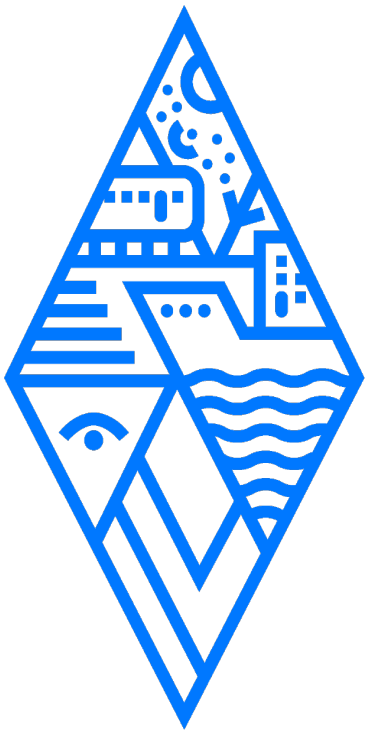

POLARIN
POLAR
RESEARCH
INFRASTRUCTURE
NETWORK

Deliverable 4.3. [POLARIN Web Portal] - POLARIN data portal, preliminary design of the POLARIN Data hub and its features

V1.1, March 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	D4.1
Deliverable title	[POLARIN Web Portal] POLARIN data portal, preliminary design of the POLARIN Data hub and its features
Version	V1.1
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	A Novellino, V Vitale
Contributors	ETT, CNR, SIOS, NILU, AU
Contributing authors	MC Paolini
Due date	28/02/2025
Delivery date	24.03.2025

Document history	
Creation Date	24/02/2025
Version	V1.1
Version Date	22/03/2025
Status	☐ Draft ☒ WP lead approved ☒ Executive Board approved ☒ Coordinator approved
Status date	24.03.2025

TABLE OF CONTENTS

1. SUMMARY 6
2. Introduction 7
3. POLARIN Data Hub (DH) 8
4. References 13

Acronyms

RI: Research Infrastructure; 4
WP4: Work Package 4; 4
PDF: Polar Data Forum; 5
SOOS: Southern Ocean Observing System; 5
POAwg: Polar Observing Assets Working Group; 5
ADC: Arctic Data Committee; 5
SAON: Sustaining Arctic Observing Networks; 5
RoPON: Registry of Polar Observing Networks; 5
DH: Data Hub; 6
FAIR: Findable, Accessible, Interoperable, and Reusable; 6
DMP: Data Management Plan; 6
TA: Transnational Access; 6
VA: Virtual Access; 6
AI: Artificial Intelligence; 7
M2M: Machine-to-Machine; 7
CAFF: Conservation of Arctic Flora and Fauna; 7
ABDS: Arctic Biodiversity Data Service; 7
ARICE: Arctic Research Icebreaker Consortium; 7
CNR: National Research Council (Italy); 7
IADC: Italian Arctic Data Center; 7
INPA: Instituto Nacional de Pesquisas da Amazônia; 7
INKODE: Interact Knowledge and Data Exchange; 7
IDP: INTERACT Data Portal; 7
NADC: National Antarctic Data Centre; 7
GFZ: German Research Centre for Geosciences; 7
POSEDA: Potsdam Seismic Data; 7
SIOS: Svalbard Integrated Arctic Earth Observing System; 7
SDMS: Svalbard Data Management System; 7
NILU: Norwegian Institute for Air Research; 7
ACTRIS: Aerosol, Clouds, and Trace Gases Research Infrastructure; 7
AU: Aarhus University; 7
ARC-MO: Arctic Monitoring and Observing; 7
GEM: Greenland Ecosystem Monitoring; 7
SPRS: Swedish Polar Research Secretariat; 8
ABISKO: Abisko Scientific Research Station; 8
ULAVAL: Université Laval; 8
CEN WK: Centre d'Études Nordiques, Whapmagoostui-Kuujjuarapik; 8
DIR-ITA: Dirigibile Italia; 8
UTU: University of Turku; 8
KEVO: Kevo Subarctic Research Station; 8
UH: University of Helsinki; 8
KILPIS: Kilpisjärvi Biological Station; 8
SAVN: The Natural History Museum of the Faroe Islands; 8
KOLTUR: Koltur Research Station; 8
UOULO: University of Oulu; 8

