**Directorate for Geosciences – Division of Earth Sciences (NSF-GEO/EAR)**

**Types of Data**

Preservation of all data, samples, physical collections and other supporting materials needed for long-term earth science research and education is required of all EAR-supported researchers.

* Give a short description of the data, including amount (if known) and format. 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, video files, reports, etc. – make sure to include file formats and any analysis software (scripts) you may use. *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?
	+ Which file formats will you use for your data, and why?
	+ 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**

Data archives must include easily accessible information about data holdings, including quality assessments, supporting ancillary information, and guidance and aids for locating and obtaining data.

* Depending on what data (or format of data) you will be producing, you may need to provide information (metadata) to make your data useable. 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*:
	+ 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?

**Polices for Access and Sharing**

It is the responsibility of researchers and organizations to make results, data, derived data products, and collections available to the research community in a timely manner and at a reasonable cost. Data may be made available for secondary use through submission to a national data center, publication in a widely available scientific journal, book or website, through the institutional archives that are standard for a particular discipline (e.g. IRIS for seismological data, UNAVCO for GP data), or through other EAR-specified repositories. See the Division of Earth Sciences (EAR) Guidance for specific requirements regarding where to share EAR data and costs.

* It is very important, the reason a DMP is required, that you specify how you will share your data with non-group members after the project is completed. EAR encourages researchers to deposit data into open access repositories – if you won’t be using a community-specific repository, explain how you will share your data. If you are using an EAR-specified database, make sure you describe it in this section (how it helps you share data), as well as in the last section (how it helps you store data). *Consider these questions*:
	+ How will you make the data available? (Include resources needed to make the data available: equipment, systems, expertise, etc.)
	+ What is the process for gaining access to the data?
	+ Will access be chargeable?

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

For those programs in which selected principle investigators have initial periods of exclusive data use, data should be made openly available as soon as possible, but no later than two (2) years after the data were collected. This period may be extended under exceptional circumstances, but only by agreement between the Principal Investigator and the National Science Foundation. For continuing observations or for long-term (multi-year) projects, data are to be made public annually.

* Describe policies surrounding the re-use of your data – the EAR division is specifically interested in how soon you will make your data available. If you will not be making the data available for re-use immediately, explain why. Remember that EAR specifies that you must make your data available no later than two years after your research is complete. If there are other policy issues regarding data access and re-use (ethical or privacy issues, for instance) elaborate on them here. *Consider these questions*:
	+ When will you make the data available?
	+ Does the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use?
	+ Will any permission restrictions need to be placed on the data?
	+ 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?
	+ What and who are the intended or foreseeable uses / users of the data?

**Plans for Archiving and Preservation of Access**

Remember - Data may be made available for secondary use through submission to a national data center, publication in a widely available scientific journal, book or website, through the institutional archives that are standard for a particular discipline (e.g. IRIS for seismological data, UNAVCO for GP data), or through other EAR-specified repositories.

* Use this section to describe your long-term strategy for archiving and preserving your data. EAR encourages PIs to submit data to an “EAR specified” repository. Such a repository may function as a long-term archive for your data. *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?