



POLARIN

POLAR
RESEARCH
INFRASTRUCTURE
NETWORK

Deliverable 4.1. Data Management Plan for generated and collected data incl. ethical guidelines for data and derived data products

V2, September 2024

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	D4.1
Deliverable title	Data Management Plan for generated and collected data incl. ethical guidelines for data and derived data products
Version	V2.0
Dissemination level	<input checked="" type="checkbox"/> PU - Public <input type="checkbox"/> PP - Restricted to programme partners <input type="checkbox"/> RE - Restricted to a group specified by the consortium <input type="checkbox"/> CO - Confidential, only for members of the consortium
Lead Beneficiary	ETT
Lead author	A Novellino, V Vitale
Contributors	ETT, CNR, SIOS, NILU, AU
Contributing authors	A Guida, G Dapuelto, G Verrazzo, C Ripa, D Kivits, I Matero, O Godoy H Savela J Koefoed Romer
Due date	31/08/2024
Delivery date	

Document history	
Creation Date	01/07/2024
Version	V2.0
Version Date	11/09/2024
Status	<input type="checkbox"/> Draft <input checked="" type="checkbox"/> WP lead approved <input checked="" type="checkbox"/> Coordinator approved <input checked="" type="checkbox"/> Executive Board approved
Status date	27 September 2024

Table of Contents

SUMMARY	4
1. Introduction.....	5
2. POLARIN Research Infrastructures.....	6
2.1. POLARIN Data Policy approach	9
3. Making POLARIN data FAIR	10
3.1. POLARIN (meta)data model	13
3.2. POLARIN (meta)data catalogues	13
3.3. POLARIN Data License	14
3.4. POLARIN Data Quality policy.....	14
3.5. POLARIN Data Management and TA projects.....	15
4. Open Science and other Research outputs	16
4.1. Scientific communications	16
4.2. Other publishable results	16
4.3. Software tools and code	17
5. Allocation of resources.....	17
5.1. Costs before making data FAIR and long-term preservation.....	17
5.2. POLARIN Data Management officer.....	17
6. Data Security.....	18
7. Ethics and GDPR.....	18
7.1. Ethics and Code of Conduct	18
7.2. Personal Data Regulation.....	18
8. References.....	20
Acknowledgements	20
Annex 1 – POLARIN Metadata Model	21
Annex 2 - DMP template for TA.....	26
Scope of the Transnational Access Data Management Plan (TA DMP)	26
PRELIMINARY TA DMP Template	27
Annex 3. Further recommendations/measures on long term preservation of informed consent	29

SUMMARY

Timely and unrestricted exchange of polar data is essential for a variety of purposes, such as weather forecasts and climate projections, climate change understanding etc.

An open and free data policy is highly promoted by the European Commission and its member states for a wide range of environmental data services targeted to a wide range of user communities. Interoperability of data systems has become a priority with the development of FAIR principles¹, *i.e.*, a set of guiding principles to make data **Findable, Accessible, Interoperable, and Re-usable**.

In line with these recommendations, POLARIN aims to make FAIR its digital resources and support its stakeholders (e.g. Transnational Access projects) to be FAIR.

This deliverable presents the first version of the POLARIN Data Management Plan (DMP). Based on the Horizon Europe Data Management Plan template, it covers the key areas necessary for proper and effective data handling throughout the project's lifecycle. Specifically, this DMP introduces the POLARIN data, outlines how existing data will be reused and accessed, how new data will be made available, describes adherence to and promotion of FAIR principles, and details the tools that will facilitate these processes.

The DMP also addresses data sharing protocols, including compliance with GDPR and other relevant regulations, to facilitate accessibility and reuse by the wider research community.

¹ <https://www.force11.org/group/fairgroup/fairprinciples>

1. Introduction

The polar regions are vital components of Earth's system, crucial for regulating our climate and serving as indicators of climate change, human expansion, and resource exploitation. These regions are experiencing rapid ice loss and significant transformations in their oceans and land, with global repercussions that impact people in diverse ways. To ensure sustainable development in the Arctic and effectively protect the Antarctic, evidence-based policy recommendations are necessary. However, the remoteness and inaccessibility of the polar regions, combined with limited research infrastructure, present significant challenges. Research data is often fragmented and dispersed across various databases, lacking sufficient interoperability. To better understand and predict key processes in the polar regions, and to provide the evidence-based information needed to support the European Green Deal and EU Arctic policy, the polar research community requires access to world-class research infrastructure in these areas.

POLARIN is an international network of polar research infrastructures and services designed to address the scientific challenges of the polar regions. This network encompasses a diverse range of top-tier research infrastructures, including Arctic and Antarctic research stations, research vessels, icebreakers operating at both poles, observatories, data infrastructures, and repositories for ice and sediment cores. POLARIN aims to provide integrated, challenge-driven access to these infrastructures, facilitating interdisciplinary research on complex polar processes.

To this end, POLARIN will:

- Offer challenge-driven transnational access to a broad portfolio of research infrastructures.
- Enhance data accessibility by improving data availability and interoperability among data infrastructures.
- Provide virtual access to data and data services.
- Deliver data products for the scientific community and decision-makers.
- Train the next generation of polar researchers to effectively utilize these infrastructures for their research.
- Actively promote the services offered by POLARIN and encourage infrastructure users to share their research outcomes with society.

A comprehensive Data Management policy is central to ensuring that data is effectively collected, managed, and shared, thereby maximizing the impact of research and supporting informed decision-making across the polar regions. This deliverable presents the first version of the POLARIN Data Management Plan (DMP). Based on the Horizon Europe Data Management Plan template, it covers the key areas necessary for proper and effective data handling throughout the project's lifecycle. Specifically, this DMP introduces the POLARIN data, outlines how existing data will be reused and accessed, how new data will be made available, describes adherence to and promotion of FAIR principles, and details the tools that will facilitate these processes.

Document Disclaimers:

This document is based on the Horizon Europe Data Management Plan Template (V1; April 2022) and adapted to provide the POLARIN partners and POLARIN stakeholders a living guide for the project Data Management Plan. POLARIN Consortium is not responsible for any use that might be made of the content of this publication.

2. POLARIN Research Infrastructures

Climate change is more pronounced in the polar regions, with the northern and southern extremes of the planet warming faster than any other area on Earth. The Arctic is warming at twice the rate of the global average. Both the Greenland and Antarctic ice sheets are losing ice mass to the ocean at an accelerating rate, contributing to global sea level rise. Despite these common trends, the polar regions differ significantly in their physical characteristics. The Arctic is an ocean surrounded by land, while the Antarctic is a continent surrounded by ocean. These differences shape the interactions between the cryosphere, atmosphere, ocean, and ecosystems—both terrestrial and marine—and influence their global connections. They also drive the evolution of distinct marine and terrestrial biology in each polar region, affecting the impacts, responses, and adaptations of polar ecosystems. Nevertheless, the polar regions remain difficult to access due to extreme weather conditions and a lack of infrastructure, making research in these areas logistically and financially challenging. Enhancing access to and integration of research infrastructures (RIs) is essential for strengthening European research and improving our observational and modeling capabilities to address the significant challenges facing these regions.

