science early in project planning and execution.

2.2 Open Access to Publications

Publications resulting from Horizon Europe projects must comply with specific open access requirements. Researchers are required to deposit either the Author Accepted Manuscript (AAM) or the Version of Record (VoR) in a trusted repository without any embargo period, thereby ensuring immediate open access. Authors retain their rights, as no Copyright Transfer Agreement is permitted, and publications must be made available under a CC BY 4.0 license. Furthermore, all outputs should include information on research data, tools, or instruments necessary to validate the publication's conclusions, as well as the project acronym or grant code in the acknowledgments section. Regarding costs, Article Processing Charges (APCs) are reimbursable when publishing in full open access venues, while hybrid journal APCs are not eligible. Finally, although the CC BY 4.0 license is standard, CC BY-NC or CC BY-ND licenses are allowed for long-text formats such as monographs, though book chapters are not eligible for reimbursement.

Horizon Europe Open Access Mandate

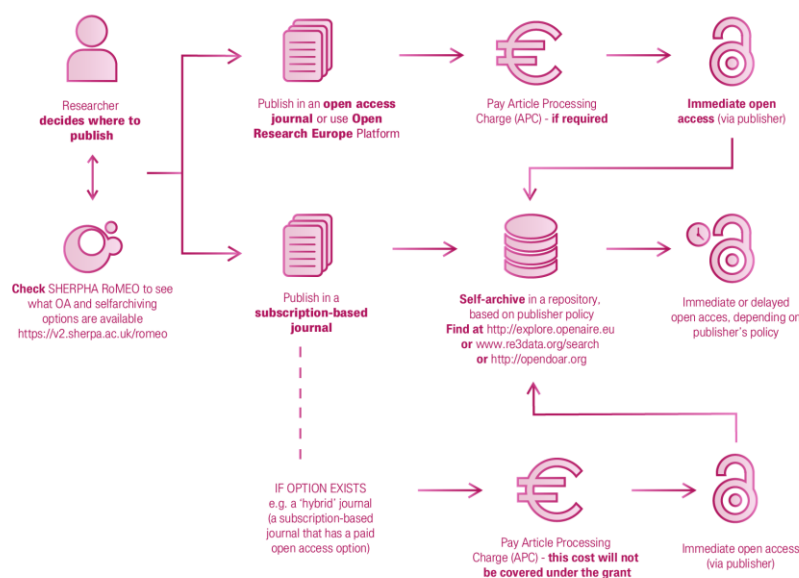


Figure 3. Horizon Europe Open Access Mandate

There are some tools that help a self-archiving for open access:

- Journal's Policy checker - Sherpa Romeo (www.sherpa.ac.uk/romeo)
- Rights Retention - <https://www.coalition-s.org/rights-retention-strategy/>

2.3 Open Science in POLARIN

POLARIN extends the concept and requirements for Open Access, Open Data, Open Tools, and Reproducible results by recommending the use of trusted data repositories that ensure long term stewardship of the data and that POLARIN contribute to main European marine data integrators (and EMODnet in particular).

Being a Horizon Europe project POLARIN designed to advance scientific understanding of polar regions by enabling access to RIs and addressing existing fragmentation in polar data access and improve the way polar research data is shared, managed, and used, POLARIN adopts and fosters the Horizon Europe recommendations for harmonizing data services and open science practices (Figure 4).

Open Science

“Open science is an approach based on open cooperative work and systematic sharing of knowledge and tools as early and widely as possible in the process”

European Commission



Figure 4. EC recommendations for Open Science

Digital research data generated within Horizon Europe projects must be managed in accordance with the FAIR principles. A Data Management Plan (DMP) is mandatory and must be submitted by month six of the project, with updates required at mid-term and at the project's conclusion. Data and metadata should be deposited as early as possible, either immediately after production or following processing and quality control, in trusted repositories. In line with the principle of being “as open as possible, as closed as necessary,” datasets should be made openly available by default, with any restrictions duly justified in the DMP.

Even when data cannot be openly shared, metadata must remain FAIR and released under a Creative Commons - CC0 license. All open datasets should be published under an open license, preferably CC-

BY or CC0. Notably, Creative Commons (CC) licenses are widely used to facilitate sharing and reuse of data in Open Science. They provide a clear and standardized way to communicate the permissions granted by the data creator. CC for data include:

- CC0 (Public Domain Dedication): this license allows creators to waive all rights and place their data in the public domain. Users can freely copy, modify and distribute the data without any restrictions or the need for attribution. CC0 is highly recommended and often required for Open Science because it maximizes data accessibility, reuse and interoperability;
- CC BY (Attribution): this license lets others use, modify and share the data, even commercially, as long as they give appropriate credit to the original creator. It supports openness while ensuring proper attribution;
- CC BY-SA (Attribution-ShareAlike): similar to CC BY, but requires that any derivative works or adaptations be shared under the same license terms. This “share alike” condition helps keep data openly available in all downstream uses;
- CC BY-NC (NonCommercial) or CC BY-ND (NoDerivatives) may restrict commercial use or modifications, which can limit the openness and reuse of data in scientific research.