OULANKA: Oulanka Research Station; 8
FMI: Finnish Meteorological Institute; 8
PAL-SOD: Pallas-Sodankylä Research Station; 8
NPI: Norwegian Polar Institute; 8
SVERDRUP: Sverdrup Research Station; 8
UAF: University of Alaska Fairbanks; 8
TOOLIK: Toolik Field Station; 8
TRS: Tarfala Research Station; 8
SU: Stockholm University; 8
ZAC: Zackenberg Research Station; 8
CSIC: Spanish National Research Council; 8
JCI: Juan Carlos I Antarctic Research Station; 8
TROLL: Troll Research Station; 8
ARI: Aurora Research Institute; 8
WARC: Western Arctic Research Centre; 8
IPEV: French Polar Institute Paul-Émile Victor; 8
CONCORDIA: Concordia Station; 8
MSZ: Mario Zucchelli Station; 8
AMUNDSEN: CCGS Amundsen Icebreaker; 9
AWI: Alfred Wegener Institute; 9
POLARTSERN: Polarstern Icebreaker; 9
MI: Marine Institute; 9
CELTIC: Celtic Explorer Research Vessel; 9
HAAKON: Kronprins Haakon Research Icebreaker; 9
UIT: UiT The Arctic University of Norway; 9
APECS: Association of Polar Early Career Scientists; 9
ZC: Zenodo Community; 9
DV: Data Viewer; 9

1. SUMMARY

POLARIN's overall aim is to provide efficient and customised research infrastructure (RI) services to address the scientific challenges of the polar regions. This includes offering access to a wide portfolio of complementary and interdisciplinary top level RIs. By integrating polar RI capacities, POLARIN facilitates scientific research to understand and predicting key processes in polar regions in the context of climate changes. This effort enhances society's problem-solving capacity and support evidence-based policy making.

One of POLARIN's four core services is improving access to data and developing online services and data products (WP4). This deliverable focuses on implementing this service through a front-end data portal designed to help users find polar data tailored to their needs.

This document describes the preliminary design of the POLARIN Data Hub and its features.

2. Introduction

Polar regions are experiencing rapid ice loss and significant transformations in their oceans and land, with global repercussions that impact people in diverse ways. The remoteness and inaccessibility of the polar regions, combined with limited research infrastructure, present significant challenges. Many initiatives (e.g. the Polar Data Forum – PDF, the Southern Ocean Observing System – SOOS, the Polar Observing Assets Working Group – POAwg, the Arctic Data Committee – ADC, the Sustaining Arctic Observing Networks – SAON, etc.) are working synergically to support and facilitate research.

Emerging efforts like the Registry of Polar Observing Networks (RoPON) highlight the ongoing effort for improving semantic interoperability, metadata standards, and community-driven data solutions.

In this framework POLARIN organizes and coordinates an international network of of top-tier polar research infrastructures and services, facilitating interdisciplinary research on complex polar processes. POLARIN has brought together a unique collection of 64 polar research infrastructures (RIs), ranging from small research stations in the Arctic and Antarctic to large icebreakers operating at both poles.

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.

POLARIN represents a joint international cooperation network which ensures excellent research contributing to identified research challenges in polar regions and is targeting four key services to users:

1) Challenge driven TA to research infrastructures
2) Improvement of data services and data products
3) VA to research infrastructures
4) Training for research infrastructure users

To this end, the project, and in particular WP4, advocates for an open and FAIR data approach (see Deliverable D4.1 - POLARIN DMP for more details) to reduce data fragmentation, to develop further data tools form previously EU founded projects, to improve data accessibility and usability. Central to this process is the design and development of the POLARIN data portal (POLARIN Data Hub – DH) that is going to provide stakeholders with data catalogues, interfaces, visualization tools, data search support tools (also empowered by AI), etc.

3. POLARIN Data Hub (DH)

The main goal of the POLARIN DH is to offer a single easy access point to discover and access the ever-growing heterogeneous collections of (polar) data and products (including observations from satellites, airplanes, drones and in-situ sensors etc) for both the Arctic and Antarctic Poles.

The POLARIN DH is central to the overall POLARIN data objectives (Figure 1) and represents a key supporting tool for many of the POLARIN activities.

