**Directorate for Mathematical and Physical Sciences – Division of Physics (NSF-MPS/PHY)**

**Types of Data Produced**

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

Give a short description of the data, including amount (if known) and content. If the project will be collecting data of a sensitive nature, note here and reflect upon it in subsequent sections. Data types could include text, spreadsheets, images, 3D models, software, audio files, video files, reports, surveys, patient records, etc. *Consider these questions*:

* What data will be generated in the research?
* What data types will you be creating or capturing?
* How will you capture or create the data?
* If you will be using existing data, state that fact and include where you got it.
* What is the relationship between the data you are collecting and the existing data?

**Data and Metadata Standards**

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).

Describe the format of your data and how it will be "documented." Think about what details (metadata) someone else would need to be able to use these files. For example, you may need a "readme file" to explain variables, structure of the files, etc. *Consider these questions*:

* Which file formats will you use for your data, and why?
* What form will the metadata describing/documenting your data take?
* How will you create or capture these details?
* Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)
* What contextual details (metadata) are needed to make the data you capture or collect meaningful?

**Policies for Access and Sharing, and Provisions for Appropriate Protection/Privacy**

Policies for access and sharing; Provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements.

* This section is very important. The main reason a Data Management Plan is required, is for you to think about how you prepare (manage) your data for sharing and describe how you will actively share your data with non-group members after the project is completed. You should explain how and when the data will become available. Will data be accessible on a web page, by email request, via open-access repository etc.? If there is an embargo period for sharing the data, make sure you provide details explaining this delay (e.g. publisher, political, commercial, patent reasons). And if the data is of a sensitive nature - human subject concerns, potential patentability, species/ecological endangerment concerns - that public access is inappropriate, address here the means by which granular control and access will be achieved (e.g. formal consent agreements; anonymiztion of data; restricted access, only available within a secure network). *Consider these questions*:
	+ How will you make the data available? (Include resources needed to make the data available: equipment, systems, expertise, etc.)
	+ When will you make the data available?
	+ What is the process for gaining access to the data?
	+ Will access be chargeable?
	+ How long will the original data collector/creator/principal investigator retain the right to use the data before making them available for wider distribution?
	+ Are there any embargo periods for political/commercial/patent reasons? If so, give details.
	+ Are there ethical and privacy issues? If so, how will these be resolved?
	+ What have you done to comply with your obligations in your IRB Protocol?
	+ Who will hold the intellectual property rights to the data and how might this affect data access?

**Policies and Provisions for Re-Use, Re-Distribution**

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

* Explain how the policies you outline above can be applied to the re-use and re-distribution of your data. In other words, you need to identify who will be allowed to use your data, how they will be allowed to use your data and whether or not they will be allowed to disseminate your data. If you are planning on restricting access, use or dissemination of the data, you must explain in this section how you will codify and communicate these restrictions. *Consider these questions*:
	+ Will any permission restrictions need to be placed on the data?
	+ Which bodies/groups are likely to be interested in the data?
	+ What and who are the intended or foreseeable uses / users of the data?

**Plans for Archiving and Preservation**

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

This portion of the Data Management Plan asks the researcher to provide a long-term strategy for archiving and preserving the data from the research described in the proposal. *Consider these questions*:

* What is the long-term strategy for maintaining, curating and archiving the data?
* Which archive/repository/database have you identified as a place to deposit data?
* What procedures does your intended long-term data storage facility have in place for preservation and backup?
* How long will/should data be kept beyond the life of the project?

Also consider these questions about the data and associated information that will be deposited:

* What data will be preserved for the long-term?
* What transformations will be necessary to prepare data for preservation / data sharing?
* What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?
* What related information will be deposited?