# Introduction

The National Science Foundation has released a new requirement for full proposal submissions regarding the management of data generated using NSF support. Starting in January, 2011, all proposals must include a data management plan (DMP). This document is a guide to writing your DMP as specified in the NSF - SBE Directorate (available here: <http://www.nsf.gov/sbe/SBE_DataMgmtPlanPolicy.pdf>). The guide has four sections:

[**Section 1**](https://docs.google.com/document/d/1zCpMpPOPg97TyLS2_eBOihTmFW0l-A8pai2yFqtStIg/edit#heading=h.tspve2a2e6zm) is a template and consideration points for completing a data management plan specifically for the social, behavior, and economic sciences.

[**Section 2**](https://docs.google.com/document/d/1zCpMpPOPg97TyLS2_eBOihTmFW0l-A8pai2yFqtStIg/edit#heading=h.z7ucugx7vvoq)is boilerplate language for your use when composing your DMP.

[**Section 3**](https://docs.google.com/document/d/1zCpMpPOPg97TyLS2_eBOihTmFW0l-A8pai2yFqtStIg/edit#heading=h.xqufr89jqhyi)is a list of additional resources (from BU and other institutions) that you may find helpful in creating your DMP.

[**Section 4**](https://docs.google.com/document/d/1zCpMpPOPg97TyLS2_eBOihTmFW0l-A8pai2yFqtStIg/edit#heading=h.dfvvaqqgzpqw)explains the BU products and services available to you.

# Section 1: Data Management Plan Template

[1. Roles and Responsibilities](#h.b4dpke1wsmwn)

[2. Expected Data](#h.yn6xvwku8mm2)

[3. Period of Data Retention](#h.cwo6tnwdcavc)

[4. Data Format and Dissemination](#h.cug54dbn2vhb)

[5. Data Storage and Preservation of Access](#h.1jlu2c3f6vz6)

[6. Additional Possible Data Management Requirements](#h.7fsbwdfotd91)

## 1. Roles and Responsibilities

The Data Management Plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data. It must also consider changes to roles and responsibilities that will occur should a principal investigator or co-PI leave the institution.

Explain how the responsibilities regarding the management of your data will be delegated. This should include time allocations, project management of technical aspects, training requirements, and contributions of non-project staff - individuals should be named where possible. Remember that those responsible for long-term decisions about your data will likely be the custodians of the repository/archive you choose to store your data. While the costs associated with your research (and the results of your research) must be specified in the Budget Justification portion of the proposal, you may want to reiterate who will be responsible for funding the management of your data. Consider the following:

* Outline the staff/organizational roles and responsibilities for implementing this data management plan.
* Who will be responsible for data management and for monitoring the data management plan?
* How will adherence to this data management plan be checked or demonstrated?
* What process is in place for transferring responsibility for the data?
* Who will have responsibility over time for decisions about the data once the original personnel are no longer available?

## 2. Expected Data

The Data Management Plan should describe the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.

Give a short description of what *data* will mean the context of your research project. Explain what types of data you plan to generate, including size, file formats, and number of files. Briefly describe your methods for collecting data. Consider the following:

* What data will be generated in the course of your 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, how will you acquire it?
* What is the relationship between the data you are collecting and the existing data?
* What data will be preserved and shared?

## 3. Period of Data Retention

SBE is committed to timely and rapid data distribution. However, it recognizes that types of data can vary widely and that acceptable norms also vary by scientific discipline. It is strongly committed, however, to the underlying principle of timely access, and applicants should address how this will be met in their DMP statement.

Describe how long you plan to retain the data produced or used in your research. If you plan to embargo the data for a period of time after the research is completed, describe why this is necessary. Consider the following:

* How long will the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use?
* Explain details of any embargo periods for political/commercial/patent or publisher reasons.

