Guide to Fulfilling NSF's Data Management Plan (DMP) Requirements

Murday@usc.edu

After January 18, 2011 a proposal to NSF must have a data management plan (DMP). FastLane will not permit submission without such a plan. Proposals must include sufficient information that peer reviewers can assess both DMPs and past performance. Strategies and eventual compliance with this policy will be evaluated not only by proposal peer review but also through project monitoring by NSF program officers, by division and directorate Committees of Visitors, and by the National Science Board.

As an illustration of the new requirement, the following paragraph was taken from a new NSF solicitation (proposal due date of 30 Jan 2011):

"Data Management Plan: The Proposal and Award Policies and Procedures Guide (PAPPG) contains a clarification of NSF's long standing data policy. All proposals must describe plans for data management and sharing of the products of research, or assert the absence of the need for such plans. FastLane will not permit submission of a proposal that is missing a Data Management Plan. The Data Management Plan will be reviewed as part of the intellectual merit or broader impacts of the proposal, or both, as appropriate. Links to data management requirements and plans relevant to specific Directorates, Offices, Divisions, Programs, or other NSF units are available on the NSF website at:

http://www.nsf.gov/bfa/dias/policy/dmp.jsp.

See Chapter II.C.2.j of the Grant Proposal Guide (GPG) for further information about the implementation of this requirement."

NSF provides answers to frequently asked questions at:
http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp

Points to Address in a Data Management Plan:

Investigators are encouraged to create plans that make their data easily accessible to the research community (typically via the web), in a format that makes the data readily useable and easy to interpret. Please write this section by thinking of the researchers who could gain access to your data, explaining how you will help the entire research community. Investigators are encouraged to cover the following points within their two page plans.

1. Products of the Research: types of data, samples, physical collections, software, curriculum materials, and other materials to be produced:

   a) Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project.
   b) Expected types of data to be retained, including
      - Analyzed data (e.g., digital information that would be published, including digital images, published tables, and tables of the numbers used for making published graphs), and
      - Metadata that define how these data were generated (e.g., data that should be published in theses, dissertations, refereed journal articles, supplemental data attachments for manuscripts, books and book chapters, and other print or electronic publication formats).

   Text for this section is specific to the topics being investigated in the proposal.

2. Standards to be used for data and metadata format and content:

   a) Data formats that will be used to make data available to others, including any metadata. Formats may include hardcopy notebook and/or instrument outputs, ASCII, HTML, JPEG or other formats).
   b) Where data are stored in unusual or not generally accessible formats, explain how the data may be converted to a more accessible format or otherwise made available to interested parties. Note any anticipated inclusion of your data into databases that mine the published literature (e.g., PubChem, NIST Chemistry WebBook).

   Text for this section is specific to the topics being investigated in the proposal.

3. Access to data and data sharing practices and policies:

   a) "Access to data" refers to data made accessible without explicit request from the interested party, for example, those posted on a web site or made available to a public database. Describe your plans, if any, for providing such general access to data, including websites maintained by your research group, and direct contributions to public databases.
   b) If maintenance of a web site or database is the direct responsibility of your group, provide information about the period of time the web site or database is expected to be maintained; a minimum of 3 years from conclusion of grant or public release is the guidance from the Engineering Directorate, but there can be exceptions. The cognizant NSF program officer should be consulted to ascertain those exceptions. Also describe your practice or policies regarding the release of data for access, for example whether data are posted before or after formal publication.
   c) "Data sharing" refers to the release of data in response to a specific request from an interested party. Describe your policies for data sharing, including where applicable provisions for protection of privacy, confidentiality, intellectual property, national security, or other rights or requirements.

   Proposals may include the following language to partially address item 3.

   All participants in this proposal will conduct research and publish the results of their work. Papers

will be published in a peer-reviewed scientific journal or book that publishes in English, or as a peer-reviewed data report. Beyond the data posted on the USC website, primary data, samples, physical collections and other supporting materials created or gathered in the course of work will be shared with other researchers upon reasonable request, at no more than incremental cost and within a reasonable time of the request or, if later, the filing of a patent application covering the results of such research.