POLARIN has brought together a unique collection of 61 polar research infrastructures (RIs), ranging from small research stations in the Arctic and Antarctic to large icebreakers operating at both poles.

POLARIN focuses on key regions that are particularly vulnerable to climate change, with an emphasis on one or more of the following processes:

1. Sea-ice and Polar Oceans
2. Sea level, glacier stability and melt
3. Carbon balance and permafrost
4. Polar ecosystems and biodiversity
5. Atmosphere dynamics
6. (Paleo)climate processes

In the northern hemisphere, POLARIN offerings cover the wide longitudinal range from Alaska to Fennoscandia, with 4 RIs located in North America, 7 in Greenland, 7 in the Svalbard Archipelago, and an additional 4 RIs located in the Atlantic sector of the Arctic Ocean (1 in the Faroe Islands, 2 in Iceland, and 1 in the Fram Strait). In the southern hemisphere, the focus is on the Antarctic Peninsula (6 RIs) and the Weddell Sea/Dronning Maud Land area (3 RIs). Additionally, Italian RIs offer access to the Ross Sea region and the East Antarctic Plateau. In addition to fixed stations, POLARIN's offerings include 12 vessels/icebreakers made available by research institutions as well as private stakeholders. Finally, 4 core repositories provide access to several thousand meters of ice and sediment samples collected and

archived from both poles. Table 1 gives full details of the available RIs. Acronyms, as well as contributions to challenges, are derived from the submitted proposal.

Table 1. POLARIN RIs

				Contribution to challenges					
	Partner (acronym)	TA or/& VA	Station	1	2	3	4	5	6
Arctic									
1	SPRS	TA/VA	ABISKO		x	x	x	x	x
2	AMU	TA	AMUPS		x	x	x	x	x
3	DTU	TA	ARC-DTU		x	x	x	x	x
4	UCPH	TA	ARCST	x	x	x	x	x	x
5	AWI/IPEV	TA	AWIPEV	x	x	x	x	x	x
6	UICS	TA	BARC	x	x	x	x	x	x
7	ULAVAL	TA/VA	CEN WK	x	x	x	x	x	x
8	CNR	TA/VA	DIR-ITA		x		x	x	x
9	GINR	TA	GINR	x	x	x	x	x	x
10	IGF PAS	TA	HORNSUND	x	x	x	x	x	x
11	UTU	TA/VA	KEVO			x	x	x	x
12	UH	TA/VA	KILPIS			x	x	x	x
13	SAVN	TA/VA	KOLTUR	x			x	x	x
14	UMK	TA	NCUPS	x	x	x	x	x	x
15	UOULO	TA	OULANKA				x	x	x
16	FMI	TA/VA	PAL-SOD				x	x	x
17	DMI	TA	QAANAAQ	x		x	x	x	x
18	RIF	TA	RIF	x			x	x	x
19	UCPH	TA	SER	x	x	x	x	x	x
20	SSLC	TA	SUDURNES				x	x	x
21	NPI	TA	SVENDRUP	x	x		x	x	x
22	UAF	TA	TOOLIK			x	x	x	x
23	SU	TA/VA	TRS		x	x	x	x	x
24	UKRI	TA	US AK	x	x		x	x	x
25	AU	TA	VRS		x	x	x	x	x
26	ARI	TA/VA	WARC	x	x	x	x	x	x
27	AU	TA	ZAC	x	x	x	x	x	x
28	AWI	TA	FRAM	x	x		x		x
Antarctic									
1	NASC	TA	AVS	x	x	x	x	x	x

2	BAI	TA	BAB	x	x	x	x	x	
3	CNR/IPEV	TA/VA	CONCORDIA		x			x	x
4	INACH	TA	ESCUADERO	x		x	x	x	x
5	MCIN	TA	GABRIEL				x	x	x
6	CSIC	TA	JC I				x	x	x
7	CNR	TA/VA	MSZ	x	x	x	x	x	
8	AWI	TA	NEUMAYER	x	x		x	x	x
9	INACH	TA	PRAT	x		x	x	x	x
10	NPI	TA	TROLL	x	x		x	x	x
11	SPRS	TA	WASA	x	x		x	x	x
Vessel & Icebreaker									
1	ULAVAL	TA	AMUNDSEN	x	x		x	x	x
2	MFRI	TA	ARNI	x	x		x	x	x
3	MI	TA	CELTIC	x	x		x	x	x
4	PONANT	TA	CHARCOT	x	x		x	x	x
5	DTU	TA	DANA	x	x		x	x	x
6	NPI	TA	HAAKON	x	x		x	x	x
7	MCIN	TA	HESPERIDES	x	x		x	x	x
8	INACH	TA	KARPUJ	x	x		x	x	
9	OGS	TA	LAURA	x	x		x	x	x
10	NASC	TA	NOOSPHERE	x	x		x	x	x
11	AWI	TA	POLARSTERN	x	x		x	x	x
12	TARA	TA	TARA	x	x		x	x	x
Core Repositories									
1	AWI	TA	AWI ICE						x
2	AWI	TA	AWI SED						x
3	UKRI	TA	PSCF						x
4	UIT/APECS	TA	UIT CORES						x

Access to these RIs, either in person or remotely, is provided through Transnational Access (TA) grants.

In addition to these Research Infrastructures (RIs) – which include 27 Arctic stations on land, 11 Antarctic stations, 12 research vessels, 1 deep-sea observatory, and 4 ice and sediment repositories – providing physical or remote access, POLARIN's offerings are further complemented by virtual access (VA) to large datasets and long-term series of polar data (spanning up to a century). These datasets, which include many relevant parameters for monitoring the status of the climate system, are made available by 12 of the 18 observatories (10 in the Arctic, 2 in Antarctica) as well as 7 data infrastructures. Additionally, virtual access is also offered for the databases of two networks operating in Greenland, GEM and ARC-MO.