O1: Enabling science for understanding and predicting key processes in polar regions [...] • *Identification of current knowledge gaps* [..]	***O2: Provide efficient challenge-driven transnational access (TA) to top level research infrastructures in the polar regions*** *[...]* • *Integrated proposal management platform for POLARIN RIs.* • *Amount of TA to polar RIs*	***O3: Improvement of data services and data products*** [...] • *Inventory of databases from polar regions.* • *Identification of workflows to improve interoperability among databases.* • *Creation of tools to facilitate the consumption of data.* • *Customised data products (e.g., atmospheric distribution of black carbon).* • *Customised data services (e.g., calibration/validation service for earth observation data).*	***O4: Provide virtual access (VA) to research infrastructures*** [...] • *A unified semantically consistent virtual data catalogue with machine interfaces.* • *A web portal providing guidance documentation and a graphical user interface to the virtual data catalogue.* • *Amount of VA to polar RIs, data infrastructures and data services.* • *Amount of data downloaded from the POLARIN data hub.*	***O5: Training for infrastructure users*** [...] *Training resources, recordings of online seminars and other training materials.* [...] ***O6: Advertise RI services and engage RI users***

Figure 1. POLARIN Data Objectives across the POLARIN WPs

An internal survey conducted during a WP4 extended meeting indicated that when looking for data (expert) users are interested in data and variables but at the same time they are trying to access these by different approaches, such as searching for platforms, searching for parameters, looking at a map, searching for a specific data format etc. (Figure 2)

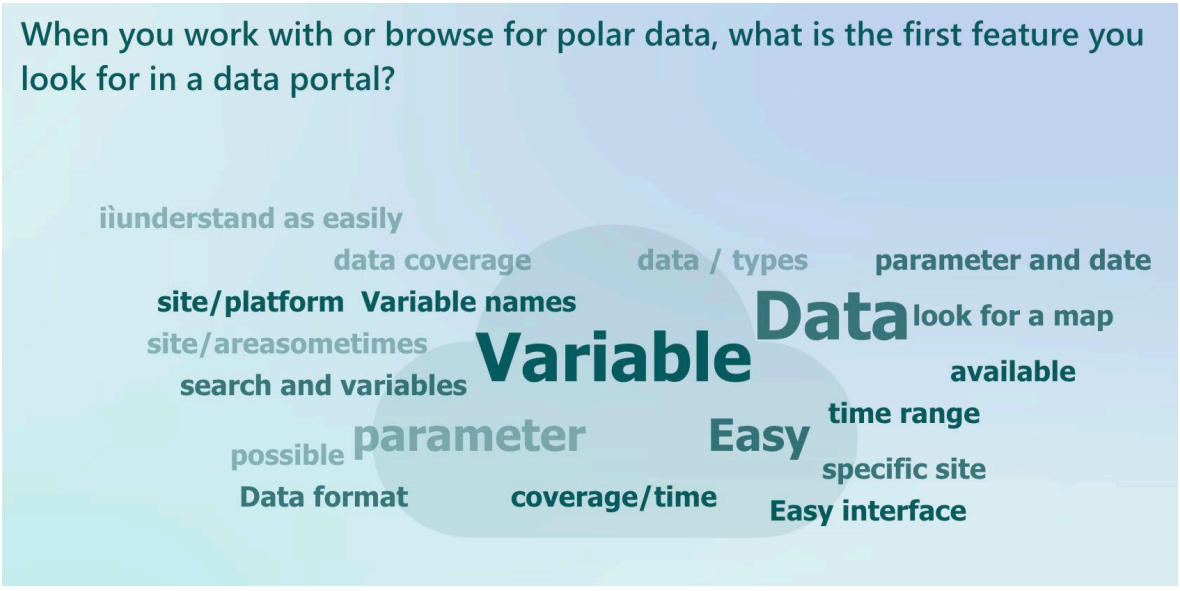

Figure 2. Word cloud outcome from the internal survey (WP4 meeting, 24/02/2025). Feature

The same survey highlights that experts do not have a single starting point for their research data search (Figure 3). This means that different experts may consider different subsets of data when

answering the same scientific question, emphasizing the importance of developing overarching tools to harmonize and facilitate the discovery of the widest range of resources from a single platform.