In addition, sufficient information about research outputs, tools, and instruments—such as software, algorithms, protocols, models, workflows, or electronic notebooks—must be provided to enable the validation, reuse, and reproducibility of the results.

As already mentioned, one specific chapter of the Open Science is the **Open Data Access**. Open data access ensures that research data outputs are freely accessible online to anyone without subscription or paywalls. It increases the visibility, impact and reuse potential of research, promoting collaboration and transparency. By sharing data openly, researchers support the FAIR principles.

POLARIN practices Open Science and Open Data effectively, indeed it:

- early and open sharing of research (GA papers, newsletters, Zenodo)
- Uses open access journals² or book³
- develops and updates a DMP
- uses open standards for data formatting and metadata to ensure interoperability;
- deposits data in trustworthy open repositories or journals that support data publication;
- applies clear licensing, Creative Commons (CC), to remove ambiguity about data reuse;
- documents datasets comprehensively to allow other researchers to understand and use the data effectively;
- registers datasets in repositories that provides DOI and long-term stewardship of the data
- links datasets with related publications to maximize discoverability and impact.
- Cites the grant agreement to award funding agency and tracks the impact of the project
- Involves all relevant stakeholders and known users

² <https://doaj.org/>

³ <https://www.doabooks.org/>

3. FAIR Principles and FAIR Guidelines

The FAIR principles—Findable, Accessible, Interoperable, and Reusable—were introduced in 2016 to increase the value and reusability of scientific data, particularly in digital and machine-actionable environments. They serve as a cornerstone for effective data stewardship and complement Open Science by providing practical guidance for managing and sharing data.

The FAIR principles are defined through 15 detailed characteristics established by the FORCE11 collective (<https://force11.org/info/the-fair-data-principles/>) and further outlined by Quimbert et al. (2022). These recommendations emphasize three essential components for ensuring interoperability: metadata, data formats and data services. It is important to clarify that FAIR data is not synonymous with Open Data. While both aim to support broader data usability, FAIR refers to how data should be managed and structured for reuse—not necessarily that it must be openly accessible. Data can be FAIR without being openly shared when legal, ethical, or commercial restrictions apply. The focus of FAIR is on enabling data discovery and interoperability, even within secure or restricted environments.

There are several reasons why data may be restricted, including the protection of personal data and privacy (such as under GDPR), intellectual property rights like patents and copyrights, confidentiality concerns, human rights or cultural sensitivities, and the protection of endangered species or sacred Indigenous knowledge. Even when data is restricted, its metadata can remain publicly accessible. This allows the data to be discoverable and enables interested parties to contact the authors. This demonstrates that Open Science and intellectual property rights can coexist harmoniously (European Commission, 2022).

Principle	Description	How to implement:
F - Findable	Data should be easy to discover for both humans and machines. This requires the description and indexing of data and metadata. To do so, you should assign persistent identifiers (such as the Digital Object Identifier - DOI - and ORCID iD), provide rich metadata, use keywords, and use versioning of data sets.	• (Meta)data are assigned globally unique and persistent identifiers (e.g., DOI, ORCID); • Data are described with rich metadata; • Metadata clearly includes the identifier of the described data; • (Meta)data are registered or indexed in searchable databases.
A - Accessible	Data should be retrievable using standard, open protocols. Metadata should remain accessible even when the data is no longer accessible, so people, institutions or publications can be tracked down	• (Meta)data are retrievable using a standardized protocol; • The protocol is open, free and universally implementable; • The protocol supports authentication and authorization, if needed;

		• Metadata remain accessible even if the dataset itself is no longer available
I - Interable	Data should be able to integrate with other datasets and systems. The data needs to interoperate with applications and workflows for analysis, storage, and processing. You should use formal, broadly applicable language, and use standard vocabularies and qualified references.	• (Meta)data use a formal, accessible, shared and broadly applicable language; • Use of vocabularies that comply with FAIR principles; • Include qualified references to related (meta)data.
R - Reusable	Data should be well-described and licensed to allow reuse. You should use a licence to permit the widest reuse possible; specify a data embargo, if needed, and describe how long the data will remain reusable and describe data-quality-assurance processes	• Data must be richly described with accurate and relevant metadata; • Clearly stated licenses for data use (e.g., Creative Commons); • Metadata must include provenance information; • Conformity with community standards within relevant domains.