Table 2. POLARIN VAs

	Partner (acronym)	TA or/& VA	Project name	Web site
Data infrastructure				
1	CAFF	VA	ABDS	https://abds.is/
2	ETT	VA	ADI	https://arice-h2020.eu/data-tools/
3	CNR	VA	IADC	https://iadc.cnr.it
4	INPA/INKODE	VA	IDP	https://dataportal.eu-interact.org/
5	CNR	VA	NADC	https://iandc.pnra.aq
6	GFZ	VA	POSEDA	http://geofon.gfz-potsdam.de/
7	SIOS	VA	SDMS	https://sios-svalbard.org/metsis/search?f%5B0%5D=dataset_level%3Alevel-1
Greenland Network Database				
	AU	VA	ARC-MO	https://gios.org
	AU	VA	GEM	https://data.g-e-m.dk/
Observatories				
1	SPRS	TA/VA	ABISKO	https://www.polar.se/en/research-support/abisko-scientific-research-station/
2	ULAVAL	TA/VA	CEN WK	https://www.cen.ulaval.ca/en/station.php?id=321&nm=wk
3	CNR	TA/VA	DIR-ITA	https://www.isp.cnr.it/index.php/en/infrastructures/research-stations/dirigibile-italia
4	UTU	TA/VA	KEVO	www.utu.fi/kevo
5	UH	TA/VA	KILPIS	www.helsinki.fi/en/research-stations/kilpisjarvi-biological-station
6	SAVN	TA/VA	KOLTUR	www.savn.fi/nature-of-koltur/
7	UOULO	TA	OULANKA	www oulu.fi/en/university/oulanka-research-station
8	FMI	TA/VA	PAL-SOD	https://en.ilmatieteenlaitos.fi/pallas-atmosphere-ecosystem-supersite
9	SU	TA/VA	TRS	www.su.se/tarfala-forskningsstation/
10	ARI	TA/VA	WARC	www.nwtresearch.com
11	CNR/IPEV	TA/VA	CONCORDIA	www.concordiastation.aq/home-1/
12	CNR	TA/VA	MSZ	www.pnra.aq/stazione-mario-zucchelli

2.1. POLARIN Data Policy approach

All digital research data and outputs generated or supported by POLARIN will be managed in accordance with the POLARIN Data Management Plan (DMP), which adheres to the FAIR principles. Applying these principles ensures that research data is Findable, Accessible, Interoperable, and Reusable throughout the entire data lifecycle.

POLARIN data and research outcomes will be accompanied by metadata. Metadata will be made available as soon as designed, data will be made available as soon as possible and in any case within 24 months after the funded project's completion. The EU recommends that data be "as open as possible, as closed as necessary"; therefore, POLARIN advocates for adopting the Creative Commons framework and using the CC BY 4.0 license for digital data wherever possible. An embargo period is

allowed, during which access to a dataset is temporarily restricted. Metadata will be made available under the CC0 license as soon as possible.

In any case, the trusted data repository where the data will be published should be specified in the metadata. Once the data is ready to be made publicly available, and in any case no later than the end of the declared embargo period, the metadata should include a PID (e.g. a DOI), and this information must be provided to the POLARIN catalogue for project records.

Requests from external users for data access during the embargo period will be forwarded to the data originators for their decision. POLARIN, through its data management system, will ensure that datasets are interoperable and accessible, in full compliance with an open data policy.

Moreover, whenever possible, POLARIN will interoperate and federate digital data with key international data infrastructures. For example, for marine data, POLARIN will support methods and tools to ensure interoperability with the European Marine Data Observation Network (the EU's hub for in situ data) through its federated data assembly and integration nodes (e.g., PANGAEA, SIOS), as well as the Copernicus Marine Service, the EU Digital Twin of the Ocean, GEOSS, and EOSC.

This approach maximizes the project's data outcomes and facilitates the use of the data beyond the project's scope.

3. Making POLARIN data FAIR

Applying the FAIR principles means making the research data Findable, Accessible, Interoperable and Reusable throughout the entire data lifecycle.

These recommendations are made according to the FAIR principles broken down into the 15 characteristics laid down by the FORCE11² collective and clearly described in Quimbert et al. (2022)³

Table 3. FAIR principles proposed by the FORCE11 community

FAIR principles proposed by the FORCE11 community		POLARIN strategy
Findable		
F1	(meta)data are assigned a globally unique and eternally persistent identifier.	POLARIN is developing a common metadata catalogue on top of its federated data infrastructure. POLARIN requires each data entry added to the catalogue to contain a PID in its metadata. This may also be a DOI associated to the data record.
F2	data are described with rich metadata.	POLARIN is adopting a common (meta)data model based on community standards. These

² <https://force11.org/info/the-fair-data-principles/> FORCE11 is a community of scholars, librarians, archivists, publishers and research funders that has arisen organically to help facilitate the change toward improved knowledge creation and sharing.

³ Quimbert et al., 2022

		metadata are published in the POLARIN catalogue that supports ISO 19115 schema and common harvesting protocols (OAI-PMH, CSW, OpenSearch, etc..).
F3	metadata specify the data identifier.	POLARIN requires each data entry added to the catalogue to contain a PID in its metadata. This may also be a DOI associated to the data record.
F4	(meta)data are registered or indexed in a searchable resource.	The POLARIN metadata catalogue will support indexing and searching of POLARIN resources.
Accessible		
A1	(meta)data are retrievable by their identifier using a standardized communications protocol.	POLARIN is promoting the use of community-trusted open tools to facilitate metadata and data retrieval for both human and machine interaction. In addition, POLARIN is developing a common catalogue (based on GeoNetWork, GeoServer, and ERDDAP™) that implements widely adopted and standardized communication protocols
A1.1	the protocol is open, free, and universally implementable.	POLARIN is adopting common standards for offering metadata, e.g. OAI-PMH, OGC standards for metadata, and data, e.g. HTTP(s), OPeNDAP, OGC, etc. Notably, the EU recommends having data as open as possible, as closed as necessary"; hence, POLARIN recommends adopting the Creative Commons schema and using the CC BY 4.0 license on digital data wherever possible. An embargo on data is permitted, which is the period during which access to a dataset is temporarily restricted. Metadata will always be made available as soon as possible and under CC0.
A1.2	the protocol allows for an authentication and authorization procedure, where necessary.	In POLARIN all the data will be publicly available, anyhow, as already anticipated, some data may require an embargo and to access these data, in some cases, authentication may be required.
A2	metadata are accessible, even when the data are no longer available.	POLARIN's legacy for metadata accessibility is based on the progressive adoption of common best practices for (meta)data publication by the federated infrastructure. This approach is essential for implementing the POLARIN catalogue, which will be maintained for at least three years after the project's completion. It will also help backend-connected RIs' data infrastructures update their metadata catalogue tools to comply with the latest international recommendations. Adopting FAIR best practices at the level closest to data production makes maintaining FAIR principles more sustainable.