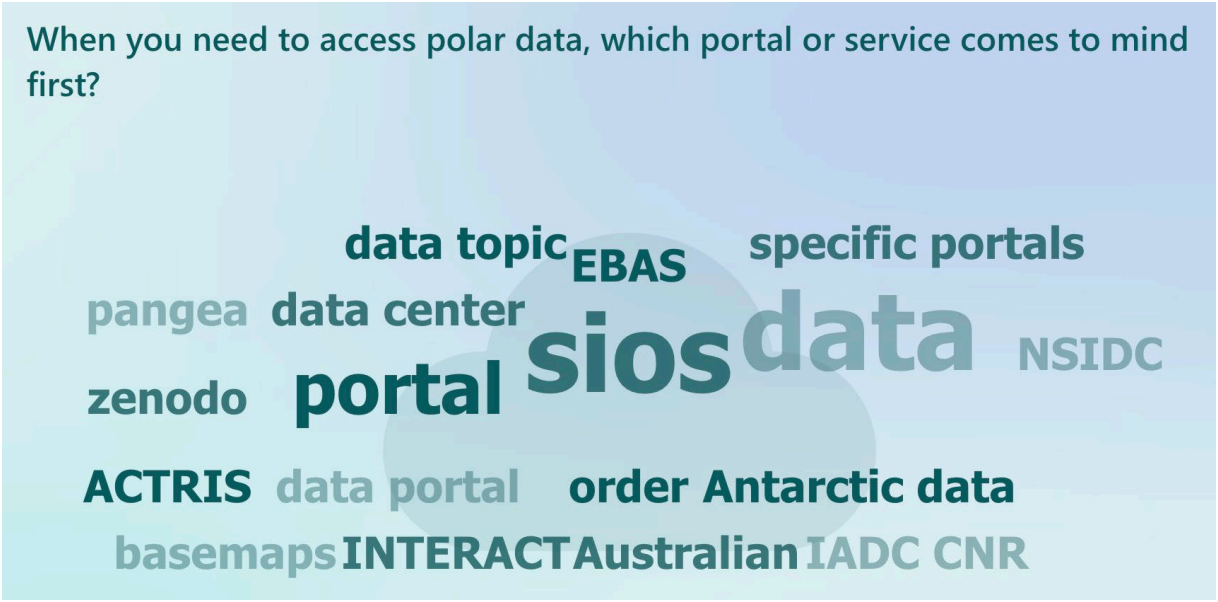

Figure 3. Word cloud outcome from the internal survey (WP4 meeting, 24/02/2025). Primary source

In line with these needs, the POLARIN DH "establishes a hub where experts as well as non-experts/intermediaries are guided ([..]) to find data and information necessary to match their specific needs" and "develop tools and services for intermediaries and end-users (both experts and non-experts)".

Figure 4 shows the POLARIN DH architecture and related backend services and tools, which range from the data catalogues, to tools to load and view data in time and space, to tools to process and compare results, to new AI based features to support data navigation and processing.

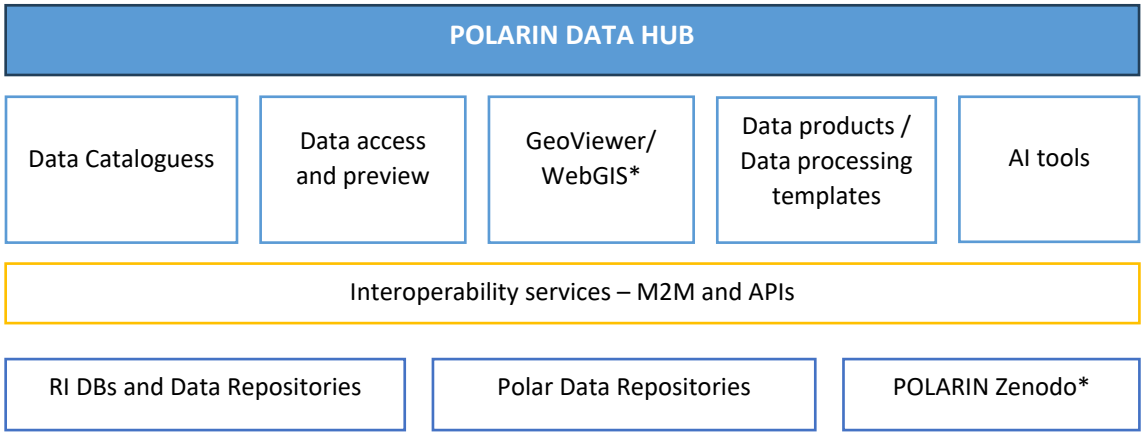

Figure 4. POLARIN Data Objectives across the POLARIN WPs

These tools interact and interoperate with the backend and M2M services from the POLARIN data nodes (e.g. the research infrastructures and databases) and open repositories that hosts polar data.

