



NATIONAL SCIENCE FOUNDATION
2415 EISENHOWER AVENUE
ALEXANDRIA, VIRGINIA 22314

NSF 22-106

Dear Colleague Letter: Office of Polar Programs Data, Code, and Sample Management Policy

July 14, 2022

Dear Colleagues:

This document replaces [NSF 16-055 Dear Colleague Letter: Data Management and Data Reporting Requirements for Research Awards Supported by the Office of Polar Programs, May 9, 2016](#).

INTRODUCTION

The recommendations and requirements laid out in this letter aim to advance open polar data to maximize the benefit of NSF's investments in research, facilitate transparency and replicability in polar science, and increase the impact of polar research. Office of Polar Programs (OPP) PIs are directed to align their research and dissemination plans and activities with the [FAIR data principles \(Findability, Accessibility, Interoperability, and Reusability\)](#) and [CARE principles for Indigenous data governance \(Collective Benefit, Authority to Control, Responsibility, Ethics\)](#). In this regard, research software/code is identified as a research object and outcome that can also help make data more interoperable and reusable. OPP also recognizes the effort that goes into data and sample management and therefore, beyond [encouraging data and sample reuse](#), OPP expects appropriate authorship, attribution, and citation of data, samples, and code.

1. OPP policy requires that metadata files, full final data sets, derived data products, physical samples, and relevant software/code be deposited in a long-lived and publicly accessible archive.
2. Samples, data, and associated code must be archived within two years of collection/creation or by the end of the award, whichever comes first, unless required sooner by a particular program or award condition.
3. Data Management Plans (DMPs) that rely on self-publication on personal, lab, or university websites/servers/archives in lieu of appropriate repositories **are noncompliant** and will require revision or the proposal may be returned without review.

OPP suggests early and regular contact with repositories and requires that PIs provide updates on the implementation of their DMPs in annual and final project reports, including any changes from the original DMP, which must be approved by the cognizant Program Officer. Full details and exceptions are detailed below.

PROPOSAL REQUIREMENTS

All proposals must include a DMP that describes how the project will provide open, ethical, and timely access to quality-controlled and fully documented data, samples, software/code, and products during and at the conclusion of the project. The DMP must be included as a Supplementary Document and be consistent with [NSF's policy on dissemination and sharing of research results](#) and [NSF Proposal & Award Policies & Procedures Guide \(PAPPG\)](#). Data management plans that are noncompliant will require revision and the proposal may be returned without review. The costs associated with data management are eligible for funding and may be included in the budget of a proposal submitted to OPP.

The DMP for OPP proposals must detail the timeline for final archiving, identify roles and responsibilities for data management, and indicate the planned long-lived repository/repositories for metadata, data, code, and physical samples. For proposals that involve the collection or generation of specimens, DMPs must include information on how specimens and associated data will be accessioned into an established, long-lived collection, if available. DMP templates, tools, expertise, and support are available from [OPP-funded data centers](#), such as the [Arctic Data Center](#) and [US Antarctic Program Data Center](#).

OPP recognizes that intermediate or experimental data products may also be produced in the course of research, and that some research will produce data at unwieldy volumes (e.g., modeling studies). DMPs should identify what raw and derived products (including educational, instructional, and training materials, if appropriate) will be made available, and PIs should archive code together with the appropriate data in their most useable forms; long-lived, open data formats should be used where possible.

DMPs that rely on self-publication on personal, lab, or university websites/servers in lieu of appropriate repositories are considered noncompliant and will require revision or the proposal may be returned without review. Specific requirements for both Arctic and Antarctic research are below, and [available resources are provided on the OPP webpage](#). PIs should describe how data management will facilitate ethical, robust, reliable, and reproducible research and improve access to research results.

