Optimizing Polar Observing with Asset-Level Metadata Interoperability Across Networks

William Manley[1], Roberta Pirazzini[2], and other Members of the SAON Polar Observing Assets Working Group[3]

[1]University of Colorado, [2]Finnish Meteorological Institute, [3]https://polarobservingassets.org

There is broadly recognized need for an integrated Arctic observing system, including a means of identifying overlaps and gaps, a "knowledge map" to clarify directions, and ways to build capacity to better meet observing goals (e.g. AOS, 2020; ASM, 2021; IARPC, 2021; EU-PolarNet, 2022). However, a fundamental challenge exists: Observing assets such as sites, tracks, research stations, projects, and programs are deployed in a diverse and distributed fashion across hundreds of networks and initiatives. At this time, it is difficult to strategically assess, plan, or synthesize because the granular information needed – specific details on activities and infrastructures – is fragmented and incomplete. Most asset-related inventories, catalogs, and portals are limited in thematic or geographic scope. Furthermore, only a fraction of networks share structured information in a way that can be accessed, harmonized, and aggregated for a comprehensive perspective.

To help address this challenge, a new Polar Observing Assets Working Group (POAwg) has been formed under the SAON Committee on Observations and Networks (CON). This group builds upon steps taken by the polar data community for the interoperability of "dataset-level" metadata, but in this case for discovery-level details in "asset-level" metadata (see Table 1). POAwg will identify and promote community-based approaches for the use of relevant standards, controlled vocabularies, crosswalks, federated search, and linkages to operational or scientific datasets. In so doing, its broader goals are to make observing-related metadata – beyond the dataset level – more Findable, Accessible, Interoperable, and Reusable (FAIR; Wilkinson et al., 2016), as well as to help showcase and integrate the summed contributions of multiple systems. For more information, see https://www.polarobservingassets.org. POAwg has identified three tasks that are focused and achievable in the short term.

The first task is to create a registry of polar observing networks – focusing on interoperability parameters and documenting: Asset-related metadata standards (e.g., ISO 19115/19139, WIGOS, INSPIRE EF, etc.), semantic technologies (observing-related vocabularies and ontologies), transfer protocols (e.g., OAI-PMH, CSW, WFS, or custom APIs), and machine-readable endpoints that are currently in practice. Use cases for the registry have been identified and will guide the development process moving forward. The registry will have a frontend allowing users to browse, search, and filter for networks, potentially with a graphical interface illustrating patterns of implementation. As a start, the group has developed a crosswalk for elements across prior inventories by SAON CON and EU-PolarNet, and in a few existing network-level metadata models (e.g. DEIMS-SDR, INTERACT). Current efforts are to prioritize metadata elements and to envision the user experience. The registry of polar observing networks will help to: Clarify best practices for observing-related metadata sharing; establish a

basis for harvesting, aggregation, & federated search; better inform local communities of active efforts nearby; and guide network assessment & planning.

Table 1. Summary of Discovery-Level, Observing-Related Metadata Elements

Network-Level	Project-Level	Site-Level
e.g., GTN-P	*e.g., TSP*	*e.g., borehole*
Network Name Network Description Discipline Observational Parameters Organization Organizational Country Time Range Region Subregion Spatial Extent Contact Info Asset Type Metadata Standards Transfer Protocols Links to organization, network, and data	Funding Agency Funding Program Funding Country Project Title Project ID Discipline Region Subregion Location Latitude & Longitude Institution Contact Info Project Start Date Project End Date Links to project summaries, sites, data, and more	Network Name Site Name Site ID Site Description Facility Type Observed Properties Country Location Latitude & Longitude Elevation Site Start Date Site End Date Institution Contact Info Links to network, institution, data, and more

The second task is to build crosswalks and facilitate existing tools for translation across standards. In contrast to a broad range of established metadata standards in use to describe scientific datasets, there are relatively few that pertain to concepts such as observing sites, major infrastructure, transects, and field-based research projects (cf. Habermann, 2018; Wohner et al. 2020). And there are even fewer tools or resources to facilitate mapping of elements from one to another (for example, see ADIwg, 2017). This task will compile crosswalks across the most common standards or custom schemas in use by networks, as identified in Task 1. Outcomes will significantly improve the ability to translate & harmonize for aggregation and federated search, and will inform the deliverable for Task 3.