Table 1. POLARIN data nodes (with the latest additions highlighted in yellow)

	Host/provider	TA or/& VA	Project name	Web site
Data infrastructure				
1	CAFF	VA	ABDS	https://abds.is/
2	ARICE	VA	ARICE	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
8	NILU	VA	ACTRIS	https://www.actris.eu/topical-centre/data-centre
Greenland Network Database				
1	AU	VA	ARC-MO	https://gios.org
2	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.fo/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	NPI	TA/VA	SVERDRUP	https://data.npolar.no
10	UAF	TA/VA	TOOLIK	https://www.uaf.edu/toolik/handbook/index.php https://arcticdata.io/catalog/data/query=toolik
11	SU	TA/VA	TRS	www.su.se/tarfala-forskningsstation/
12	AU	TA/VA	ZAC	https://data.g-e-m.dk
13	CSIC	TA/VA	JCI	http://cndp.utm.csic.es/geonetwork/srv/eng/catalog.search#/home
14	NPI	TA/VA	TROLL	https://data.npolar.no/
15	ARI	TA/VA	WARC	www.nwtresearch.com
16	CNR/IPEV	TA/VA	CONCORDIA	www.concordiastation.aq/home-1/
17	CNR	TA/VA	MSZ	www.pnra.aq/stazione-mario-zucchelli
Research vessels				
1	ULAVAL	TA/VA	AMUNDSEN	https://www.polardata.ca/
2	AWI	TA/VA	POLARTSERN	https://www.pangaea.de/
3	MI	TA/VA	CELTIC	https://erddap.marine.ie/erddap/index.html
4	NPI	TA/VA	HAAKON	https://data.npolar.no/ ; https://nmdc.no/
Core repositories				
1			UIT/APECS-UIT CORES	https://geodata.uit.no/core_repository

Importantly, the POLARIN Zenodo community (POLARIN ZC)[1] has been awarded as an Open Pilot Research Project Community meaning that each of the uploaded digital objects (e.g. data products) can be up to 200GB, thus, besides offering easy access to POLARIN public documents, it can represent a further potential data node for the system.

Figure 5 shows the current design of the POLARIN DH landing page. The tools organized in areas, each providing the user with preliminary information or links to the specific service. A key goal is to inform the user about the amount of available and integrated resources, their temporal and geographical distribution, the access policy. Some pre-defined filters (right hand), i.e. POLARIN scientific challenges, will guide the user towards a selection of data and products. The navigation bar (top) provides the user with link to the viewer, catalogue, and cookbook repositories[2]. The catalogue (Figure 6) is designed as a shopping basket where the user finds basic information (metadata) about the data and products and can add this to the shopping list for successive data pre-view or data request/download. Notably the catalogue[3] is dynamic and based on the metadata offered by the linked backend POLARIN nodes (Figure 7).

