Data Management
Excerpted from: https://www.nsf.gov/bio/pubs/BIODMP_Guidance.pdf

CONTENT of the DMPs Data Management Plans should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results3. DMPs submitted to BIO programs should describe how PI(s) will manage data (digital and analog) and physical materials (samples and collections) gathered or generated during the time of the award. This should include description of data handling processes to protect the data (e.g. to ensure quality and/or security), as well as preparations for dissemination and access after the period of the award.

DMPs must include detail sufficient for evaluation of the plan (and past performance if any) during merit review. To facilitate the merit review process and post-award management, and as appropriate, please organize the DMP as follows:

1. Describe the types of data, physical samples or collections, software, curriculum materials, and other materials to be produced in the course of the project. (For collaborative proposals, the DMP must cover all the
various data types being collected by each collaborator
2. Describe the standards to be used for all the data types anticipated, including data or file format and metadata. [Note: Where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies.]
3. Describe the roles and responsibilities of all parties with respect to the management of the data (including contingency plans for the departure of key personnel from the project).
4. Describe the dissemination methods that will be used to make data and metadata available to others during the period of the award, and any modifications or additional technical information regarding data access after the grant ends.
5. Describe the PI's policies for data sharing, public access and re-use, including re-distribution by others and the production of derivatives. Where appropriate, include provisions for protection of privacy, confidentiality, security, intellectual property rights and other rights.
6. Where relevant, describe plans for archiving data, samples, software, and other research products, and for on-going access to these products through their lifecycle of usefulness to research and education. Consider which data (or research products) will be deposited for long-term access and where. (What physical and/or cyber resources and facilities (including third party resources) will be used to store and preserve the data after the grant ends?)

POST-AWARD MANAGEMENT As noted above, after an award is made, implementation of the DMP will be monitored through the annual and final report process by BIO Program Directors and Committees of Visitors. Annual project reports required for all NSF multi-year awards must include information about progress made in data management and sharing of research products (e.g., identifier or accession numbers for data sets, citations of relevant publications, conference proceedings, and other types of data sharing and dissemination).

Final project reports required for all NSF awards should describe the implementation of the DMP including any changes from the original DMP and contain the following information: • the data produced during the award period, and which data that will be retained after the award expires; • how the data is to be disseminated and made available for sharing; • the standards that will be used to make the data available to others, including data format and any metadata; • where the data generated by

the project has been deposited/is being stored for long-term public access, including data identifiers and/or accession numbers, and current URLs.

Data Storage:
Excerpted from: https://purr.purdue.edu/dmp/dmpoverview

A new project begins with a quota of 10 GB of working space for 3 years with the capability to publish and archive up to 1 GB of curated data. If you have an externally awarded current grant that has been processed and categorized as a research grant by Sponsored Programming Services (SPS) you can register it and receive an increase in your project quota to 100 gigabytes for the duration of the grant and increase your publication allocation to 10 gigabytes of curated data. When you receive your grant award, click on "Register Your Grant" to increase your allocation.

You can also extend the life of a project and purchase additional storage with grant or departmental funds upon request. See the Storage and Pricing section of the PURR site for full details.

The owners of a project will receive an email notification one month before their project is scheduled to expire. Published datasets will remain online for at least 10 years, after which time the datasets will be selected or deselected for permanent inclusion in the Libraries' collections.

If you have an externally awarded current grant that has been processed and categorized as a research grant by Sponsored Programming Services (SPS) you can register it and receive an increase in your project quota to 100 gigabytes for the duration of the grant and increase your publication allocation to 10 gigabytes of curated data. When you receive your grant award, click on "Register Your Grant" to increase your allocation.

You can also extend the life of a project and purchase additional storage with grant or departmental funds upon request. See the Storage and Pricing section of the PURR site for full details.

The owners of a project will receive an email notification one month before their project is scheduled to expire. Published datasets will remain online for at least 10 years, after which time the datasets will be selected or deselected for permanent inclusion in the Libraries' collections.

Types of data:

Excerpted from: https://purr.purdue.edu/legal/file-format-recommendations

PURR accepts all file formats and will provide preservation support for as many formats as possible. Upon the publication of a dataset, PURR assigns one of three levels which represent the level of support:

Sustainable: PURR recognizes and fully supports these file formats because they have the highest probability of long-term stability. Sustainable formats are openly documented, supported by a wide range of software platforms, widely adopted, have no data compression (or lossless data compression), and are widely accepted within the archival community. Sustainable file formats will receive Full Preservation

Supported: PURR recognizes these file formats. These formats not meet the minimum requirements for a Sustainable ranking but come close and for practical reasons, may be necessary for long-term care. These formats are more likely to require migration in order to remain viewable. Supported formats are proprietary, widely adopted, publically and commercially important, have lossy data compression, or may be a format which has been depreciated in favor of a newer version. Supported file formats will receive Limited Preservation.

Unsustainable: PURR cannot recognize these formats. Unsustainable formats are not viable for long term storage or accessibility. These formats are proprietary, have little publically documented information, are not widely adopted, have lossy data compression, and are only supported by a single or very few software platforms. Unsustainable file formats will receive Bit-level Preservation.