OPP ARCHIVING REQUIREMENTS

The Office of Polar Programs policy requires that metadata files, samples, full final data sets, derived data products, and relevant software/code be deposited in a long-lived and publicly

accessible archive, well documented, assigned with interlinked persistent unique identifiers (e.g., [DOI](#)), and labeled with an open/appropriate license. PIs should make use of community-accepted, disciplinary repositories where possible and seek early contact with repository managers to discuss (implementation of) data management plans. Information regarding appropriate data centers can be found on [the Office of Polar Programs website](#) or through contact with the cognizant Program Officer(s).

Except as noted below, all samples, data, and associated code must be provided to an appropriate repository within two years of collection/creation or by the end of the award, whichever comes first.

Any limitation on access to the information beyond these dates must be based on a compelling justification and documented in the DMP and approved by the cognizant Program Officer. Exceptions to the data management requirements may be granted for social science data, Indigenous knowledge, other Indigenous data, and where sensitivity, privacy, sovereignty, and/or intellectual property rights might take precedence. Such requested exceptions must be documented in the OPP-approved DMP.

ADDITIONAL ARCHIVING REQUIREMENTS FOR ARCTIC SCIENCES

Metadata for all Arctic supported data sets must be submitted to the NSF [Arctic Data Center](#).

Arctic Observing Network (AON) data are public and not subject to any embargo period. All AON data must be quality controlled and deposited in a long-lived and publicly accessible archive within 6 months of collection. All AON datasets and derived data products must be accompanied by a metadata profile and full documentation. PIs should report data usage statistics in their annual and final reports.

DMPs for projects involving social science and/or Indigenous knowledge data should describe plans for non-sensitive data deposit in the [Arctic Data Center](#), or in a discipline-specific durable, publicly accessible, digital repository. DMPs should also explicitly describe potential data sensitivities and outline plans for archiving such data. Options include data content labels, data embargoes, archiving with a repository capable of handling sensitive data, and/or deposit with the community of origin. DMPs for projects involving Indigenous knowledge must address data sovereignty with respect to the CARE Principles. PIs should describe how they will collaboratively determine procedures for data collection, handling, and archiving (for example, through informed consent, tribal consultation, or community meetings). In the absence of a digital repository, PIs should prioritize a repository accessible to the community of origin.

ADDITIONAL ARCHIVING REQUIREMENTS FOR ANTARCTIC SCIENCES

Upon award, the awardee must create a project page with the [US Antarctic Program Data](#)

Center (USAP-DC); registration should include submission of the DMP agreed to at the time of award. Metadata for all Antarctic supported data sets and derived data products must be submitted or linked to the USAP-DC project page. Project registration within the USAP-DC will ensure registration within the Antarctic Metadata Directory (AMD) and will fulfill data sharing obligations under the Antarctic Treaty. Proof of submission must be included in the Final Project Report to NSF in the form of a link to the metadata and data archives. Antarctic PIs should update their USAP-DC project page over the course of their award, as well as include any updated DMP on their project page.

Antarctic measurements and observations that are collected routinely and automatically as part of ongoing projects or operations are expected to be made available to the community in their raw instrumental output format without any artificial delay.

REPORTING REQUIREMENTS

PIs are required to provide updates on the implementation of their DMPs in annual and final project reports, including any changes from the original DMP, in the sections for "Accomplishments" and/or "Changes," as appropriate. Persistent unique identifiers / links for archived materials in the ADC/USAP-DC should be included in these reports in the section titled "Products-Websites." PIs are reminded that they are also expected to include information about archived products in the final [Project Outcomes Report](#).

VALIDITY & DISCLAIMER

This DCL applies to proposals and supplements submitted after 90 days from publication of this DCL and until it is replaced. The above guidelines are not intended to replace the guidance given in the [NSF Proposal & Award Policies & Procedures Guide](#) (PAPPG) and specific program solicitations. In any perceived conflict, the PAPPG or the solicitation will take precedence, as appropriate, for the proposal.

Questions regarding this policy should be directed to oppdata@nsf.gov and the relevant cognizant Program Officer(s).

Sincerely,

Alexandra R. Isern
Assistant Director for Geosciences