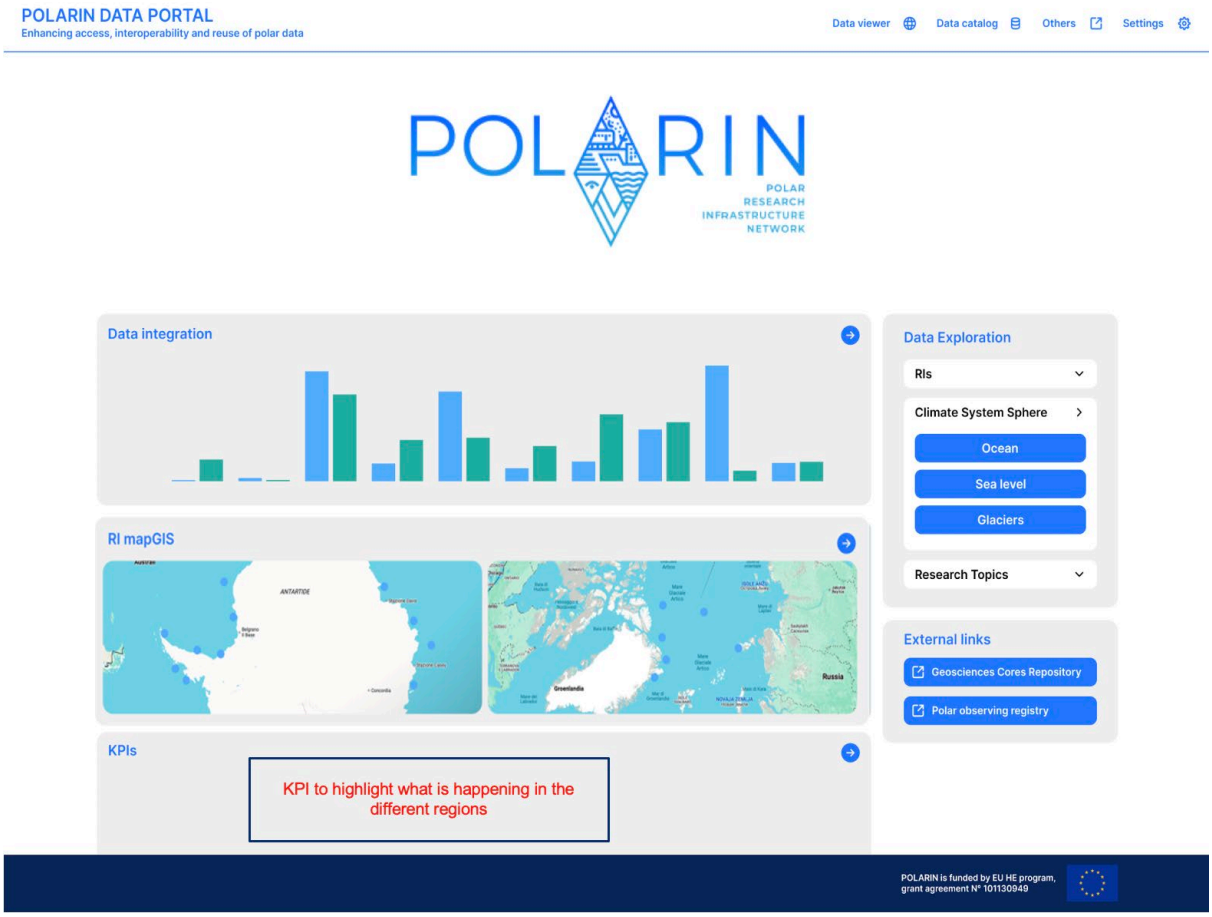

Figure 5. POLARIN DH Landing Page

[1] https://zenodo.org/communities/eu-polarin/records?q=&l=list&p=1&s=10&sort=newest
[2] https://github.com/he-polarin
[3] The catalogue middle layer is based on GeoNetwork - https://geonetwork.s4polarin.eu/geonetwork/

