**National Science Foundation GENERIC Data Management Plan Template**

[This plan is based on the "NSF-GEN: Generic" template provided by National Science Foundation (NSF) - (ver: 6, pub: 2021-10-25).]

***Instructions for using this template****: highlighted in yellow are the section headers for the information to address in the data management plan. Under each section header in this document is guidance for the content of that section. The intent of the guidance is as a prompt to help you consider all the types of data produced and how to manage the data. With the exception of the section headers,* ***delete******all the guidance and prompts from your final document, including these instructions and the information above this paragraph.***

**DATA MANAGEMENT PLAN**

### Types of data produced

**The types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project.**

Guidance about data description

* Give a summary of the data you will collect or create, noting the content, coverage and data type, e.g., tabular data, survey data, experimental measurements, models, software, audiovisual data, physical samples, etc.
* Consider how your data could complement and integrate with existing data, or whether there are any existing data or methods that you could reuse.
* Indicate which data are of long-term value and should be shared and/or preserved.
* If purchasing or reusing existing data, explain how issues such as copyright and IPR have been addressed. You should aim to minimize any restrictions on the reuse (and subsequent sharing) of third-party data.

Guidance about data format

* Clearly note what format(s) your data will be in, e.g., plain text (.txt), comma-separated values (.csv), geo-referenced TIFF (.tif, .tfw).
* Explain why you have chosen certain formats. Decisions may be based on staff expertise, a preference for open formats, the standards accepted by data centers, or widespread usage within a given community.
* Using standardized, interchangeable, or open formats ensures the long-term usability of data; these are recommended for sharing and archiving.
* See DataONE Best Practices for [file formats](https://www.dataone.org/best-practices/document-and-store-data-using-stable-file-formats).

### Data and metadata standards

**The standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies).**

Follow data format guidance above

Guidance about data volume

* Note what volume of data you will create in MB/GB/TB. Indicate the proportions of raw data, processed data, and other secondary outputs (e.g., reports).
* Consider the implications of data volumes in terms of storage, access, and preservation. Do you need to include additional costs?
* Consider whether the scale of the data will pose challenges when sharing or transferring data between sites; if so, how will you address these challenges?

Guidance about metadata & documentation

* What metadata will be provided to help others identify and discover the data?
* Researchers are strongly encouraged to use community metadata standards where these are in place. The Research Data Alliance offers a [Directory of Metadata Standards](http://rd-alliance.github.io/metadata-directory/). Data repositories may also provide guidance about appropriate metadata standards.
* Consider what other documentation is needed to enable reuse. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data, and software used to collect and/or process the data.
* Consider how you will capture this information and where it will be recorded, e.g., in a database with links to each item, in a "readme" text file, in file headers, etc.

### Policies for access and sharing

**Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements.**

Guidance on Ethics & Privacy

* Investigators carrying out research involving human participants should request consent to preserve and share the data. Do not just ask for permission to use the data in your study or make unnecessary promises to delete it at the end.
* Consider how you will protect the identity of participants, e.g., via anonymization or using managed access procedures.
* Ethical issues may affect how you store and transfer data, who can see/use it, and how long it is kept. You should demonstrate that you are aware of this and have planned accordingly.
* See [ICPSR approach to confidentiality](http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/confidentiality/index.html) and Health Insurance Portability and Accountability Act [(HIPAA) regulations for health research](https://privacyruleandresearch.nih.gov/).

Guidance on intellectual property rights

* Investigators carrying out research involving human participants should request consent to preserve and share the data. Do not just ask for permission to use the data in your study or make unnecessary promises to delete it at the end.
* Consider how you will protect the identity of participants, e.g., via anonymization or using managed access procedures.
* Ethical issues may affect how you store and transfer data, who can see/use it, and how long it is kept. You should demonstrate that you are aware of this and have planned accordingly.
* See [ICPSR approach to confidentiality](http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/confidentiality/index.html) and Health Insurance Portability and Accountability Act [(HIPAA) regulations for health research](https://privacyruleandresearch.nih.gov/).

Guidance on storage & security

* Describe where the data will be stored and backed up during the course of research activities. This may vary if you are doing fieldwork or working across multiple sites so explain each procedure.
* Identify who will be responsible for backup and how often this will be performed. The use of robust, managed storage with automatic backup, for example, that provided by university IT teams, is preferable. Storing data on laptops, computer hard drives, or external storage devices alone is very risky.
* See the DataONE Best Practices for [storage](https://www.dataone.org/best-practices/storage).
* Also consider data security, particularly if your data is sensitive e.g., detailed personal data, politically sensitive information or trade secrets. Note the main risks and how these will be managed. Also note whether any institutional data security policies are in place.
* Identify any formal standards that you will comply with, e.g., ISO 27001. See the DCC Briefing Paper on Information Security Management - [ISO 27000](http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/information-security-management-iso-27000-iso-27k-s) and UK Data Service guidance on [data security](https://www.ukdataservice.ac.uk/manage-data/store/security).

Guidance on data sharing

* How will you share the data, e.g., deposit in a data repository, use a secure data service, handle data requests directly, or use another mechanism? The methods used will depend on a number of factors such as the type, size, complexity, and sensitivity of the data.
* When will you make the data available? Research funders expect timely release. They typically allow embargoes but not prolonged exclusive use.
* Who will be able to use your data? If you need to restrict access to certain communities or apply data sharing agreements, explain why.
* Consider strategies to minimize restrictions on sharing. These may include anonymizing or aggregating data, gaining participant consent for data sharing, gaining copyright permissions, and agreeing a limited embargo period.
* How might your data be reused in other contexts? Where there is potential for reuse, you should use standards and formats that facilitate this, and ensure that appropriate metadata is available online so your data can be discovered. Persistent identifiers should be applied so people can reliably and efficiently find your data. They also help you to track citations and reuse.

### Policies for re-use, re-distribution, derivatives

**Policies and provisions for re-use, re-distribution, and the production of derivatives.**

See guidance above for Access and Sharing

### Plans for archiving and preservation

**Plans for archiving data, samples, and other research products, and for preservation of access to them.**

Guidance for data repository

* Where will the data be deposited? If you do not propose to use an established repository, the data management plan should demonstrate that the data can be curated effectively beyond the lifetime of the grant.
* It helps to show that you have consulted with the repository to understand their policies and procedures, including any metadata standards, and costs involved.
* An international list of data repositories is available via [re3data](http://www.re3data.org/) and some universities or publishers provide lists of recommendations, e.g., [PLOS ONE recommended repositories](http://journals.plos.org/plosone/s/data-availability#loc-recommended-repositories).

Guidance for Preservation

* Outline the plans for data sharing and preservation - how long will the data be retained and where will it be archived? Will additional resources be needed to prepare data for deposit or meet any charges from data repositories?
* See the DataONE Best Practices for [Identifying data with long-term value](https://www.dataone.org/best-practices/identify-data-long-term-value).

Budget

* Carefully consider and justify any resources needed to deliver the plan. These may include storage costs, hardware, staff time, costs of preparing data for deposit, and repository charges.
* Outline any relevant technical expertise, support, and training that is likely to be required and how it will be acquired.
* If you are not depositing in a data repository, ensure you have appropriate resources and systems in place to share and preserve the data. See UK Data Service guidance on [costing data management](https://www.ukdataservice.ac.uk/manage-data/plan/costing).