The third task is to create recommendations for adoption and implementation of established standards and solutions. The practice of building and deploying asset-level metadata catalogs is still in its infancy. Many networks and related organizations have done so, primarily for their own internal goals of resource tracking, logistics, and operational management. However, in many cases these databases (or spreadsheets, or tables on webpages and PDFs) have been constructed from the ground up — with custom approaches that limit utility and impact. To address this, POAwg will create a guidance document for implementation of standards, protocols, vocabularies, crosswalks, open-source platforms, and more. This guide will draw from, and build upon, similar efforts acting outside the high latitudes (e.g., ENVRI-FAIR, 2021; IODE ODIS, 2021; eLTER RI, 2022; JERICO-CORE, 2022; cf. Jones et al., 2021), but tailored to a polar observing audience. The end result will communicate an easier path for networks to

populate, expand, and share metadata catalogs of observing assets — improving overall interoperability while saving considerable time and effort.

Participation in POAwg is open and encouraged. Participants are network coordinators, data managers, and others with familiarity of network assets – and those with experience, or just an interest, in metadata management. Monthly web meetings include presentations, discussion, and collaboration. Sign up at https://www.polarobservingassets.org.

<u>References</u>

ADIwg. 2017. Alaska Data Integration Working Group. https://www.adiwg.org.
AOS. 2020. Arctic Observing Summit: Observing for Action.
        https://arcticobservingsummit.org/summits/aos-2020/.
ASM. 2021. Joint Statement of Ministers: On the occasion of the Third Arctic Science
        Ministerial. https://asm3.org/library/Files/ASM3_Joint_Statement.pdf.
eLTER RI. 2022. Integrated European Long-Term Ecosystem, critical zone and socio-ecological
        Research, Research Infrastructure. https://elter-ri.eu.
ENVRI-FAIR. 2021. Report On Fairness Implementation Activities In The Biodiversity And
        Ecosystem Subdomain.
        https://envri.eu/wp-content/uploads/2021/04/ENVRI-FAIR_D_11-2.pdf
EU-PolarNet. 2022. Procedure for ongoing collection and collation of European Polar observing
        capacities and activities: EU-PolarNet 2 - Co-ordinating and Co-designing the European
        Polar Research Area, Deliverable No. D6.1. in prep.
Habermann, T. 2018. Metadata Life Cycles, Use Cases and Hierarchies: Geosciences 8(5):179.
        http://dx.doi.org/10.3390/geosciences8050179
IARPC and NSTC. 2021. Arctic Research Plan 2022-2026: Interagency Arctic Research Policy
        Committee and National Science and Technology Council.
        https://www.iarpccollaborations.org/plan/.
IODE ODIS. 2021. The Ocean InfoHub Project: IODE Ocean Data and Information System.
        https://book.oceaninfohub.org/index.html.
JERICO-CORE. 2022. Coastal Ocean Resources Environment, JERICO Research
        Infrastructure. https://www.jerico-ri.eu/jerico-core/
Jones, M.B., Richard, S., Vieglais, D., Shepherd, A., Duerr, R., Fils, D., and McGibbney, L.
        2021. Science-on-Schema.org v1.2.0. Zenodo. https://doi.org/10.5281/zenodo.4477164
Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg,
        N., et al. 2016. The FAIR Guiding Principles for scientific data management and
        stewardship. Scientific Data 3:160018. https://doi.org/10.1038/sdata.2016.18
Wohner, C., Peterseil, J., Genazzio, M.A., Guru, S., Hugo, W., and Klug, H. 2020. Towards
        interoperable research site documentation – Recommendations for information models
        and data provision, Ecological Informatics 60:101158.
        https://doi.org/10.1016/j.ecoinf.2020.101158