USC policies for the protection of human subjects in research govern research being conducted at USC, with USC facilities, or by USC faculty, staff, or students. USC adheres to the Federal Regulations for the Protection of Human Subjects (HHS OHRP/FDA), as well as the protection of identifiable patient information under the Health Insurance Portability and Accountability Act (HIPAA), and will restrict access accordingly**.
http://www.usc.edu/admin/provost/oprs/policies/hspp.html.**

4. <u>Policies and provisions for re-use, re-distribution, and the production of derivatives:</u>

   a) Describe your policies regarding the use of data provided via general access or sharing. 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. For example, if you plan to provide data and images on your website, will the website contain disclaimers, or conditions regarding the use of the data in other publications or products? Describe these disclaimers and/or terms of use.

   Proposals may include the following language to partially address item 4.

   USC's policy is to encourage, wherever appropriate, research data to be shared with the general public through internet access. This public access will be regulated by the university in order to protect privacy and confidentiality concerns, as well to respect any proprietary or intellectual property rights. Administrators will consult with the university's legal office to address any concerns on a case-by-case basis, if necessary. Terms of use will include requirements of attribution along with disclaimers of liability in connection with any use or distribution of the research data, which may be conditioned under some circumstances

   Proposers should expand on the above with any specifics related to the particular project.

5. <u>Plans for archiving and for preservation of access:</u>

   a) Describe how data will be archived and how preservation of access will be handled. For example, will hardcopy notebooks, instrument outputs, and physical samples be stored in a location where there are safeguards against fire or water damage? Is there a plan to transfer digitized information to new storage media or devices as technological standards or practices change? Will there be an easily accessible index that documents where all archived data are stored and how they can be accessed?
   b) NSF-supported large research facilities and other focused research programs may specify more stringent data-sharing and archiving procedures for research conducted using these facilities or under these programs. Some NSF Divisions require data to be archived in specific sites. For instance inventories (metadata) of all marine environmental data collected should be submitted to the designated National Data Centers within sixty (60) days after the observational period/cruise. For continuing observations, data inventories should be submitted periodically if there is a significant change in location, type or frequency of such observations.

Proposals may include the following language to partially address item 5.

Samples and other research products will be made available immediately after publication. Final peer-reviewed journal manuscripts, and supplemental information such as data tables for graphical information in manuscript figures and for statistically processed averages, which arise from NSF funds, will be posted in the USC Digital Repository and will be available on this publically available website no later than 12 months after publication. Authors will ensure their publishing agreement allows the paper to be posted to the archive; alternatively the USC website will provide a reference to the journal articles coupled with the supplemental information. These records will be durable, accessible through web protocols, and made safe from tampering or falsification. The storage media will be updated as necessary to keep it current.

The USC Digital Repository (USCDR) provides fee-based consulting and services to help USC researchers meet NSF requirements. Services include digitization, cataloging, preservation, archiving, and online access. Consulting is also available to help researchers determine appropriate web hosting and web design solutions. USCDR is a center that is jointly operated by the USC Libraries, the USC Shoah Foundation Institute (SFI), and USC's Information Technology Services (ITS) division. As such, USCDR is able to offer researchers access to the professional expertise and technological infrastructure of the SFI, Libraries, and ITS.

After consultation with the appropriate NSF program officer to ascertain any exceptions, items will be discarded no sooner than 3 years after the conclusion of the grant or the public release, whichever is later. Research data that support patents will be retained for the entire term of the patent.

(Note: Researchers should contact Sam Gustman, associate dean for the USC Libraries and chief technology officer for the USC Shoah Foundation Institute, at sgustman@usc.edu for additional information about services and pricing for the USCDR.)

Appendix A: Resources on Data Management Plans

The URL http://www.nsf.gov/bfa/dias/policy/dmp.jsp leads to the following information:

Dissemination and Sharing of Research Results

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. (page 6 below)

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. See (page 7 below) Grant Proposal Guide (GPG) Chapter II.C.2.j for full policy implementation.