## 4. Data Format and Dissemination

The Data Management Plan should describe data formats, media, and dissemination approaches that will be used to make data and metadata available to others. Policies for public access and sharing should be described, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Research centers and major partnerships with industry or other user communities must also address how data are to be shared and managed with partners, center members, and other major stakeholders.

Describe the format of your data. Ideally, data formats will be chosen that are openly and freely available, and/or non-proprietary in nature. Consider the following:

* Which file formats will you use for your data, and why?
* What transformations (to more shareable formats) will be necessary to prepare data for preservation and data sharing?
* What form will the metadata 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?

Describe how you will ensure dissemination of your data. Consider the following:

* How and when will you make the data available? (Include the resources needed to make the data available: equipment, systems, expertise, etc.)
* What other types of information should be shared regarding the data, e.g. the way it was generated, analytical and procedural, information?
* What is the process for gaining access to the data?
* Will any permission restrictions need to be placed on the data?
* How will you manage data with sensitive information?
* 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?

## 5. Data Storage and Preservation of Access

The Data Management Plan should describe physical and cyber resources and facilities that will be used for the effective preservation and storage of research data. These can include third party facilities and repositories.

Describe your long-term strategy for storing, archiving and preserving the data you will generate or use. Consider the following:

* 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?
* What data will be preserved for the long-term?
* On what basis will data be selected for long-term preservation?
* 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?

## 6. Additional Possible Data Management Requirements

More stringent data management requirements may be specified in particular NSF solicitations or result from local policies and best practices at the PI’s home institution. Additional requirements will be specified in the program solicitation and award conditions. Principal Investigators to be supported by such programs must discuss how they will meet these additional requirements in their Data Management Plans.

Describe here any additional program-specific data management requirements. If none exist you may leave this section blank.

# Section 2. Boilerplate Language

## Introduction

This plan describes the management, dissemination, retention, and archiving of the research data produced during the proposed project. The staff of [INSERT YOUR DEPARTMENT OR LAB NAME HERE], with the assistance of [BU Libraries Research Data Management] and [BU Information Services & Technology/Open@BU/Other], will provide for sustainable discovery, access to, and preservation of these data for use by other researchers, instructors, and interested members of the public for the length of this project and at least three years beyond. This will be facilitated through data and publication deposits in existing open-access disciplinary and/or institutional repositories.

## Data Formats and Description

We will utilize the [Metadata scheme] metadata scheme to capture information about the data collected during the course of our research. We will work with a metadata expert from [BU Libraries] to create a working template that captures each dataset’s metadata throughout the research process. Upon completion, we will export this data to [format], which conforms to the data submission requirements of [any discipline-specific depositories or other relevant sources].

## Access, Re-Use, Re-Distribution, and Derivative Works Policies

[If no sensitive or personally-identifiable information is used, include this:]

All data produced during this research will be available freely to the public; we anticipate no sensitive or confidential data. Under the terms of the [Open Data Commons Public Domain Dedication and License (PDDL; http://www.opendatacommons.org/licenses/pddl/1-0/)], users may share, create, and/or adapt these data/databases.

[If sensitive or personally-identifiable information is used, include this instead:]

Results, data, and collections will be made available to other researchers in a timely basis with [EXAMPLE] limitations. Sensitive and confidential data collected will be treated following [HIPAA/IRB] regulations, and an added layer of security will be implemented using [strategies such as data encryption, restricted access, or the separation of identifiable data]. Under the terms of the [Open Data Commons Public Domain Dedication and License (PDDL; http://www.opendatacommons.org/licenses/pddl/1-0/)], users may share, create, and/or adapt data/databases made freely available.

## Data Archiving and Preservation

To increase access to the published research that has been funded, the research collaborators will deposit peer-reviewed or pre-print manuscripts (with linked supporting data where possible) in the [Open@BU] institutional repository. Other works, including presentations and white papers, will also be made accessible via the [Open@BU] institutional repository.