Interoperable		
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	POLARIN adopts international best practices and vocabularies to empower interoperability with its stakeholders. Notably, POLARIN has established open collaboration with key representatives of EU marine data infrastructures and programs. This collaboration also facilitates the two-way implementation and adoption of formal, accessible, shared, and broadly applicable language for knowledge representation. It also includes the adoption of common (meta)data transport formats.
I2	(meta)data use vocabularies that follow FAIR principles.	POLARIN supports research aimed at predicting key processes in polar regions in the context of climate change. These processes range from sea-ice-atmosphere interactions and glacier stability and melt, to ecosystems and biodiversity, and (paleo)climate processes. POLARIN is networking various research domains that have developed and consolidated domain-specific vocabularies. POLARIN adopts these common bases and extends these standards whenever needed to facilitate even broader adoption.
I3	(meta)data include qualified references to other (meta)data.	The POLARIN metadata model is subject to continuous evolution and updates throughout the project lifecycle. The POLARIN DMP serves as a tool to track these changes and updates.
Re-usable		
R1	(meta)data have a plurality of accurate and relevant attributes.	POLARIN (meta)data model defines the minimum set of the relevant attributes to be used.
R1.1	(meta)data are released with a clear and accessible data usage license.	The EU recommends making data "as open as possible, as closed as necessary." Accordingly, POLARIN recommends adopting the Creative Commons framework and using the CC BY 4.0. An embargo is permitted, during which access to a dataset is temporarily restricted. For new data collection supported by POLARIN, metadata must be made available under the CC0 license as soon as possible.
R1.2	(meta)data are associated with their provenance.	POLARIN metadata model includes specific fields to provide POLARIN stakeholders full details for data provenance.
R1.3	(meta)data meet domain-relevant community standards.	POLARIN (meta)data model is adopting and promoting domain-specific community standards.

3.1. POLARIN (meta)data model

If we refer to the ISO standards for cataloguing the information the key references are:

- ISO 8601 Representation of date and time
- ISO 19108 Temporal characteristics of geographic information
- ISO 19113 revised by 19157 standards for geographic information
- ISO 19115 Geographical information metadata
- ISO 19119 Taxonomy of services
- ISO 19139 Geographical information metadata implementation specification - XML implementation of ISO19115

ISO 19113 defines quality principles, which are applied in ISO 19115 (geographic metadata). The metadata records in the current GEOSS use the ISO 19115 data model and its companion XML encoding (ISO 19139). ISO 19115 Standard requires a basic minimum number of metadata elements that are essential for the data presentation:

- Dataset or dataset series on specific challenges ('what'),
- Geographic bounding box ('where'),
- Temporal extent ('when'),
- Contact point to learn more about or order the dataset ('who').

Obviously, additional elements increase (re)usability. The adoption of ISO standards and the use of shared controlled vocabularies are a key prerequisite towards consistency and interoperability.

POLARIN DMP recommends the adoption of the common vocabularies developed by its federated networks, hence e.g. it supports and promotes the adoption of the EMODnet standards as well as some more general ISO standards for e.g. time and datum. In general, any POLARIN related data should include metadata related to the following three categories:

- Discovery and identification
- Spatial and temporal dimension
- Citation and traceability

Annex 1 – POLARIN Metadata Model presents the current version of the metadata model for the POLARIN catalogues.

3.2. POLARIN (meta)data catalogues

GeoNetWork - GeoNetwork offers a web interface that enables users to search geospatial data across multiple catalogues. The search offers the option of a full-text search as well as a faceted search on keywords, resource types, organisations, scale, and so on. The catalogue is designed to accommodate a range of data types, including geospatial layers, services, maps, and non-geographic datasets. GeoNetwork has been designed to implement a number of standards, including WxS, OGC, ISO 19115/119/110, which are commonly used for spatial resources. It has also been developed to align with the Dublin Core format, which is typically employed for open data portals

Geoserver - GeoServer is an open source tool that implements several Open Geospatial Consortium protocols including Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS) and Web Map Tile Service (WMTS) and that was lately updated with the INSPIRE module.

ERDDAP™- The ERDDAP data server is open-source software written in Java that builds upon the open-source ideals of the OPeNDAP, WCS, SOS and OBIS standards. The ERDDAP data server is compatible with a number of commonly used data file formats, including HTML table, NetCDF, CSV, TXT, MAT, JSON and more. Output files are generated on the fly in any of these formats. Additionally, the ERDDAP server implements the FGDC Web Accessible Folder (WAF) with FGDC STD 001 1998 and ISO 19115 WAF with ISO 19115-2/19139 standards.

GOOS Observation Coordination Group (OCG), is encouraging the use of ERDDAP as a potential solution to enhance the interoperability of global ocean datasets.

POLARIN (meta)data MapViewer - In addition to these data cataloging tools, POLARIN is also developing a central map viewer, designed as a user-friendly tool to help discover POLARIN's available resources. This tool will integrate and build upon the legacy of several predecessor initiatives to POLARIN, such as the H2020 ARICE project, the H2020 INTERACT project, and the H2020 EuroFLEET project.

3.3. POLARIN Data License

POLARIN is an EU-funded research project that includes already collected data (provided by some RIs) and new data to be collected during the project. All data must be freely available to the community without cost or restriction. Data licence is a key to accessibility, interoperability and re-use of the data. The data licence should consider:

- provide open and free access to the data, where possible. Note that this access may be through authorisation or authentication, if necessary.
- provide a standardised way for the data actor (creator) to grant permission to use their work under copyright.
- be clear and accessible to the user or data actor and machine-readable.

The Creative Commons (CC) licence summarises these characteristics. It lists 6 different licence types, from the most permissive to the least, with the common point that the creator must be credited. The most permissive: CC-BY (with the only restriction that credit must be given to the creator) should be preferred, following the principle of "as open as possible, as closed as necessary".

An embargo period may also be applied. An embargo is a period of time during which access to a dataset is temporarily restricted. Typically, embargoes are applied while researchers are awaiting publication or pursuing a patent. Typical embargo periods range from 6 to 24 months from the date of data collection.

Metadata will be open under CC 0 or equivalent (to the extent that legitimate interests or constraints are protected), in accordance with the FAIR principles and provide information about the licensing terms and persistent identifiers, amongst others.

3.4. POLARIN Data Quality policy

POLARIN is based on collaboration between European institutions representing research excellence. These are composed of highly trained and qualified researchers who are familiar with the principles of the quality standard, that encourage all researchers to maintain, and to respond adequately to possible threats or violations of research integrity. In addition, WP7 will provide materials for the training of

staff in the correct handling of data, in compliance with the regulations in force. With these, researchers must ensure that at each stage of the project it respects all the following principles.

The methodology used must apply suitable methods, documented protocols and standard procedures to ensure the best possible quality of work, dissemination and replication.

Researchers must acknowledge all those who contribute, such as colleagues, collaborators and others. The basic principle of scientific activity is that researchers should be honest and respect their own actions and those of other scientists.