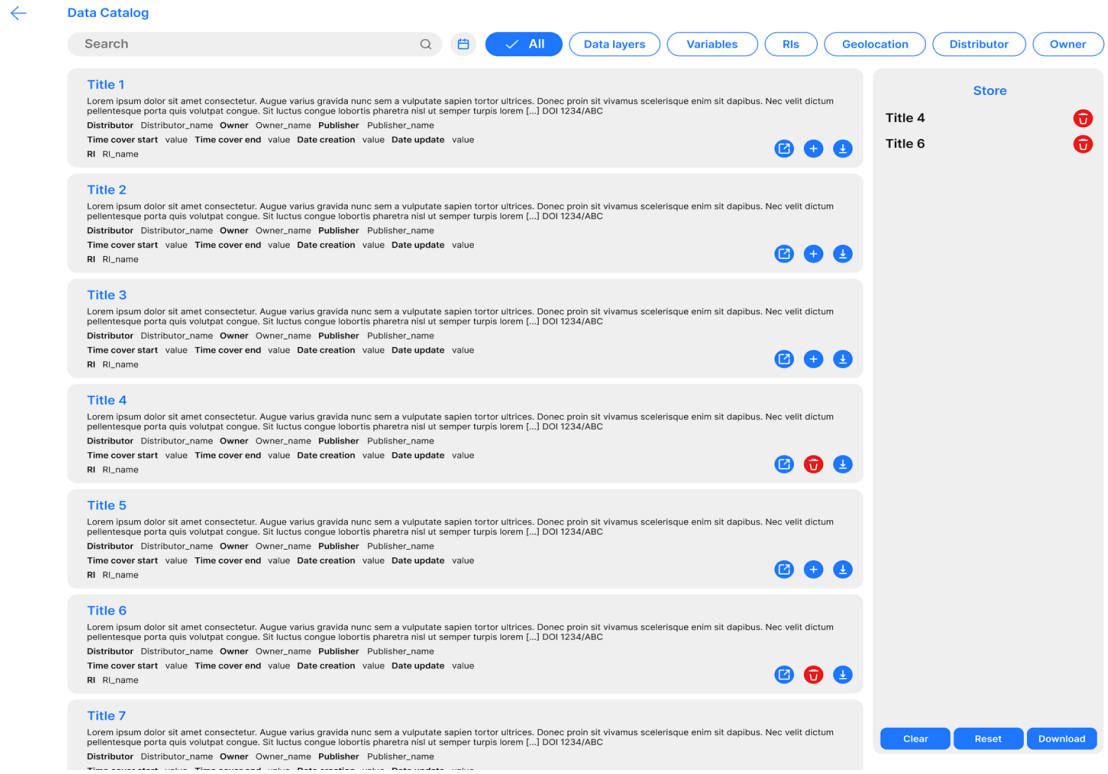

Figure 6. POLARIN DH Catalogue

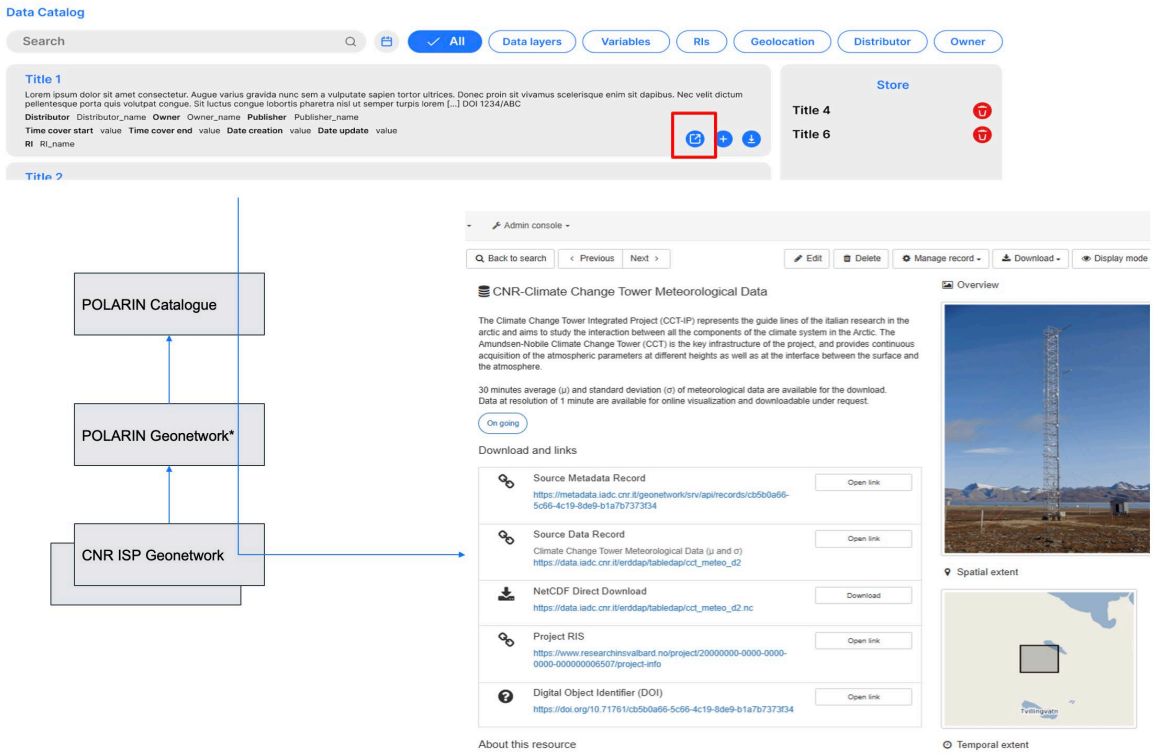

Figure 7. POLARIN DH Catalogue – metadata harvesting and presentation workflow

The POLARIN Data Viewer (POLARIN DV) provides the user with a further level of data interaction and access. The tool (Figure 8) is designed to have a control box to select data by type, RI, parameter, timerange, etc and on top of these there will be a free-text searching tool to wide ist usability. The body shows the dataset in space and offers controls to navigate them in time and depth.