Requirements by Directorate, Office, Division, Program, or other NSF Unit
Links to data management requirements and plans relevant to specific Directorates, Offices, Divisions, Programs, or other NSF units, are provided below. If guidance specific to the program is not provided, then the requirements established in Grant Proposal Guide, Chapter II.C.2.j apply.

Presently posted NSF division requirements (see pages 8-10 for selected illustrations of information requested in these guidances):

- Engineering Directorate (ENG)
    - Directorate-wide Guidance

- Geological Sciences Directorate (GEO)
    - Division of Earth Sciences
    - Integrated Ocean Drilling Program
    - Division of Ocean Sciences

- Mathematical and Physical Sciences Directorate (MPS)
    - Division of Astronomical Sciences
    - Division of Chemistry
    - Division of Materials Research
    - Division of Mathematical Sciences
    - Division of Physics

- Social, Behavioral and Economic Sciences Directorate (SBE)
    - Directorate-wide Guidance

Note that a program solicitation may provide specific guidance on preparation of data management plans, such guidance must also be followed.

Other pertinent Documents
- Council on Governmental Relations, Access to and Retention of Research Data: Rights and Responsibilities, March 2006. http://206.151.87.67/docs/CompleteDRBooklet.htm
- National Science Foundation, Proposal and Award Policies and Procedures Guide, http://www.nsf.gov/publications/pub_summ.jsp?ods_key=papp
- Office of Management and Budget, Circular A-110, September 30, 1999. White House Website, OMB Home. http://www.whitehouse.gov/omb/circulars/a110/a110.html

Appendix B. Award & Administration Guide (AAG) Chapter VI.D.4.

4. Dissemination and Sharing of Research Results

a. Investigators are expected to promptly prepare and submit for publication, with authorship that accurately reflects the contributions of those involved, all significant findings from work conducted under NSF grants. Grantees are expected to permit and encourage such publication by those actually performing that work, unless a grantee intends to publish or disseminate such findings itself.

b. 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. Privileged or confidential information should be released only in a form that protects the privacy of individuals and subjects involved. General adjustments and, where essential, exceptions to this sharing expectation may be specified by the funding NSF Program or Division/Office for a particular field or discipline to safeguard the rights of individuals and subjects, the validity of results, or the integrity of collections or to accommodate the legitimate interest of investigators. A grantee or investigator also may request a particular adjustment or exception from the cognizant NSF Program Officer.

c. Investigators and grantees are encouraged to share software and inventions created under the grant or otherwise make them or their products widely available and usable.

d. NSF normally allows grantees to retain principal legal rights to intellectual property developed under NSF grants to provide incentives for development and dissemination of inventions, software and publications that can enhance their usefulness, accessibility and upkeep. Such incentives do not, however, reduce the responsibility that investigators and organizations have as members of the scientific and engineering community, to make results, data and collections available to other researchers.

e. NSF program management will implement these policies for dissemination and sharing of research results, in ways appropriate to field and circumstances, through the proposal review process; through award negotiations and conditions; and through appropriate support and incentives for data cleanup, documentation, dissemination, storage and the like.

Appendix C: Grant Proposal Guide (GPG) Chapter II.C.2.j

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:

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

2. 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);

3. policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;

4. policies and provisions for re-use, re-distribution, and the production of derivatives; and

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

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.

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. Proposals for supplementary support to an existing award are not required to include a Data Management Plan.

A valid Data Management Plan may include only the statement that no detailed plan is needed, as long as the statement is accompanied by a clear justification. Proposers who feel that the plan cannot fit within the supplement limit of two pages may use part of the 15-page Project Description for additional data management information. Proposers are advised that the Data Management Plan may not be used to circumvent the 15-page Project Description limitation. The Data Management Plan will be reviewed as an integral part of the proposal, coming under Intellectual Merit or Broader Impacts or both, as appropriate for the scientific community of relevance.