Researchers must respect accountability to the general public and must take all reasonable steps to ensure that their research complies with any agreements, related policies and guidelines of professional bodies, and allows for proper governance and transparency.

Finally, researchers should avoid any unreasonable risk or harm to research subjects and researchers themselves.

3.5. POLARIN Data Management and TA projects

Central to POLARIN is the provision of transnational access to polar research infrastructures (RIs). POLARIN builds on the successful and well-established transnational access (TA) to Arctic RIs provided by INTERACT and ARICE, and, for the first time, offers access to Antarctic RIs.

Access to Antarctica is, on one hand, heavily regulated by the Antarctic Treaty to minimize the impact of human activities on the continent's pristine environment. On the other hand, it is challenging and expensive due to difficult logistics and time-consuming travel. POLARIN offers access to such a resource for the first time. Moreover, POLARIN RIs are operated not only by national polar programs but also by private entities such as the Foundation Tara, which provides access to a sailing boat drifting through the Arctic Ocean, and the French cruise ship company Ponant, which contributes 1,400 units of access to the modern and state-of-the-art icebreaking cruise ship, Le Commandant Charcot.

POLARIN promotes excellence and interdisciplinarity in research, combining optimal resource use with multiplatform access (e.g., research stations, vessels, data infrastructures). POLARIN develops the scientific framework, defines the evaluation criteria for selecting applications, performs the scientific evaluation of submitted proposals, and assesses the contributions of POLARIN's funded projects towards addressing targeted challenges. In line with this goal, the TA project should also develop a Data Management Plan (DMP) that aligns with the POLARIN DMP, which includes a TA DMP template (see Annex).

The TA DMP template (see Annex 2) will be provided in the POLARIN Transnational Access Platform (POLARIN TAP) and a draft/preliminary version must be uploaded to POLARIN TAP by the applicants at the proposal submission, together with the project description and appendices and Curriculum Vitae of the user-group leader.

The TA DMP submitted will be assessed by WP4 members only if the proposal is selected for implementation. If needed, WP4 will support the TA applicant to complete the TA DMP with missing details (and consolidate the final version of the TA DMP).

4. Open Science and other Research outputs

4.1. Scientific communications

POLARIN partners are committed to ensure open access to peer-reviewed scientific publications relating to their results. In particular, they must ensure that an electronic copy of the publication is freely available from a trusted repository for scientific publications (open access journal), the POLARIN project is clearly acknowledged, and if any POLARIN related data is used, this is acknowledged and the dataset PID (e.g. DOI) is reported.

4.2. Other publishable results

POLARIN supports the open sharing of research outputs and facilitates open science by recommending its researchers to make publicly available a wide range of materials, such as presentations, newsletters, and more. While this material may not be classified as scientific papers, it remains valuable for POLARIN stakeholders.

A candidate technology for facilitating this outcome is the development of a “community” under the ZENODO system. ZENODO is an open-access repository developed and operated by CERN (the European Organization for Nuclear Research) and the EC through the OpenAIRE infrastructure. It is designed to support the open sharing of research outputs.

A “community” on ZENODO is essentially a thematic or topical group that helps organize and manage collections of research outputs and other materials related to a specific area of interest.

The ZENODO community can organize Open Access for presentations, posters, proceedings of meetings, images/diagrams, Promotional material of the project (factsheets...), public deliverables, after they are approved by the EC reviewers⁴, data (in case these data cannot be hosted by other POLARIN federated data nodes)⁵. ZENODO is not for peer-reviewed articles (full article should not be duplicated and the original should stay in the publisher trusted repository), confidential deliverables/milestones.

In ZENODO, each document is given a unique DOI, also deliverable reports and presentations are given a DOI which identifies the content and provides a persistent link to its location on the Internet (RETRIEVABILITY). ZENODO guarantees long-term full open access to the project documents (OPEN ACCESSIBILITY). ZENODO enables tagging a document with the grant agreement number (“Funding”), this ensures that each of the documents in Zenodo will be automatically harvested by the EC Participant Portal of the Commission and be listed there as “publications” linked to project. This useful feature ensures that each of the documents in Zenodo is automatically harvested by the EC Participant Portal of the Commission and is listed there as “publications” linked to project. This would allow POLARIN to automatize the process to show to the EC we are complying with the open access rules (LINK TO THE EC Portal).

⁴ These will be uploaded to Zenodo by ETT after the review

⁵ For uploading Data into Zenodo, the researcher has first to contact WP4 leaders (A. Novellino or V.Vitale)

4.3. Software tools and code

Open Science also deals with code sharing. Sharing code on open platforms like GitHub is crucial for advancing open science and research because it enhances transparency, allowing others to verify and replicate findings. It fosters collaboration by enabling researchers to contribute to and improve each other's work. Open code also accelerates innovation by providing a foundation for new developments and applications. Additionally, sharing code helps prevent duplication of efforts and promotes reproducibility in research. In line with this POLARIN DMP recommendation, a GitHub repository has been opened and is going to host the programming codes used for some of the data-processing-tools that POLARIN is developing. It is currently private but will be open in the future. Notably, Github is fully integrated with ZENODO, allowing publishing of final versions of software tools and code to the ZENODO community

5. Allocation of resources

5.1. Costs before making data FAIR and long-term preservation

POLARIN WP4 and WP5 are designed to manage the POLARIN (meta)data catalogue and (virtual) data access. They are developing a FAIR compliant central node of a federated data infrastructure (already designed to enable long-term data preservation).

Whenever needed, WP4 and WP5 will develop middleware tools (scripts, routines, methodologies) to facilitate the harmonization of the integration of these (meta)data federated POLARIN (meta)data sources into the common catalogue.

Additionally, all partners have allocated a dedicated budget for the promotion and dissemination of (meta)data and for supporting data preservation after the project's end.

Regarding data preservation, POLARIN is developing a system where metadata and data are hosted long-term in national (trusted) data repositories. Moreover, when relevant, a copy of the data will be shared or submitted to key European and international initiatives that address POLARIN's data and goals.

As an example, ocean data collected by the POLARSTERN German Research Vessel will be hosted by the PANGAEA German National Data Repository, which is federated with EMODnet Physics. EMODnet Physics implements interoperability with the Marine Copernicus Service and GOOS stakeholders. Thus, POLARIN data collected by the POLARSTERN are both long-term preserved (by PANGAEA) and made available to a broad community of stakeholders.

5.2. POLARIN Data Management officer

Data Management is included in a dedicated project work package (WP4). The POLARIN Data Management Officer responsible for the data management is Antonio Novellino (ETT), who has long term experience in ocean data management and sharing.

