**Data Management Plan**
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

# Using this Data Management Plan (DMP) Template

- Outlined below are questions for each of the five DMP sections presented in NSF's Grant Proposal Guide, Chapter II - Proposal Preparation Instructions, Section j. Special Information and Supplementary Documentation (second bulleted point). Consider and answer these questions as appropriate to develop your data management plan.

- Where possible, text has been adapted from actual data management plans (although not necessarily plans for NSF) under the heading *Sample Text*. Such examples are not available for every section. Consider adapting these text examples for your plan as appropriate.

  **Please Note:** The sample text examples provided below are not officially sanctioned. They are only intended to serve as examples for what you *might* do.

- Create your DMP using the section headings (e.g. Types of Data, Data Description, etc) and your corresponding answers. Remove the questions and modify the answers into meaningful paragraphs.

From the NSF's Grant Proposal Guide: "Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan." This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see AAG Chapter VI.D.4) and may include the following.

# I. Types of Data

## Data Description

1. What is the information output of this grant (e.g. database, dataset, website, software)?

2. What file formats will be used?

3. What is the projected size of the collection (e.g. MB, GB, TB)?

   *Sample text*: Over the course of the project, data will be collected and entered into two relational databases.

   *Sample text*: Over the course of the project, data will be generated from sensors and recorded in X format.

**Data Management Plan**
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

### Existing Data

1. Is there existing data relevant to the project?

2. Will the data be integrated? How?

3. What other information (e.g. computer code, analysis, procedures) need to be shared with the data to make it usable?

# II.  Standards

### Format

1. In what format will data be generated, maintained, and made available?

2. Why are these formats appropriate?

3. Are the data in unusual format?

4. How will they be converted or made accessible for future users?

5. Will multiple file formats be generated?  If so, how will related files in different formats be linked?

   *Sample text*: The associated data types will be captured using X survey software and analyzed using X data analytics tools.

   *Sample text*: Research data will be stored using X file formats. Related files in different formats will be linked by file naming conventions, e.g., ...

### Metadata

1. What standards are in place in your discipline for describing data?

   *Sample text*: Metadata will be generated to describe the data produced in X format and will be stored alongside the data.  X metadata standards will be applied during the creation of the metadata.

   *Sample text*: Discovery metadata for all datasets will be included and recorded as "headers" in each Excel spreadsheet used to store the acquired data.

**Data Management Plan**
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

### Data Organization

1. What best practices are out there in your field regarding data collection and organization?

2. How will you implement data management during the active phase of your grant?

3. Have you consulted with a librarian or archivist on organizing your data and files for the project?

4. How will you manage transfers and synchronization of data between different machines?

5. How will you keep track of the different versions of your data files and documents?

   *Sample text*: Data will conform to best practices and standards from the X community.

   *Sample text*: An on-line Field Catalog will be functional during the field project to support real-time planning. The real-time Field Catalog will contain data of three basic types: operational data, images from experimental real-time numerical weather prediction models, and field reports.

### Quality Assurance

1. What are the procedures for ensuring data quality?

   *Sample text*: Internal calibration (for geophysical data), instrument calibrations, duplicate samples and field blanks (for hydro-chemical data) will be recorded and tested against collected/recorded data to ensure their validity. Qualitative descriptions (lithological data) will be validated through comparative descriptions of collected materials.

### Responsibility

1. Will responsibilities be assigned to someone who is generating or managing the data?

   *Sample text*: Day-to-day quality assessment will be the responsibility of the Lab Director who in turn is overseen by the Project Director.

# III.  Access and Sharing (Including Protected Data)

### Storage and Backup

**Data Management Plan**
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

1. What are the procedures for data storage and backup?

   *Sample text*: Active research materials will be stored on a secure file server that is backed up nightly. Sensitive data will be stored on a machine with whole disk encryption. At present, research data are regularly backed up on a daily basis to a separate password-protected secure server.

