WHAT SHOULD NSF DATA MANAGEMENT PLANS LOOK LIKE

Bin Ye, College of Agricultural and Life Sciences University of Wisconsin



Diane Winter, Inter-university Consortium for Political and Social Research (ICPSR), University of Michigan



OUTLINE

- Key elements of NSF data management plans
- How to prepare a data management plan for your research proposal?
- Data archiving at a domain repository

Key Elements of NSF Data Management Plans

2011 NCURA REGION IV

PRINCIPLES UNDERLYING DATA SHARING

- Reinforces open scientific inquiry.
- Encourages diversity of analysis and opinions.
- Promotes new research and allows for the testing of new or alternative methods. Improves methods of data collection and measurement through the scrutiny of others.
- Reduces costs by avoiding duplicate data collection efforts.
- Provides an important resource for training in research.

NSF DATA MANAGEMENT PLAN POLICY

• 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:

Proposals submitted or due on or after January 18, 2011, 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 Grant Proposal Guide (GPG) Chapter II.C.2.j for full policy implementation.

KEY ELEMENTS OF NSF DATA MANAGEMENT PLANS

- All proposals must include a data management plan in the form of no more than two-page supplementary document.
- A data management plan is needed for a research project that is not anticipated to generate data.
- The data management plan will be subject to peer review.
- Allowable costs of implementing the data management plan can be included in your budget.
- NSF doesn't have particular requirement for archiving and accessibility of data and samples.
- All data needs to be available within a reasonable length of time.

KEY ELEMENTS OF NSF DATA MANAGEMENT PLANS

- No particular requirements for what types of data need to be saved.
- Don't need to share data or samples while the data or samples are being analyzed.
- Need to follow regulations and guidelines for protecting privacy rights (such as HIPAA) if the data are related to individuals' information, locations of endangered species, etc.
- Don't need to share data which have potential commercial value.
- All reporting to NSF must include how you are following the data management plan (this is an implied rule based on language preceding 'rules').

How to prepare a data management plan for your research proposal?

Data Management Plan may include (GPG, Chapter II, C,2,j):

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.

YOUR DATA MANAGEMENT PLAN

- What are your data?
- How are the data organized and saved?
- How to preserve the data?
- How will the data be shared?
- Who is responsible for the Data Management Plan?

TYPES OF DATA

- Data collected during research
- Publications
- Samples
- Physical collections
- Software
- Models

HOW ARE THE DATA ORGANIZED AND SAVED?

- Research data, collected information, software and model
 - Hard disks in a PC
 - Campus network drive
 - Third party database
- Biological samples
 - In freezers at the lab
 - Animal care facility on campus
- Physical collections
 - File cabinets in the lab
 - Climate controlled facility on campus

HOW TO PRESERVE THE DATA?

- A plan to store raw data, information and research data.
- How and when to back up all data.
- Resources to preserved biological samples and other physically collected samples.

HOW WILL THE DATA BE SHARED?

- Biological samples
 - American Type Culture Collection (ATCC)
 - AddGene
- Data and Information are in digital format:
 - Pubmed, BioMed Central, Wikipedia
 - UW-Madison Research Data Service (researchdata.wisc.edu)
 - The Inter-university Consortium for Political and Social Science at the University of Michigan
- A plan to follow guidelines and regulations before confidential or private personal information are disclosed to the public.

OTHER IMPORTANT ELEMENTS

• 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.

• Only a single supplemental combined Data Management Plane is needed for Simultaneously submitted collaborative proposals and proposals that include subawards.

• 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.

• Part of the 15-page Project Description may be used for additional data management information. However, the Data Management Plan may not be used to circumvent the 15-page Project Description limitation.

Data Archiving at a Domain Repository

2011 NCURA REGION IV

GUIDELINES FOR EFFECTIVE DATA MANAGEMENT PLANS

Framework

- List of elements
- Questions to consider
- Examples of elements
- Information on why elements are important
- Suggested additional reading



DATA MANAGEMENT PLAN RESOURCES

- Resources for Development
- Templates and Tools
- Guidance on Funder Requirements
- University Data Management Web Sites
- Good Practice Guidance



NAVIGATING THE ICPSR WEBSITE

Taking a walk in the Cloud (computing that is)



http://www.icpsr.umich.edu/icpsrweb/ICPSR/dmp/index.jsp

WHY DEPOSIT DATA WITH A DOMAIN REPOSITORY?

- Data Curation. Enhances and adds value to data by making them easier to use. We also describe data fully for Web discovery and protect respondent privacy.
- Long-Term Preservation. Ensures long-term data availability.
- Worldwide Dissemination. Offers data in the major statistical package formats and online analysis. Usage statistics are available on request.
- User Support. Staff are available to answer questions about downloading and using data.
- Levels of Access. Offers restricted access data services, and a secure data enclave
- Aggregation of Publications. Creates a database of citations based on analyses of your data.

CONFIDENTIAL DATA – RISKS OF DISCLOSING SUBJECTS' IDENTITIES ARE INCREASING

- Geographically reference data
- Longitudinal data
- Multi-level data
 - Student, teacher, school, school district
 - Patient, clinic, community

CONFIDENTIAL DATA – HOW TO PROTECT

- Remove direct identifiers name, address, identification numbers, etc.
- Evaluate risk of indirect identification linking to public databases
- Apply disclosure risk limitation measures recoding extreme values

RESTRICTING ACCESS

- Restricted-use agreements
- Remote execution
- Virtual data enclave
- On-site data enclave

WHAT WE DO AT ICPSR

- Acquire, curate and archive social science data
- Distribute data to researchers
- Preserve data for future generations
- Link data to publications
- Provide training in quantitative methods
- Resources for using data in teaching

??? QUESTIONS ???

2011 NCURA REGION IV

THANK YOU

Bin Ye

University of Wisconsin College of Agricultural and Life Sciences binye@cals.wisc.edu 608-890-1648

Diane Winter

University of Michigan Inter-university Consortium for Political and Social Research (ICPSR) diwinter@umich.edu 734-615-2315