Antonio Novellino is member of the EuroGOOS DATAMEQ group and contributes to several EuroGOOS Task Teams for advising (<http://eurogoos.eu/>) on operational oceanography data management procedures and standards. He is co-chair of the SOOS DMSC (South Ocean Observing System Data Management Steering Committee), member of the DOOS DMTT (Deep Ocean Observing System Data Management Technical Team). He serves on the EMODnet Steering Committee, the EMODnet

Technical Working Group. He serves the Expert Team on WIS Centres (ET-WISC) and Task Team on Data Centres (TT-DC), and an EOSC-FAIR champion.

6. Data Security

The POLARIN central services (GeoNetwork, ERDDAP, mapviewer) are going to be hosted on ETT cloud that is on the ARUBA.it infrastructure. Since 2015, Aruba is running a dedicated service to private business clients and it provides the client with top level services such as Data Centre (Virtual Servers, Real Servers, hosting infrastructures), Back up and Disaster Recovery etc. ARUBA also provides us with most recent services for data security cryptography (AES), security protocols (AES, SSL) and bandwidth balance. The main characteristics of the service are SLA: 99,80%; security: crypted transmission channel; bandwidth unlimited (upload/download); Certifications: ISO 9001:2015, ISO 27001:2013; service desk: 24h; Cloud security certification: ISO/IEC 27017:2015; data privacy: ISO/IEC 2018:2014; security incidents: ISO/IEC 27035:2016.

7. Ethics and GDPR

7.1. Ethics and Code of Conduct

Partners must carry out their actions in accordance with ethical principles that include the highest standards for research integrity and the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.

Partners must pay particular attention to the principle of proportionality, the right to privacy, the right to the protection of personal data, the right to the physical and mental integrity of persons, the right to non-discrimination, the need to ensure protection of the environment and high levels of human health protection.

In addition, the partners must respect the fundamental principle of research integrity as set out in the European Code of Conduct for Research Integrity.

This means following the principles of the reliability of quality of research in design, methodology, analysis and use of resources. As well as honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent way. Furthermore, the respect for colleagues, society, ecosystem, cultural heritage⁶ and the accountability for the research from idea to publications, management, organisation, training, supervision, mentoring and for impacts.

This also means that the partners must ensure that persons, carrying out research tasks, follow the good research practices including ensuring, where possible, openness, reproducibility and traceability and refrain from the research integrity violations described in the Code of Conduct.

7.2. Personal Data Regulation

POLARIN is an international network of polar research infrastructures and services designed to address the scientific challenges of the polar regions. POLARIN research data encompasses measurements and analyses conducted using various research infrastructures, including Arctic and Antarctic research

⁶ Partners may also refer to CARE principles (<https://www.gida-global.org/care>)

stations, research vessels, icebreakers operating at both poles, observatories, data infrastructures, and repositories for ice and sediment cores.

This data is not related to personal information and, in general, does not fall under the GDPR.

However, according to the POLARIN metadata model, outcomes should also include information about the principal investigators (e.g., Name, Surname, ORCID). While this constitutes personal data, it does not fall under the GDPR, as it is in the interest of the researcher to be recognized for their contributions to science and to progress in their career.

Notably, some POLARIN activities, such as workshops, meetings, the TA program, training program, etc. may collect personal data that falls under GDPR. In these cases, POLARIN develops a document specifying the reasons for personal data collection and its intended use. Personal data will not be used for any other purposes or shared outside the project and the specific task that collected the data. This document also informs participants of their GDPR rights, including the rights to access, rectify, erase, restrict processing, data portability, and object to data processing. Derived anonymized/grouped figures (e.g., the number of attendees) may be used for POLARIN reporting purposes.

In these cases, Article 37 of the GDPR, “Designation of the Data Protection Officer,” requires the appointment of a Data Protection Officer (DPO). Therefore, the POLARIN DMP mandates that each organization managing lists of people (e.g., for attending a meeting or entering an infrastructure) appoint a DPO. The contact details of the DPOs should be included as an annex to the POLARIN DMP. DPOs are responsible for implementing technical and organizational measures that safeguard the rights of data subjects/research participants, following their organization's best practices and GDPR and security directives. DPOs may be required at any time to provide documents (or links to them) describing the methods applied.

Generally, research institutes manage DPOs at the institute level. In such cases, the DPO should add the POLARIN project to their list of managed projects.

DPOs are also responsible for responding promptly to participants whenever they choose to exercise their rights (e.g., removal from the list).

Moreover, POLARIN adopts the following approach to GDPR:

- Perform a periodic Data Protection Impact Assessment, to verify that POLARIN partner are applying the recommendation on data protection
- Informed Consent, to ensure that participants provide explicit, informed consent for the collection and use of their personal data, and maintain records as evidence that proper consent was obtained.
- Collect Only Necessary Data, to ensure that the personal data collected is adequate, relevant, and limited to what is necessary for the purposes of the research
- Keep the documentation (DMP, consent forms, data protection policy, etc) updated

8. References

- Quimbert, E., Fichaut, M., Maudire, G. (2022). Guide principes FAIR. Principes FAIR dans le contexte du pôle ODATIS <https://doi.org/10.13155/87107>

Acknowledgements

POLARIN has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101130949. Please visit www.eu-polarin.eu for more information.

Annex 1 – POLARIN Metadata Model

Variable name	Attribute name	Data type	Description	Mandatory	Vocabulary	Notes
NC_GLOBAL	cdm_data_type	string	Data type according to Common Data Model (e.g., TrajectoryProfile, TimeSeries, Grid, Point, Other)	yes		
NC_GLOBAL	Conventions	string	Conventions	yes		
NC_GLOBAL	contributors_email	string	Email of any individuals or institutions that contributed to the collection of this data (separated by comma)			It should include the principal investigator (coordinator of the data collection). The PI guarantees that the list is complete and that other authors have agreed to publish their email
NC_GLOBAL	contributors_name	string	Name of any individuals or institutions that contributed to the collection of this data (separated by comma)			It should include the principal investigator (coordinator of the data collection). The PI guarantees that the list is complete and that other authors have agreed to publish their names
NC_GLOBAL	contributors_orcid	string	ORCID of any individuals or institutions that contributed to the collection of this data (separated by comma)		ORCID	It should include the principal investigator (coordinator of the data collection). The PI guarantees that the list is complete and that other authors have agreed to publish their orcid
NC_GLOBAL	contributors_role	string	Role of any individuals or institutions that contributed to the collection of this data (separated by comma)			It should include the principal investigator (coordinator of the data collection). The PI guarantees that the list is complete and that other authors have agreed to publish their role
NC_GLOBAL	creator_email	string	Email of data creator	yes		It is a generic (not personal) email to contact the institute/project/...
NC_GLOBAL	creator_name	string	Name of data creator	yes		
NC_GLOBAL	creator_type	string	Type of data creator (e.g., person, institution)	yes		
NC_GLOBAL	creator_url	string	URL of data creator	yes		