The FAIR principles guide the effective sharing and management of research data. To foster their knowledge and adoption of FAIR practices, researchers are invited to use tools like:

- FAIR-Aware (<https://fairaware.dans.knaw.nl>) - An online tool that assesses your knowledge of FAIR principles. It also helps you understand how making your data (set) FAIR can increase the potential value and impact of your data and helps you understand how to apply them to your data.
- F-UJI FAIR Data Assessment Tool (<https://www.f-ujl.net>) - A free, open-source service that evaluates how well a dataset meets FAIR principles. Just input a dataset's DOI or URL and the tool provides a detailed score and suggestions for improvement. It also supports integration into research workflows.

In POLARIN, researchers are expected to integrate these principles throughout the data lifecycle, from planning and collection to archiving and publication in line with POLARIN's Data Management Plan (DMP). More specifically, FAIR data policy POLARIN requires:

1. Use of Persistent Identifier (PID) – to ensure that project and contributors remain identifiable and citable. This includes:
 - a. use DOIs (Digital Object Identifiers) for datasets;
 - b. link researcher profiles with ORCID iDs;
 - c. when possible, also use ROR IDs and/or EDMO codes for institutions and other PIDs for software or funding references.

2. Use exhaustive metadata using standards - metadata is essential to make data understandable, reusable and interoperable. This includes:
 - a. POLARIN encourages use of ISO 19115 metadata schema, including:
 - i. What - title, abstract, keywords;
 - ii. Where - spatial coverage (geographic bounding box);
 - iii. When - temporal extent (ISO 8601);
 - iv. Who - creator/contact information.
 - b. Ensure metadata includes:
 - i. licensing terms;
 - ii. persistent identifiers;
 - iii. provenance (methodology, tools, version).
3. Use controlled vocabularies and classification systems – that ensure consistent interpretation and data integration. This includes:
 - a. EMODnet, INSPIRE, CF conventions;
 - b. ISO thematic vocabularies;
 - c. GCMD keywords for environmental sciences.
 - d. Use of common data formats
 - i. NetCDF, CSV, JSON, XML, TXT, GeoTIFF, Shapefile;
 - ii. OGC WMS, WFS, WCS
 - e. Use of open data management systems
 - i. ERDDAP, GeoNetWork, GeoServer
4. Use versions control policy - data is rarely static, record changes and maintain transparent history. This includes:
 - a. use version numbers for datasets and metadata;
 - b. maintain clear logs (e.g., changelogs or Git-based records);
 - c. document provenance: origin, data collection methods, transformations, software used.
5. Use clear data license according Common Creative.
 - a. default license for metadata: CC0
 - b. default license for metadata: CC-BY
6. Limit the use of embargos⁴ and when applied, provide clear description of:
 - a. embargo end date – which anyhow should be before the end of the POLARIN project

⁴ according to Open Science Practice in Horizon Europe, this means that while beneficiaries are expected to do their utmost to ensure that their project will contribute to the EU's ambitions for open science, it recognises that there are certain cases in which making such information or data available in open access is not desirable – or indeed not legal

- b. reason for restriction (that can be 1) Legitimate interest of the beneficiary, including commercial exploitation; 2) Data protection, including personal data protection; Other constraints such as Union competitive interests, security rules or any other obligation under the Grant Agreement) – which should be clearly described
- c. contact information;

Notably, to move towards this Open Science approach, POLARIN set up a Zenodo Community⁵.

Zenodo is an open-access repository developed by CERN and enable researchers to share and preserve a wide range of research outputs. It accepts publications, datasets, software, presentations, and other research materials from all disciplines. Zenodo assigns DOIs (Digital Object Identifiers) to uploaded items, making them citable and discoverable. Zenodo is free to use and promotes open science by enabling unrestricted access to research results. Zenodo is supported by OpenAIRE (Open Access Infrastructure for Research in Europe) which is a European initiative (funded by European Commission) that provides infrastructure, services, and guidance to support open science and the free sharing of research outputs. OpenAIRE plays a key role in implementing the European Open Science Cloud (EOSC) and promotes the FAIR principles. Moreover, OpenAIRE is directly interconnected with the CORDIS (Community Research and Development Information Service) which is the European Commission's primary public repository for information on EU-funded research projects. The native interoperability between Zenodo, OpenAIRE and CORDIS make all the Zenodo results immediately available and shareable in the CORDIS project page promoting open and free science across Europe and beyond.

