

EAR Data and Sample Policy
Division of Earth Sciences
National Science Foundation
Updated July 2023

Overview

The Division of Earth Sciences (EAR) is committed to achieving the broadest benefit of its research investments. Adherence to open, inclusive, and transparent research practices, including those articulated through the *FAIR Guiding Principles (Findable, Accessible, Interoperable, and Reusable)*¹ and the *CARE Principles for Indigenous Data Governance (Collective Benefit, Authority to Control, Responsibility, Ethics)*², is critical for maximizing the scientific value of data, samples, and other research products supported through EAR awards. The 2020 National Academies of Sciences, Engineering, and Medicine (NASEM) Report, *A Vision for NSF Earth Sciences 2020-2030: Earth in Time*, noted that “FAIR data standards will improve the longevity, utility, and impact of EAR-funded data.”³ In 2022, a memorandum published by the Office of Science and Technology Policy (OSTP) articulated the importance of ensuring free, immediate, and equitable access to federally funded research.⁴ In 2023, NSF published an updated Public Access Plan 2.0, which describes NSF’s expected approach to achieving the goals of the OSTP memorandum.⁵ EAR believes that projects committed to open sharing of results and conducted in alignment with *FAIR* and *CARE* principles will accelerate scientific discovery, broaden data access, and ensure reproducibility and replicability of research in the Earth Sciences. Deposit of data and associated metadata (including sample-based metadata) in repositories that fulfill FAIR principles, as articulated in the OSTP “Desirable Characteristics of Data Repositories for Federally Funded Research” guidelines, is a straightforward way for Earth Scientists to adhere to open science principles.⁶

This document defines data and sample policies for all proposals submitted to, and awards managed by, EAR programs. These policies supplement NSF-wide requirements in the Proposal and Award Policies and Procedures Guide (PAPPG).⁷ In the PAPPG, NSF requires that all proposals include a Data Management Plan (DMP) describing how the project will conform to

¹ Wilkinson, M., Dumontier, M., Aalbersberg, I., *et al.* (2016), The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

² Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., *et al.* (2020), The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19(1), 43. <http://doi.org/10.5334/dsj-2020-043>

³ National Academies of Sciences, Engineering, and Medicine (2020), *A Vision for NSF Earth Sciences 2020-2030: Earth in Time*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25761>

⁴ White House Office of Science and Technology Policy (August 25, 2022), Ensuring Free, Immediate, and Equitable Access to Federally Funded Research. <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>

⁵ NSF 23-104 (February 21, 2023), NSF Public Access Plan 2.0: Ensuring Open, Immediate and Equitable Access to National Science Foundation Funded Research. <https://new.nsf.gov/public-access>

⁶ White House Office of Science and Technology Policy (2022), Desirable Characteristics of Data Repositories for Federally Funded Research. <https://doi.org/10.5479/10088/113528>

⁷ NSF Proposal & Award Policies & Procedures Guide (PAPPG). https://www.nsf.gov/publications/pub_summ.jsp?ods_key=papppg

the NSF PAPPG policy on dissemination and sharing of research results.⁸ NSF considers the DMP to be an integral part of the proposal, to be considered under Intellectual Merit and/or Broader Impacts, as appropriate, and as part of the proposal evaluation process. As such, EAR Program Directors ask proposal reviewers to carefully evaluate proposal DMPs relative to the guidance set forth in this policy, the PAPPG, relevant program solicitation(s), community-specific standards, and open science principles (e.g., FAIR and CARE). During the period of the award, EAR awardees are responsible for adhering to the DMP, and EAR Program Directors monitor such adherence through Annual and Final Project Reports.

EAR requirements for Data Management Plans (DMPs)

This section summarizes key requirements for DMPs as described in the PAPPG and supplemented by this EAR policy. Specific guidance on how to achieve these requirements is provided below in “Data Management Plan (DMP) content for EAR proposals.”

EAR requirements:

1. Proposals must include a document of no more than two pages, titled "Data Management Plan," in the supplementary documentation section of the proposal. In cases of collaborative proposals or proposals involving subawards, the lead PI submits a single DMP for the entire project. In cases where no data or samples will be produced (for example, in conference proposals), the DMP may simply state that no detailed plan is needed, as long as such statement is clearly explained.
2. The DMP should demonstrate consistency with open science principles (e.g., FAIR and CARE) and community-specific standards. While variation in DMPs is expected across research communities, each DMP should be appropriate for the data and samples being generated and reflect community best practices. Deviations from open science principles and/or community standards must be justified.
3. The DMP must address plans for all types of data and samples to be collected and/or generated through the proposal, including roles and responsibilities for managing such data and samples, as well as relevant metadata standards to be followed. EAR defines “data” and “samples” expansively while acknowledging differences across disciplines. Possible types of “data” to be addressed in the DMP include, but are not limited to: observational, experimental, analytical, and model outputs; derived and compiled datasets; software and code; educational materials; and any other relevant digital products resulting from the project. Possible types of “samples” to be addressed in the DMP include, but are not limited to: physical samples and collections; drilling cores; specimens; and any other relevant physical, chemical, and/or biological materials resulting from the project. For purposes of this policy, sample-derived digital products are considered “data.”
4. All new data resulting from the project must be made publicly accessible within two (2) years after completion of data collection or generation, via appropriate long-lived FAIR-aligned repositories. Expected timelines for data collection or generation may vary by

⁸ NSF has announced its intent to retitle “Data Management Plans (DMPs)” as “Data Management and Sharing Plans (DMSPs).” This EAR Data and Sample Policy applies to both current DMPs and expected future DMSPs.

data type and should align with appropriate disciplinary expectations. All new data collected via continuing observations, large-scale community projects, or RAPID awards must be made accessible as close to the time of initial collection as is practicable. All data in support of peer-reviewed scholarly publications resulting from the project must also be made publicly accessible at or before the time of publication. Exceptions to this policy must be justified (e.g., if an appropriate repository does not exist, or if data access must be restricted). ***“Data available upon request” is not acceptable.***

5. Metadata describing all new samples resulting from the project must be publicly indexed within two (2) years after sample collection is considered complete, via appropriate long-lived FAIR-aligned repositories. Metadata describing samples collected via continuing observations, large-scale community projects, or RAPID awards must be indexed and made accessible as close to the time of collection as is practicable. All sample metadata in support of peer-reviewed scholarly publications must also be publicly indexed at or before the time of publication. Publicly indexed sample metadata should specify provisions for sample access, including the expected period and location of sample preservation, preferably via a repository appropriate for the specific sample type. The samples themselves should also be made publicly accessible within the above timeframes; situations in which samples cannot be made publicly accessible should be explained.

Some programs within the Division of Earth Sciences have specific guidelines regarding data and sample acquisition, permitting, and repository selection. Please see the relevant program solicitation(s) and consult with the cognizant Program Director(s) for further information.

Data Management Plan (DMP) content for EAR proposals

To fulfill the requirements described above and to ensure alignment with open science principles, PIs are encouraged to structure their DMPs around the following 2 sections:

1. Data and sample types. Describe the types of data and samples expected to result from the proposed work:

- a) List the types of data and samples to be collected and/or generated. The listing of each data/sample type should briefly identify what metadata will be provided and when data/sample preparation will be considered complete. (Definitions of “data” and “samples” are explained above within “EAR requirements.”) For proposals providing community-serving infrastructure or research services, the DMP should describe the data/sample types to be managed and what guidance or support will be provided to help users meet their data/sample sharing obligations. EAR recognizes that data/samples may undergo multiple transformations in the research process (including destructive analyses), and disciplinary expectations for assignment of metadata and retention of intermediate data and sample products may vary.
- b) For each data or sample type, identify which personnel and institution(s) will be designated for its management, including contingency plans for the departure of key personnel from the project. For collaborative projects, PI(s) of the award(s) associated with the designated personnel and institution(s) are ultimately responsible for overseeing and reporting on their data and sample management activities.

2. Data/sample deposit, access, and preservation. Describe how each type of data or sample will be deposited, made accessible, and preserved:

- a) For each data type listed, identify an appropriate long-lived FAIR-aligned repository for data deposit, the timeframe for public data access, and the expected period of data preservation. For each sample type listed, identify an appropriate long-lived FAIR-aligned repository for indexing sample metadata, the location for sample storage (preferably a repository appropriate for the specific sample type), and the expected period of sample preservation. (Required timeframes for data/sample access are specified above within “EAR Requirements.”) Many repositories commit to preserve access to data/samples indefinitely; any deviations from this expectation should be explained. PIs are encouraged to coordinate with designated repositories in advance of planned data/sample submission.
- b) In most cases, it is sufficient for the DMP to identify the repository(ies) to be used and the timeframe for access and preservation for each type of data/sample identified. In these cases, the selected repository(ies) should align with FAIR principles and community-specific standards. Occasionally, appropriate long-lived FAIR-aligned repositories do not exist for certain types of data or samples. In such cases, it may be necessary to adopt alternative approaches to data access and retention, such as via use of a local computer server. In such cases, the DMP should explain how the proposed approach fulfills important attributes for FAIR-aligned repositories, consistent with OSTP guidance on “Desirable Characteristics of Data Repositories for Federally Funded Research.”⁹ These attributes include, but are not limited to, the following:
 - *Findability.* Data should be findable via standard search tools, such as through the assignment of globally unique persistent identifiers (e.g., Digital Object Identifiers (DOIs) and International Geo Sample Numbers (IGSNs)) and rich metadata that is indexed in a searchable resource.
 - *Accessibility.* Data should be publicly accessible to other researchers, at no more than incremental cost, within the specified timeframe. Any data access limitations must be justified. “Data available upon request” is not acceptable.
 - *Interoperability.* To ensure interoperability, data should be described via appropriate metadata standards, in alignment with expectations of the associated scientific discipline(s).
 - *Reusability.* To facilitate the broadest possible data reuse, data should be assigned clear and accessible usage licenses and metadata descriptors that identify provenance. EAR expects the adoption of unrestrictive open licenses except with specific justification.

Costs associated with data and sample management

NSF recognizes that data management activities require time and expense, including upfront curation costs that may be charged by data and sample repositories. Expenditures for such

⁹ White House Office of Science and Technology Policy (2022), Desirable Characteristics of Data Repositories for Federally Funded Research. <https://doi.org/10.5479/10088/113528>

activities are allowable and should be documented and budgeted appropriately. See NSF 19-069, “Dear Colleague Letter: Effective Practices for Data,” for further guidance.¹⁰

Award reporting

PIs and co-PIs are responsible for providing updates to NSF, within Annual and Final Project Reports, on data and sample management activities carried out by personnel and institution(s) associated with their awards, as designated in the DMP. PIs and co-PIs are also expected to provide updates to the general public by reporting on data and sample management activities within the Project Outcomes Report and by indexing research products within the NSF Public Access Repository (NSF-PAR).¹¹ The Project Outcomes Report and NSF-PAR entries can be viewed on the public-facing award page.

Annual and Final Project Reports to NSF should contain the following information:

- Describe ongoing data and sample management activities, including data/samples in preparation that have not yet been shared, in the “Accomplishments” section of the report. Describe any significant deviations from the proposal DMP in the “Changes/Problems” section of the report.
- For data and/or samples that have been made available within the prior year, list these within the “Products” section of the report. Such listings, which may be facilitated by indexing associated metadata within the NSF-PAR, should include globally unique persistent identifiers, such as Digital Object Identifiers (DOIs) or International Geo Sample Numbers (IGSNs).
- For data or samples that will be made available after submission of the Final Project Report, the PI should include plans for data and sample availability in the “Accomplishments” section of the report. Once the data and/or samples are made available, the PI should index associated metadata in the NSF-PAR and notify the cognizant Program Director by e-mail.

Annual and Final Project Reports that do not adequately address data and sample management activities may be returned to PIs to provide the required information.

Implementation of prior DMPs may be considered during evaluation of subsequent proposals. Description of data and sample management for prior awards should be included in the “Results from Prior NSF Support” section of the Project Description of the proposal. When appropriate, this section should include evidence that data, samples, and other products have been made accessible in appropriate repositories. All products that are specifically listed in the Results from Prior NSF Support section must be referenced in the References Cited section of the proposal with globally unique persistent identifiers (e.g., DOIs or IGSNs).

¹⁰ NSF 19-069 (May 20, 2019), Dear Colleague Letter: Effective Practices for Data.

https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf19069

¹¹ NSF Public Access Repository (NSF-PAR). <https://par.nsf.gov/>; see also Research.gov guidance on adding datasets to NSF-PAR:

https://resources.research.gov/common/attachment/Desktop/How_to_Add_Datasets_to_NSF-PAR.pdf

Resources

To facilitate adherence to the EAR Data and Sample policy and open science practices, EAR maintains a website with a list of resources for EAR proposers and awardees: [Resources for EAR Data and Sample Policy](#)