

## FH Data Science Lab NIH Data Management and Sharing Plan Checklist

(adapted from NIH DMSP Guidance Working Group, <https://osf.io/uadxr/>)

For latest versions of docs <https://hutchdatascience.org/projects/nihdatasharing/>

To be notified of relevant updates sign up: <https://fhdata.substack.com/>

**Note that this is intended to be a collection of items to potentially include; if components of this checklist do not apply to your data you do not need to include them!**

**Disclaimer: These materials have been developed at Fred Hutch for the purpose of assisting our faculty scientists with the NIH data sharing policy requirements enacted in 2023. While we are confident in the content of these materials, we make no guarantee that they will fully comply with NIH's policy. We will revise these materials on an on-going basis, as appropriate.**

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on [sharing.nih.gov](https://sharing.nih.gov). The Plan is recommended not to exceed two pages.

### Element 1: Data Types

*Summarize the scientific data necessary to validate your findings.*

List or create a table to describe the datasets that will be created or used as part of the study, including:

- Data type, format, size, and number of files (estimate quantities as necessary)
- Which datasets will be shared
- The level of aggregation, de-identification, or processing/cleaning that will be done prior to sharing
- The source of any secondary data, previously collected data reused in this project
- List the metadata and other documentation that will be shared with your data to facilitate interpretation

### Element 2: Related Tools, Software, and/or Code

*Identify tools, software, and/or code necessary to access or manipulate the shared data.*

State whether or not specialized tools are needed. For those listed include:

- Version numbers and operating systems
- How they can be accessed (i.e., open source and freely available, generally available for a fee in the marketplace, or available only from the research team or some other source)
- How long they will be available

### Element 3: Standards

*List the standards that will be used for sharing the data and metadata*

State whether or not there are data standards for your field that are applicable to your project. Typical data standards include:

- Metadata schemas
- Standard terminologies
- Content/Encoding standards
- Common data elements

Standards sometimes are data type specific, arise from the data repository in which it is shared, or the [FAIR guidelines](#) can provide guidance when none is available.

### Element 4: Data Preservation, Access, and Associated Timelines

*Provide details and timelines for sharing and preserving data for long term usability.*

- Name of the repository(ies) where data will be archived (note: a specific NIH repository may be required in the funding opportunity announcement)
- Specify which type of unique identifier is used by the repository, if known
- State when the data will be made available (portions of the data may be released at different times). Timelines required by the policy are:
  - Data will be available when the work is published or the award/support period ends (whichever comes first), OR
  - Data will be made available earlier
- State the minimum number of years data will be available, based on repository policies

### Element 5: Access, Distribution, or Reuse Considerations

*Describe how sharing will be maximized while respecting restrictions such as Legal, Technical or Ethical restrictions.*

- Consider whether data can be shared with access controls or, if there are intellectual

property concerns, an embargo period, rather than refraining from sharing altogether

- If you have human subjects data, describe how you will protect the privacy, rights, and confidentiality of study participants (de-identification, etc.)

**Element 6: Oversight**

*Identify who will be responsible for plan compliance and oversight.*

- List names and titles/roles of everyone responsible for complying with the data management plan
- State how often compliance with the data management plan will be verified.

**Summary:**

Try to create a plan that is feasible for you and your team to implement, appropriate for your data type and study purpose, and has FAIR data sharing principles in mind. Where sharing is infeasible or needs restrictions to be legally or ethically acceptable, be specific about what CAN be done.