Keene State College
Office of Sponsored Projects and Research

Guidance on Creation of Data Management Plans

This document compiles helpful guidance from many different sources, all designed to help you write effective Data Management Plans for your research projects. Attributions for sources are provided by linking to the original source, and you are encouraged to explore these resources more fully by visiting those websites. We gratefully acknowledge the work of our professional colleagues.

What is a data management plan and why do I need one?

Simply put, a data management plan (DMP) is a clear description of the kinds of data you expect to collect through your studies, a description of how you will make that data available to others (and under what conditions), and how you will archive that data to safeguard it for future users.

Why do you need one? One could argue that thinking about the elements of a DMP is simply part of being a good scholar, researcher and/or scientist. Data should be freely available to the research community as a whole in the hopes that such open sharing will lead to significant new knowledge. On the other hand, a very practical, operational answer is that an increasing number of funding agencies are requiring you to file a DMP in one form or another as a required part of your application. For example, NIH has long expected researchers to document how the will disseminate their data and make it available to others as an integral part of a proposal. More recently, NSF has required all proposals to include a two-page DMP as a required supplementary document in a proposal.

From: http://www.sewanee.edu/dspace/dmpg

What is typically included in a DMP?

This depends on the conventions of your specific discipline, as well as requirements of the agency funding your research (if applicable). The generic guidelines put forth by the Inter-university Consortium for Political and Social Research (ICPSR) provide a good starting point for an effective DMP, and suggest the investigators consider including the following key elements, depending on the nature of your research/data.

- Data Description
- Access and Sharing
- Metadata
- Intellectual Property
- Ethics and Privacy
- Data Format
- Archiving and Preservation
- Storage and Backup

- Security
- Responsibility
- Existing Data
- Selection and Retention Periods
- Audience

For more detail, including succinct sample language for each element, see:
http://www.icpsr.umich.edu/icpsrweb/content/ICPSR/dmp/framework.html

For agency-specific guidance, see later sections specific to NSF and NIH.

Can Keene State College Assist with Archiving my Data?

The Mason Library manages the Keene State College Institutional Repository known as KDig, which offers one option for archiving your research data. KDig is a digital resource tool used to collect and archive the electronic records of enduring value produced by the academic and administrative departments of the College. KDig also provides a means to accumulate and share the research, scholarly, and artistic output of the Keene State academic community.

This digital repository can accept multiple file types, including spreadsheets, tab-delimited formats, XML format, jpeg, and others. It is Dublin Core and Open Archives Initiative (OAI) compliant. Moreover, it is flexible enough to accommodate most discipline-based meta-data standards, and can be password protected or open (unprotected) depending on your needs. Once your materials are deposited in the archive, you can choose to have a landing page for your archived materials on your departmental or other webpage. Please note that all confidential information would need to be stripped from the data by the investigator prior to deposit.

Individuals are encouraged to contact Kara Young, Head of Library Systems (kyoung1@keene.edu) to discuss their archival needs and determine if KDig is an appropriate data management solution for your project.

http://www.keene.edu/library/kdig.cfm

What if KDig does not meet my archival needs?

A range of other options exist, though they may involve a cost. Investigators should carefully consider the expected costs of utilizing an outside archival service and include those costs in your proposed grant budget (if applicable). Digital data repositories often specialize in certain types of data or disciplines, so you will want to have a conversation with any service you are considering to make certain it will meet your needs and to receive accurate pricing information for the services you will need.

To begin searching for a repository, you might use the following websites:

> Simmons University listing of data repositories by discipline:
> http://oad.simmons.edu/oadwiki/Data_repositories

University of North Carolina listing of data repositories by discipline:
http://www.lib.unc.edu/reference/data_services/researchdatatoolkit/repositories.html

OpenDOAR, a directory of open access academic repositories:
http://www.opendoar.org/index.html

Additionally, a good place to start is by checking with your discipline-specific professional organization to determine their standards for archiving of data, and possible repositories they might suggest as appropriate for your data.

Where can I find sample data management plans for various disciplines?

Several institutions have samples or templates posted on their websites:

Rice University: http://osr.rice.edu/forms/dataManagementPlans.pdf

University of Virginia:
http://www.google.com/url?sa=t&source=web&cd=8&sqi=2&ved=0CF4QFjAH&url=http%3A%2F%2Fwww2.lib.virginia.edu%2Fbrown%2Fdata%2FNSF_Data_Management_Plan_Template.doc&rct=j&q=sample%20data%20management%20plan%20nsf&ei=bXawTfnGHIjDgQf7y7X3Cw&usg=AFQjCNFNhbm0RnwPiaH1fD5ds5g9UEee0w&sig2=7VtadiyOJDGfGoCj1jOGyw

University of Michigan: http://www.lib.umich.edu/research-data-management-and-publishing-support/nsf-data-management-plans#examples

What Should I know about the National Science Foundation's Expectations?

NSF Data Sharing Policy
Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. See Award & Administration Guide (AAG) Chapter VI.D.4.

NSF Data Management Plan Requirements
Beginning January 18, 2011, proposals submitted to NSF must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results, and may include:

- the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;
- the 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);
- policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
- policies and provisions for re-use, re-distribution, and the production of derivatives; and
- plans for archiving data, samples, and other research products, and for preservation of access to them.

Directorate-Specific Guidance Available
Data management requirements and plans specific to the Directorate, Office, Division, Program, or other NSF unit, relevant to a proposal are available at: http://www.nsf.gov/bfa/dias/policy/dmp.jsp. If guidance specific to the program is not available, then the requirements established in this section apply.

DMPs for Collaborative Proposals
Simultaneously submitted collaborative proposals and proposals that include subawards are a single unified project and should include only one supplemental combined Data Management Plan, regardless of the number of non-lead collaborative proposals or subawards included. Fastlane will not permit submission of a proposal that is missing a Data Management Plan.

See Grant Proposal Guide (GPG) Chapter II.C.2.j for full NSF policy implementation.

What should I know about the National Institutes of Health's Expectations?

NIH has had a data sharing policy since October 2003:

"Investigators submitting an NIH application **seeking $500,000 or more in direct costs in any single year** are expected to include a plan for data sharing or state why data sharing is not possible."

From NIH's Data Sharing Policy and Implementation Guide:
The precise content of the data-sharing plan will vary, depending on the data being collected and how the investigator is planning to share the data. Applicants who are planning to share data may wish to describe briefly:

- the expected schedule for data sharing,
- the format of the final dataset,
- the documentation to be provided,
- whether or not any analytic tools also will be provided,
- whether or not a data-sharing agreement will be required and, if so, a brief description of such an agreement (including the criteria for deciding who can receive the data and whether or not any conditions will be placed on their use),
- and the mode of data sharing (e.g., under their own auspices by mailing a disk or posting data on their institutional or personal website, through a data archive or enclave).

For more information:
http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm