

National Science Foundation (nsf.gov): NSF-EAR: Earth Sciences

Data and sample types

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.

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.

Data/sample deposit, access, and preservation

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.

Guidance:

NSF EAR Guidance:

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 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.

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.

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.

Guidance:

[White House Office of Science and Technology Policy \(2022\), Desirable Characteristics of Data Repositories for Federally Funded Research.](#)