Although the POLARIN Zenodo is preliminary designed to manage the official project documents, primarily deliverables, after they are approved by the EC reviewers, it also provides the project an easy and powerful infrastructure to support and facilitate the storage of selected datasets (notably, adopting Open Data and Open Science policies, POLARIN Zenodo as a storage allowance of up to 200GB per item), where data deposit is controlled by the Data Manager and the Steering Committee to ensure completeness and validity.

Zenodo community will be also used in combination with software repositories services to support the full data value chain FAIRness.

3.1 FAIR (Data) implementation in European Marine Research Projects

In marine research, adhering to the FAIR principles involves not only proper data management but also the strategic use of established European infrastructures. For projects under Horizon Europe, such as POLARIN, this includes depositing data and metadata into recognized platforms to ensure long-term accessibility and interoperability. These platforms include National Data Centers or National Databases (e.g., PANGAEA in Germany, the National Antarctic Data Center and the Italian Arctic Data Center in Italy, and the SIOS Svalbard Integrated Observation System). For ocean-related projects, interoperability with the EU marine data infrastructure and the EU Digital Twin of the Ocean is also

⁵ <https://zenodo.org/communities/eu-polarin/records?q=&l=list&p=1&s=10&sort=newest>

required. More specifically, it is essential that all newly recorded data be made interoperable and accessible through the European Marine Observation and Data Network (EMODnet)⁶, which is the designated platform of the European Commission for in situ data. EMODnet serves as the central repository for marine data, ensuring that datasets meet EU standards for openness and usability. EMODnet is organized in thematic areas (Physics, Chemistry etc.) and operates a Data Ingestion platform to facilitate the interoperability and discoverability of data from federated sources.

While EMODnet Physics serves as the central repository for ensuring that marine environmental data are FAIR within Horizon Europe, in line with the Open Science policy, POLARIN must demonstrate full FAIRness towards its stakeholders.

To support the coordinated use of shared marine research resources such as vessels and expeditions, POLARIN may contribute to DUE South⁷, a platform developed by the Southern Ocean Observing System (SOOS) that maps and provides visibility of upcoming research cruises and strengthens international collaboration, reduces duplication of effort, and aligns with FAIR principles in operational planning.

Equally important is the proper documentation and dissemination of data collection methodologies. The European Commission's Ocean Best Practices and Standards System (EC-OBPSS)⁸ provides a repository for standardized protocols used during field activities. By publishing its data collection protocols on EC-OBPSS.

⁶ <https://emodnet.ec.europa.eu>

⁷ <https://soos.aq/activities/duesouth>

⁸ <https://ecobpss-pre.ctnaval.com/home>

4. References

- de San Román (2021). Open Science in Horizon Europe. Train-the-trainer workshop. Zenodo. CC-BY 4.0 <https://doi.org/10.5281/zenodo.5549524> Directory of Open Access Books (DOAB) – <https://www.doabooks.org>
- Directory of Open Access Journals (DOAJ) – <https://doaj.org>
- DUE South (SOOS) – <https://soos.aq/activities/duesouth>
- EMODnet (European Marine Observation and Data Network) – <https://emodnet.ec.europa.eu>
- EMODnet Physics – <https://emodnet.ec.europa.eu/en/physics>
- European Commission's Ocean Best Practices and Standards System (EC-OBPSS) – <https://ecobpss-pre.ctnaval.com/home>
- F-UJI FAIR Data Assessment Tool – <https://www.f-uji.net>
- FAIR-Aware: Self-assessment Tool for FAIR Data – <https://fairaware.dans.knaw.nl>
- FORCE11. (2016). The FAIR Data Principles – <https://force11.org/info/the-fair-data-principles/>
- Guidelines for uploading research outputs to OpenAIRE and the use of POLARIN Zenodo Community – (link not provided; please add if available)
- INTERACT (International Network for Terrestrial Research and Monitoring in the Arctic) – <https://eu-interact.org>
- Knoth, P., & Pontika, N. (2015). Open Science Taxonomy. Figshare – <https://doi.org/10.6084/m9.figshare.1508606.v3>
- Ocean Prediction UN Decade Collaborative Centre (OceanPrediction-DCC) – <https://www.unoceanprediction.org/en/homepage>
- ORCID – <https://orcid.org>
- Quimbert, E., Fichaut, M., & Maudire, G. (2022). Guide principes FAIR – Principes FAIR dans le contexte du pôle ODATIS. <https://doi.org/10.13155/87107>
- Southern Ocean Observing System (SOOS) – <https://soos.aq>
- Sustaining Arctic Observing Networks (SAON) – <https://www.arcticobserving.org>
- UNESCO. (2021). UNESCO Recommendation on Open Science – <https://en.unesco.org/science-sustainable-future/open-science/recommendation>