Deliverable D1.4

NC_GLOBAL	geospatial_lat_max	double	Max latitude expressed in WGS84 (the highest possible precision)	yes, if spatial data	WGS84	
NC_GLOBAL	geospatial_lat_min	double	Min latitude expressed in WGS84 (the highest possible precision)	yes, if spatial data	WGS84	
NC_GLOBAL	geospatial_lat_resolution	double	Latitude resolution expressed in WGS84 (for grid data)	yes, if spatial grid data		
NC_GLOBAL	geospatial_lat_units	string	Degrees north	yes, if spatial data		
NC_GLOBAL	geospatial_lon_max	double	Max longitude expressed in WGS84 (the highest possible precision)	yes, if spatial data	WGS84	
NC_GLOBAL	geospatial_lon_min	double	Min longitude expressed in WGS84 (the highest possible precision)	yes, if spatial data	WGS84	
NC_GLOBAL	geospatial_lon_resolution	double	Longitude resolution expressed in WGS84 (for grid data)	yes, if spatial grid data		
NC_GLOBAL	geospatial_lon_units	string	Degrees east	yes, if spatial data		
NC_GLOBAL	geospatial_vertical_max	double	Max vertical extension (in case of vertical profile) (the highest possible precision)	yes, if spatial data		
NC_GLOBAL	geospatial_vertical_min	double	Min vertical extension (in case of vertical profile) (the highest possible precision)	yes, if spatial data		
NC_GLOBAL	geospatial_vertical_positive	string	Positive direction of vertical extension ("up" means that z increases up - height, "down" means that z increases downward - pressure or depth)	yes, if spatial data		
NC_GLOBAL	geospatial_vertical_units	string	Units used for the vertical extension	yes, if spatial data		
NC_GLOBAL	infoUrl	string	URL of data information background (es. project web page, dataset page, ...)	yes		
NC_GLOBAL	inspire	string	INSPIRE spatial data or SeaDataNet vocabulary	yes, if spatial data	INSPIRE Spatial Data Themes (GEMET)	
NC_GLOBAL	institution	string	Institution principally responsible for this data (owner or provider)	yes		

Deliverable D1.4

NC_GLOBAL	institution_edmo_code	string	EDMO code of the institution principally responsible for this data (owner or provider)	yes	EDMO SeaDataNet	See specific project list on EDMERP SeaDataNet
NC_GLOBAL	institution_country	string	Country of the institution principally responsible for this data (owner or provider) expressed according to ISO 3166	yes	ISO 3166	
NC_GLOBAL	keywords	string	List of keywords and phrases (separated by comma - in case of multiple vocabularies, insert keywords following the same order listed in "keywords_vocabulary")	yes		
NC_GLOBAL	keywords_vocabulary	string	Identifies the controlled list of keywords from which the values in the "keywords" attribute are taken (in case of multiple vocabularies, separated by comma)	yes	e.g.: Global Change Master Directory (GCMD)	
NC_GLOBAL	licence	string	Licence that describes the restrictions to data access and distribution	yes	Creative commons	
NC_GLOBAL	naming_authority	string	Name of who defines the data set and the standards to be applied	yes		Fixed value
NC_GLOBAL	references	string	Description of how the dataset was created: published or web-based references that describe the data or methods used to produce it. Recommend URIs (such as a URL or DOI) for papers or other references. This attribute is defined in the CF conventions	yes		
NC_GLOBAL	source	string	The method of collection and production of the dataset (e.g., types of instrument, model, collection). If it was model-generated, source should name the model and its version. If it is observational, source should characterize it. This attribute is defined in the CF Conventions	yes		
NC_GLOBAL	standard_name_vocabulary	string	Name and version of standard vocabulary (e.g., CF Standard Name Table v70)	yes		
NC_GLOBAL	summary	string	A paragraph describing the dataset, analogous to an abstract for a paper	yes		
NC_GLOBAL	time_coverage_duration	string	Time coverage duration using ISO 8601 (in alternative to time_coverage_start/time_coverage_end)	yes	ISO 8601	
NC_GLOBAL	time_coverage_end	string	Time coverage end using ISO 8601	yes	ISO 8601	

Deliverable D1.4

NC_GLOBAL	time_coverage_resolution	string	Time coverage resolution using ISO 8601 (if applicable)	yes, if applicable	ISO 8601	
NC_GLOBAL	time_coverage_start	string	Time coverage start using ISO 8601	yes	ISO 8601	
NC_GLOBAL	title	string	Dataset title (a short phrase or sentence describing the dataset)	yes		
NC_GLOBAL	variables	string	List of variables id (separated by comma)	yes		
data	data_creation	string	Date of dataset creation	yes	ISO 8601	
data	data_update	string	Date of dataset update	yes	ISO 8601	
data	data_version	string	Dataset update version	yes		
data	data_doi	string	Data DOI/PID	yes, if existing		
project	project_name	string	Project name	yes	CORDIS for European Project	Fixed value
project	project_code	string	Project code/acronym	yes	CORDIS for European Project	Fixed value
project	project_id	double	Project Grant agreement number	yes	CORDIS for European Project	Fixed value
project	project_statement	string	Project Grant agreement statement	yes	CORDIS for European Project	Fixed value
project	Project_DOI	string	Project DOI	yes	CORDIS for European Project	Fixed value
project	project_edmerp	string	Project EDMERP code	yes	EDMERP SeaDataNet	Fixed value

Deliverable D1.4

project	oceanographic_campaign	string	Name of the oceanographic campaign during which the data of the specific project were collected			
ship	ship_name	string	Name of the ship (e.g., "Belgica")			
ship	ship_imo	double	IMO ship identification number (unique ship identifier) - report only the seven-digit number (e.g. "9871294")		Marine Traffic Research	
ship	ship_call_sign	string	Maritime call sign assigned as unique alphanumeric identifier to the ship (e.g., "ORCO")			
platform	platform_type	string	Type of platform		NERC Vocabulary L06	
platform	platform_id_orig	string	Pre-existing platform id, if applicable			
sensor	sensor_model	string	Sensor URL		NERC Vocabulary L22	

Annex 2 - DMP template for TA

Scope of the Transnational Access Data Management Plan (TA DMP)

The user-group leader should develop a single Data Management Plan (DMP) to cover the project's overall approach to managing data. For specific datasets that require unique considerations (e.g., regarding openness or embargo periods), these should be clearly outlined.

This document provides the required sections and fields to ensure that POLARIN can effectively access and utilize the incoming data. The DMP is a living document, updated as the project progresses or when significant changes occur. Updated versions must be submitted whenever major modifications are made, and a final version should clearly state:

- What data has been produced,
- What data has been made available to POLARIN,
- Which data have been integrated into the POLARIN nodes.