Digital data will be stored using the [Boston University storage archive/3rd party]. Data stored on the SDA that will be made freely available will be archived in the [Open@BU] repository, which will provide a user-friendly interface for the organization, context, and discoverability of data. This combination of Open@BU] and the SDA provides mirroring, redundancy, media migration, access control, file integrity validation, embargoes, and other security-based services that ensure the data are appropriately archived for the life of the project and beyond the project if necessary.

# Section 3: Additional Resources

General Resources on Data Management Plans

* [NSF Data Management FAQ](http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp)
* [NSF Grant Proposal Guide on DMPs](http://www.nsf.gov/pubs/policydocs/pappguide/nsf15001/gpg_2.jsp#dmp)
* [BU Data Management Plans and Resources](http://www.bu.edu/datamanagement/outline/)
* [DataONE Best Practice: Revisit Data Management Plan Throughout the Project Life Cycle](http://www.dataone.org/best-practices/revisit-data-management-plan-throughout-project-life-cycle)

Defining Data and Metadata

* [What is data?](http://www.bu.edu/datamanagement/background/whatisdata/)
* [What are metadata?](http://www.bu.edu/datamanagement/outline/elements/metadata/)
* [DataONE Best Practice: Define Expected Data Outcomes and Types](http://www.dataone.org/best-practices/define-expected-data-outcomes-and-types)
* [DataONE Best Practice: Identify and Use Relevant Metadata Standards](http://www.dataone.org/best-practices/identify-and-use-relevant-metadata-standards)
* [DataONE Best Practice: Document and Store Data Using Stable File Formats](http://www.dataone.org/best-practices/document-and-store-data-using-stable-file-formats)
* [DataONE Best Practice: Define the Data Model](http://www.dataone.org/best-practices/define-data-model)
* [DataONE Best Practice: Identify Data with Long-Term Value](http://www.dataone.org/best-practices/identify-data-long-term-value)
* [DataONE Best Practice: Define Expected Data Outcomes and Types](http://www.dataone.org/best-practices/define-expected-data-outcomes-and-types)

Roles and Responsibilities, Effective Data Management

* [DataONE Best Practice: Define Roles and Assign Responsibilities for Data Management](http://www.dataone.org/best-practices/define-roles-and-assign-responsibilities-data-management)
* [ICPSR – Effective Data Management](http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/index.html)
* [University of Edinburgh “How to Manage Research Data”](http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/research-data-mgmt)

Storage

* [Storage of Ongoing Data](http://www.bu.edu/datamanagement/outline/elements/storage/)
* [Document and Store Data Using Stable File Formats](http://www.dataone.org/best-practices/document-and-store-data-using-stable-file-formats)
* [BU IS&T Storage and Archiving Services](http://www.bu.edu/datamanagement/resources/ist-services/)
* [External Archiving Services and Data Repositories](http://www.bu.edu/datamanagement/resources/archiving-services/)
* [re3data: Registry of data repositories](http://service.re3data.org/search/results?term=)
* [Open Access Directory - Disciplinary repositories](http://oad.simmons.edu/oadwiki/Disciplinary_repositories)
* [Open Access Directory - Data repositories](http://oad.simmons.edu/oadwiki/Data_repositories)
* [DataONE Best Practice: Identify Suitable Repositories for your Data](http://www.dataone.org/best-practices/identify-suitable-repositories-data)

Dissemination, Access, and Derivative Works

* [NSF - Dissemination of results](http://www.nsf.gov/bfa/dias/policy/dmp.jsp)
* [DataONE: Sharing Data. Legal and policy considerations](https://www.dataone.org/best-practices/sharing-data-legal-and-policy-considerations)

[BU: Access to Data](http://www.bu.edu/datamanagement/outline/elements/access/)

# Section 4: BU Resources

* [IS&T Storage and Archiving Services](http://www.bu.edu/datamanagement/resources/ist-services/)
* [External Archiving Services and Data Repositories](http://www.bu.edu/datamanagement/resources/archiving-services/)
* [OpenBU](http://open.bu.edu/xmlui/)