For each dataset a popup, whenever possible, will present plots or graphs.

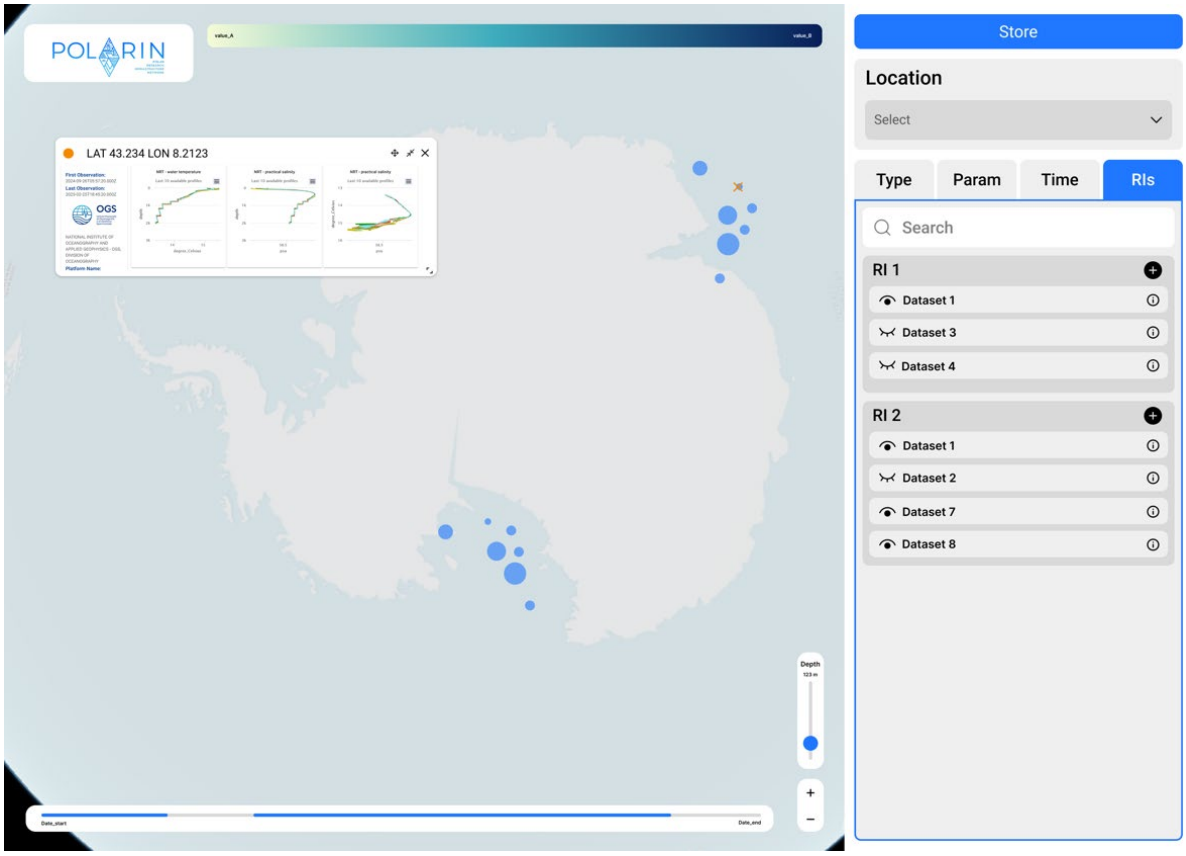

Figure 8. POLARIN DH Data Viewer

4. References

Deliverable D4.1 – POLARIN DMP

Deliverable D4.2 - POLARIN graphic products package for multiple audiences and gap analysis

Deliverable D4.4 - Guidance on dataset granularity

Deliverable D5.1 – A unified semantically consistent virtual data catalogue with machine interfaces