The template contains the following sections:

- **Executive Summary:** Outlines the project scope, technologies, and goals.
- **Data Summary:**
 - A description of the data produced, including licensing, availability, and access methods.
 - Details of any external data consumed, with clear provenance and usage notes.
 - Methods and technologies adopted to ensure that the data is FAIR (Findable, Accessible, Interoperable, and Reusable).
- **Other Considerations:** Including data security, privacy, and ethics concerns, if applicable.
- **Embargo Period:** If an embargo period is requested (between 6-24 months), a justification should be provided.

The TA DMP template will be accessible via the POLARIN Transnational Access Platform (POLARIN TAP). Applicants should upload a draft or preliminary version, alongside other required documents (e.g., research plan, CV).

The TA DMP will not be reviewed until the proposal has been selected for implementation. At this time, POLARIN WP4 will check the completeness and compliance with the POLARIN DMP and if needed will support the user-group leader in completing the TA DMP with missing details (and consolidate the final version of the TA DMP).

PRELIMINARY TA DMP Template

Project Information

- **Project Name:**
 - **Project Acronym:**
 - **User-group leader:**
 - **Hosting research infrastructure:**
 - **Hosting research infrastructure contact:**
-

Executive Summary

- Provide a brief project summary (copy from the TA proposal form).
-

Data Summary

- What data will the project collect or generate?
 - Is this new data or re-analysis of existing data?
 - Will other external data sources be reused? (Provide references if applicable).
-

Data Type & Acquisition

- What kind of data is being collected (e.g., in-situ, remote sensing, model output)?
 - Is data collection continuous or campaign-based? Include start and end dates.
 - Expected data size.
 - Will real-time data be provided?
-

Data Storage & Format

- What format will the data be in (e.g., NetCDF, CSV)?
 - Where will the data be stored (e.g., public or private repository)? Include a link if possible.
-

Metadata

Specify the metadata format used (e.g., XML, JSON, CSV).

List the standards and conventions followed for attributes (e.g., time, feature description, parameters).

FAIR Data

- How will the data be made accessible (e.g., software tools required)?
 - Will you apply a data license (e.g., Creative Commons)?
 - Will the data have a DOI?
 - Is an embargo period required? If yes, for how long and why (provide a justification)?
-

Data Security & Ethics

- How will the data be secured?
 - Is personal or sensitive data being collected? If yes, how will privacy be protected?
 - Are there any ethical concerns related to the project (e.g., human participants, sensitive ecosystems)?
-

Additional Notes

- This DMP is preliminary and will not be evaluated until the project is accepted for implementation. Then, POLARIN WP4 will support the TA applicant to complete the TA DMP with missing details (and consolidate the final version of the TA DMP).
- The maximum embargo period is 24 months after project completion. [If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible. If data is declared to be CC, embargo cannot exceed the project duration]
- **Metadata and metadata conventions:**
 - Metadata must be provided to describe the data (who, what, where, when) and should be made available as soon as possible under a **CC0 license**.
 - Follow metadata standards like **ISO19115/19115-2** and use international conventions (e.g., **ISO8061 for time**, **WGS84 for coordinates**).
 - Metadata should include the citation and acknowledgement to POLARIN.
- Once the data is published in a trusted repository, applicants must submit the DOI reference, along with any publication records, to the POLARIN Transnational Access Platform (TAP).

Annex 3. Further recommendations/measures on long term preservation of informed consent

Whenever POLARIN will implement surveys, questionnaires, or collect personal data for any reason (e.g., attendance to events), European GDPR regulation will be used as reference and the user will be informed about the use of personal data. In general, POLARIN will not transfer personal data (e.g., email addresses) to other entities and the only use will be setting up a distribution list to inform users about project progress. User will be always able to change their consensus and ask for being removed from the distribution channel.

For stakeholder engagement activities, coordinators will collect only the necessary information, including name and surname, affiliation, and activity sector/field of expertise.

Information will be collected by using information sheets and consent forms. Any further contact by the coordinator the activity should be explicitly authorised by the data subject, and it will be strictly limited to the purposes stated in the consent form.

Contact information will be used to involve stakeholders who gave their consent in follow up communications and efforts, in order to guarantee consistency with POLARIN multi-stakeholder approach.

Workshops and dissemination events will be conducted according to two different approaches throughout the project timeframe, based on the purpose and scope of the initiative:

- i. Events without registration: for these events, coordinators will collect no data about participants, in line with the principle of data minimisation. This approach will be followed whenever it does not impair the fulfilment of POLARIN objectives.
- ii. Events with registration: in order to take part in these events (e.g., workshops), the participants will need to submit a registration form, with a limited number of personal data fields (up to 7). In this case, basic personal information and contact data will be used to facilitate the management of the meeting, give participants access to preparatory materials. Participants will have the possibility to give their consent to receive information on the organisation of further initiatives and related activities.

In accordance with the EU Charter of Fundamental Rights and with the GDPR, POLARIN will safeguard the following rights throughout the project duration: (i) right to be informed about the personal information shared, (ii) right of access to such information, (iii) right to rectification (in case the information is incorrect or incomplete), (iv) right to erasure of the data, (v) right to restrict processing of the data, (vi) right to data portability from a service to the other, (vii) right to object.

Concerning the right to be informed (art. 13 and art. 14 GDPR) about the personal data shared with the Consortium (or with one of the partners), when giving the consent to data processing, data subjects will be adequately informed and will be able to access information on data sharing on the portal of the activity for which data are collected.

Concerning the right to access information shared (art. 15 GDPR), data subjects will have the possibility to request a copy of such data to the relevant DPO. Based on the internal procedures of the DPO, the data stored by the organisation will be collected within 30 days. Any potential delay or obstacle (due to the anonymisation/pseudonymisation procedure or to the data storage techniques) will be adequately notified and justified to the data subject.

Similarly, concerning the right to rectify incorrect or incomplete information (art.16 GDPR), POLARIN will ensure that data subject have the possibility to submit a request of rectification to the relevant DPO, that will carry out the necessary procedures within 30 days. Any potential delay will be adequately notified and justified to the data subject.

Concerning the right to be forgotten (art. 17 GDPR), POLARIN will ensure the erasure of the data collected and of the files stored in local archives within 30 days from the official request. When sharing personal data, the subject will be informed of the procedures to obtain the erasure of such data, according to the internal policies and data management infrastructures of the partner leading a certain activity.

Concerning the right to portability of data and their limitation (art. 18 and art 20 of GDPR), the subject will be informed on the procedures to obtain the transmission of data to the physical or legal person designated, based on the internal policies of the responsible partner. The partner will be responsible for the assessment and management of the request and for the aggregation of data in an intelligible format in due time.