## Data Access

1. In X amount of time after the grant ends, where will you make this data available for researchers? (e.g. working with a repository to set a schedule for opening access to data, or having a timeline in the project to prepare data for re-use)

2. How will data be made accessible?

3. Are data posted before or after formal publication?

4. Who will have access?

5. What resources/capabilities are necessary to meet requests for data?

   *Sample text*: Data will be posted on a website within three months of the grant closing. Data will be contributed to X public database. Data will be submitted to supplementary materials sections of peer-reviewed journals.

   *Sample text*: After the 16-month initial data-analysis period, all VORTEX2 data will be considered public domain.

## Ethics and Privacy

1. How will you protect the privacy of individually-identifiable information? (e.g. de-identifying datasets, etc.)

2. How will informed consent be handled?

Data Management Plan
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

### Proprietary Data

> 1. Are there contracts, agreements, or other obligations that will prevent or delay the sharing of data?

### Intellectual Property

> 1. Who holds IP rights to the data? How will this be protected?
>
> *Sample text*: The main output from this project is field data. We recognize that these data are the property of X and hence we will be asking their permission to license these data to Y for use in their exploration program.
>
> *Sample text*: X and third party copyright will be protected. The PI will be responsible for ensuring that all project members are aware as to the ownership of data and who may access them and under what conditions. On-line access to the data will be password protected.
>
> *Sample text*: Users of field data should acknowledge and/or offer co-authorship to the investigators who collected the data.

### Legal Requirements

> 1. What are the relevant legal requirements for data management and sharing?
>
> *Sample text*: We will comply with all applicable HIPAA and FSMA regulations.

## IV. Re-use

### Access and Sharing

> 1. Who will be overseeing data distribution? You or your project team? A consortium? The federal government?
>
> *Sample text*: Data will be available and cited in publication. Researchers will be able to contact the PI for access to data. Data will be maintained in an open XML format to enable open re-use of the data.

**Data Management Plan**
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

**Re-Use**

1. Are there disclaimers, conditions for use, or copyright issues?

2. How will these be noted?

   *Sample text*: There is an agreement regarding the right of the original data collector, creator, or PI for first use of the data. The specified embargo period associated with the data being submitting extends from [ ] until [ ]. The embargo will be lifted by [ ].

# V. Archiving Data

### Archiving and Preservation

1. Have you identified a discipline-specific archive or data repository where similar data is housed? Is there a succession plan in place should the archival entity go out of existence?

   *Sample text*: We plan to use the Dryad public repository for the long-term preservation and dissemination of data underlying publications from this funded research project. Data submitted to Dryad is made publicly available upon online publication** of the associated article. All data in Dryad is released to the public domain without legal restrictions on reuse, through a Creative Commons Zero waiver. There is a (legally non-binding) expectation of attribution of the Dryad data record and associated article. A one-time data deposit charge is paid by the authors or the associated journals, which allows Dryad data to be available for download without cost to users. *[**Researchers may instead choose to stipulate an embargo period of 1 year.]*

### Disaster Preparedness

1. Will hard copies be protected from fire or water damage?

2. Is there a plan to transfer digital information to new media or devices as standards or practices change?

3. Will there be a well-organized index for the data?

4. Will responsibility change from one institution or entity to another? How will this be managed?

### Budget

**Data Management Plan**
Generic Template (with example text)
Adapted from a template from the UNC-CH Libraries (rev. 5/16/2011)

1. What are the costs of preserving the data? How will these costs be paid?

**Selection and Retention**

1. How will data be selected for long-term preservation?

2. How long will data be held and who will be responsible for it once it is inactive?

3. How will you dispose of data that is not needed over the long-term?

   *Sample text*: Project data and metadata will be stored in a distributed system with two remote locations. Our deposit agreement with the Carolina Digital Repository (CDR) specifies preservation of the original format. When archivists there judge the original format to be approaching obsolescence, they will transfer the data set to ICPSR or another appropriate repository that plans for the long-term migration of format if the CDR has not developed that ability by that time.