

US AON Benefit Tool – Data Management Plan

Project Description and Overview

The US Arctic Observing Network's (US AON) [Benefit Tool](#) creates diagrams that illustrate the links between observing systems, data products, and applications and show how they connect to societal benefit. Capabilities and gaps are given numerical ratings and written descriptions throughout the evaluation. The process of creating these diagrams can help managers at those various levels examine the FAIR compliance of their own data and metadata - leading to additional improvements in data management.

Data Management - Overview

The US Arctic Observing Network (US AON) is committed to both the Findable, Accessible, Interoperable, Reusable (FAIR; Wilkinson 2016) and Collective benefit, Authority to Control, Responsibility, Ethics (CARE; Russo Carroll 2019) data principles. Data and data-derived products have been developed on a software tool and website hosted and designed in partnership with the National Snow and Ice Data Center (NSIDC). US AON will work with the US AON Expert Committee on Methods, which includes data management experts, to refine data management practices. The team will also work with the experts to ensure that the data management reflects CARE principles. US AON will make efforts to align and integrate with international efforts, including the Arctic Observing Summit, SAON ROADS process, and the Polar Observing Assets working group (POAwg).

Roles and Responsibilities

Data Management - (name, email, role)

Data Storage (prior to archiving) - (name, email, role)

Metadata and Documentation creation - (name, email, role)

Archiving and Preservation - (name, email, role)

Information to be created & collected

The primary data to be assembled, created, and managed by this project is (1) expert evaluations of the value of Arctic observations, data products, and applications, and shared benefits, and of the links between them (2) a registry/library of objects supporting the Shared Benefit Analyses, and (3) Shared Benefit Analysis tool software. An example of data to be created and collected as part of this process is: Subject matter experts create an analysis and then walk through the results with a US AON staff member. The Shared Benefit Analysis tool also includes a library of objects that describe observations, data products, and applications. Once the analysis is considered final, it is published and the underlying data (e.g. quantitative and qualitative ratings, object-level metadata) is made available.

File formats and standards: As a default, the expert elicitation data, object library, and software will be publicly available via widely accessible specifications and file formats. In all cases, we will ensure that data are stored using the appropriate encoding system such as Unicode Transformation Format-8. This will ensure that all natural language text, including those with non-Latin characters (e.g., Inuktitut syllabics), will be correctly represented.

Data storage and metadata Data supporting the expert elicitations will be hosted at NSIDC, stored in a PostgreSQL database with regular backups, and made downloadable in non-proprietary, interoperable file formats (e.g., JSON, CSV, SQL) where appropriate. Standards compliant metadata records for specific analyses may also be created at NSF's Arctic Data Center as EML (ecological metadata language) to improve findability and interoperability. In addition to the data created, the tool's registry/library of objects will conform to standards-based metadata wherever possible, e.g. discovery level metadata using the schema.org, ISO 19115 metadata standards, or the POLDER *Best Practices*. This will improve interoperability with parallel efforts. Wherever possible ethical, standard, widely adopted vocabularies (e.g. RDF, OGC, Schema.org) will be used to represent concepts and terms.

Software and source code management The Shared Benefit Analysis tool is being developed using an open-source approach. Source code, documentation, and project planning information is available on GitHub, a widely used development platform. Community volunteers interested in contributing to the development and maintenance of *earthaccess* have been and will continue to be given maintainer privileges to enable *earthaccess* to persist and grow without requiring explicit funding agreements. With each release, the source code and associated documentation is automatically archived for long-term storage on Zenodo and issued a new DOI. The latest release will always be available at this DOI: **10.5281/zenodo.8403759**.

Role FAIR and CARE play in data management¹ - For all data activities, FAIR and CARE principles will be applied.

US AON Benefit Tool research team acknowledges Force11 [FAIR Data principles](#)², the GIDA [CARE principles](#)³, and the First Nations [Principles of OCAP](#)⁴, applying these information management frameworks however is most appropriate to protect the interests and considerations of Indigenous peoples, Indigenous communities, and this research project. The FAIR Data principles prioritize research outputs that are findable, accessible, interoperable, and reusable for the broader research community. FAIR's secondary concerns are with regards to reproducibility of science, and the long-term preservation/archiving of scientific data. The CARE principles call for information sharing that addresses collective benefit, authority to control, and responsible and ethical sharing, with an emphasis placed on Indigenous self-determination and sovereignty. The First Nations Principles of OCAP, like the CARE principles, seek to apply best practice with regards to Indigenous and tribal ownership, control, access, and possession of research products on behalf of Indigenous communities, lands, cultural heritage or other

¹ Raymond, V. (2022). Data Management Plans: How FAIR, CARE, and OCAP are applied in this research. (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.7109286>

² Wilkinson, M. D. *et al.* The FAIR guiding principles for scientific data management and stewardship. *Sci. Data*. 3, 160018 (2016).

³ Carroll, SR, Rodriguez-Lonebear, D and Martinez, A. 2019. Indigenous Data Governance: Strategies from United States Native Nations. *Data Science Journal*, 18(31): 1–15. DOI: <https://doi.org/10.5334/dsj-2019-031>

⁴ First Nation's Information Governance Centre. 2018. The First Nations Principles of OCAP®. Available at <https://fnigc.ca/ocap>

— DRAFT —

intellectual property. Below, the US AON Benefit Tool research team has described how these guidelines and principles are being applied (or not applied) with regards to this research and its information management.

Guideline/Principle	Framework	Priority in this research project	Notes / Details
Findable	FAIR	Medium	
Accessible / Access	FAIR / OCAP	High	
Interoperable	FAIR	Medium	This is a priority and efforts are underway, but interoperability of metadata is a persistent challenge
Re-usable	FAIR	High	
Collective Benefit	CARE	High	
Authority to Control / Control	CARE / OCAP	High	
Responsibility	CARE	High	
Ethical	CARE	High	
Ownership (sovereign right)	OCAP	Medium	
Possession (physical control of data)	OCAP	Low	This is a future goal. For now, with all the data being held within US AON's software tool, there is not a substantial path for community possession of data.

Policies for access, sharing, and reuse: As a default, data/products/software produced through US AON's expert elicitation process will be fully open for access and reuse unless expert and/or community preferences override this policy on a case-by-case basis. In the case of sensitive or confidential data, US AON will work collaboratively to document an appropriate policy for access, sharing, informed consent, archival, and reuse. As a general policy, CARE principles will take precedence over FAIR principles.

Plans for archiving data: We will archive the data (excluding sensitive or Indigenous-controlled data) and software code with the National Science Foundation's Arctic Data Center to ensure continuity beyond the lifetime of the project. The underlying SBAT data will be archived at the NSF Arctic Data Center.