NATIONAL SCIENCE FOUNDATION – DATA MANAGEMENT PLAN

Suggested proposal language for NSF's Data Management Plan, which is required in the form of a two-page supplementary document for all proposals submitted, or due, on or after January 18, 2011. The plan will be subject to peer review and will be reviewed under the merit review criteria. This plan should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results, 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.

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.

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 provided, then the requirements established in Grant Proposal Guide, Chapter II.C.2.j apply.

Please see the UC Curation Center (UC3) for general Data Management Plan information and suggestions for organizing, managing, sharing and preserving your data (http://www.cdlib.org/services/uc3/datamanagement/index.html).

▪▪▪

Draft "boilerplate" language and template for NSF-required supplementary Data Management Plan (Note: Individual Data Management Plans should be tailored to the specific research activities described within each proposal)
(Adapted from ICPSR Guidelines for Effective Data Management Plans)

This Data Management Plan addresses NSF's policy on the dissemination and sharing of research results within a reasonable time. In accordance with this policy, this plan does not include preliminary analyses (including raw data), drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. Furthermore, data to enable peer review and publication/dissemination and/or to protect intellectual property may be temporarily withheld from distribution and other proposed data management. This plan will make certain that the data produced during the period of this project is appropriately managed to ensure its usability, access and preservation.

1. **Data description** (Highly Recommended)

Provide a description of the information to be gathered; the nature and scale of the data that will be generated or collected. Describe the characteristics of the data, their relationship to existing data, and any disclosure risks that may apply.

2. **Format** (Highly Recommended)

Specify the formats in which the data will be generated, maintained, and made available, including a justification for the procedural and archival appropriateness of those formats. Preservation formats should be platform-independent and non-poprietary to ensure that they will be usable in the future.

3. **Metadata** (Highly Recommended)

Describe the metadata to be provided along with the generated data, and a discussion of the metadata standards used. Structured or tagged metadata, like the XML format of the Data Documentation Initiative (DDI) standard, are optimal because the XML offers flexibility in display and is also preservation-ready and machine-actionable.

4. **Storage and backup** (Highly Recommended)

Describe the physical and cyber resources and facilities that will be used for the effective preservation and storage of the research data. How and where will you store copies of your research files to ensure their safety? How many copies will you maintain and how will you keep them synchronized?

5. **Intellectual property rights** (Highly Recommended)

Identify the entities or persons who will hold the intellectual property rights to the data, and how IP will be protected if necessary. Any copyright constraints (e.g., copyrighted data collection instruments) should be noted.

6. **Access and sharing** (Highly Recommended)

Describe how data will be shared, including access procedures, embargo periods, technical mechanisms for dissemination and whether access will be open or granted only to specific user groups. A timeframe for data sharing and publishing should also be provided. Possible mechanisms for archiving and sharing include: 1) Domain repository like ICPSR (social science); 2) Self-dissemination through a dedicated Web site that the research team will create and maintain. If this option is chosen, it is recommended that the data producer arrange for eventual archiving of the data after the self-dissemination period terminates and specify the schedule for data sharing in the grant application; 3) Preservation with delayed dissemination. Under such an agreement the data producer makes an arrangement with a public data repository for archival preservation of the data with dissemination to occur at a later date, usually within a year; 4) Institutional repositories. Institutional repositories at academic institutions have the goal of preserving and making available some portion of the academic work of their students, faculty, and staff. Note that not all IRs have the capacity to accept and curate data. Will your data be free of direct and indirect identifiers? If not, how will you share your restricted data? Will special terms of use be required? Indicate when the data will be made available to others.

7. **Archiving and preservation** (Highly Recommended)

Describe the procedures in place or envisioned for long-term archiving and preservation of the data, including succession plans for the data should the expected archiving entity go out of existence.

8. **Ethics and privacy** (Highly Recommended)

If applicable, discuss how informed consent will be handled and how privacy will be protected, including any exceptional arrangements that might be needed to protect participant confidentiality, and other ethical issues that may arise. If applicable, what are your plans to obtain IRB approval? Are there legal constraints (e.g., HIPAA) on sharing data? If applicable, how will you manage disclosure risk in the data to be shared and archived?

9. **Existing data** (Optional)

Provide a survey of existing data relevant to the project and a discussion of whether and how these data will be integrated.

10. **Data organization** (Optional)

Describe how the data will be managed during the project, with information about version control, naming conventions, etc.

11. **Quality assurance** (Optional)

Specify the procedures for ensuring data quality during the project.

12. **Security** (Optional)

Describe the technical and procedural protections for information, including confidential information, and how permissions, restrictions, and embargoes will be enforced.

13. **Responsibility** (Optional)

Names of the individuals responsible for data management in the research project.

14. **Budget** (Optional)

Describe the costs of preparing data and documentation for archiving and how these costs will be paid. Requests for funding may be included.

15. **Legal requirements** (Optional)

Indicate whether there are any relevant federal or funder requirements for data management and data sharing.

16. **Audience** (Optional)

Describe the potential secondary users of the data you will produce.

17. **Selection and retention periods** (Optional)

Describe how data will be selected for archiving, how long the data will be held, and plans for eventual transition or termination of the data collection in the future.