

Research Data Management Group Survey of NSF Principal Investigators at Cornell University

Developed and administered by: Dianne Dietrich, Eric Chen, Florio Arguillas, Stefan Kramer, and Gail Steinhart

This document contains the survey instrument used in a 2011 survey of NSF PIs at Cornell University.

Block 0

As you may be aware, on October 1, 2010, the National Science Foundation announced a new policy requiring a supplementary document for all grant proposals outlining the proposal's data management plan (see <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>). This requirement will take effect on January 18, 2011. Individual programs and directorates within NSF may have additional guidelines. Other major research funders can be expected to implement similar policies, if they have not already done so.

The Research Data Management Service Group (RDMSG, <http://data.research.cornell.edu/>) is conducting this survey to estimate the demand on campus services for data management, and to identify potential gaps in existing services.

It should take you approximately 10 minutes to complete this survey and your participation is voluntary. You will not be required to provide any identifying information unless you choose to.

Your answers will provide valuable information for use in the RDMSG's planning efforts. Some results from this survey, such as general trends, may be used in external reports, but no identifying information or direct quotes will be used without your consent.

This survey will be closed and no further submissions will be accepted after February 1st, 2011.

Information Sessions

You are also invited to attend an informational session on the National Science Foundation's (NSF) new policy requiring a data management plan with all grant proposals. The new policy goes into effect January 18, 2011.

Staff from RDMSG will review the new requirement, describe how researchers can obtain assistance from the RDMSG to create data management plans, and answer questions.

Three sessions will be offered:

Thursday, January 13, 1:30-2:30pm, G01 Biotech
Tuesday, January 18, 9:00-10:00am, 102 Mann Library
Thursday, January 20, 12:30-1:30pm, 312 Hollister

By clicking the next button below, you voluntarily agree to participate in this online survey.

Block 1

Please answer the following questions with your most recent NSF award in mind.

Please specify the NSF directorate of your most recent award.

- Directorate for Biological Sciences
 - Directorate for Computer & Information Science & Engineering
 - Directorate for Education & Human Resources
 - Directorate for Engineering
 - Directorate for Geosciences
 - Directorate for Mathematical & Physical Sciences
 - Directorate for Social, Behavioral & Economic Sciences
 - Office of the Director (includes Office of Cyberinfrastructure, Polar Programs, and others)
-

Would you be interested in any sort of guidance, including consultation, for writing a data management plan in support of an NSF grant application?

- Yes
- No
- I'm not sure

Additional comments

Block 2

According to the NSF, a data management plan may include a description of "the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project."

Please specify the types of data you have produced or anticipate producing for this project that you intend to share with others. Check all that apply.

- Text
- Image
- Audio
- Video
- Spreadsheets
- Databases
- Code
- Other
- I'm not sure

Please specify other data types

Please list the file extensions you have produced or anticipate producing for this project that you intend to share with others.

Block 3

According to the NSF, a data management plan may include a description of "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)."

Does the data you have produced or intend to produce conform to known standards in your discipline?

- Yes
 - No
 - I'm not sure
-

Please specify the standard(s) you are using.

"Metadata" refers to descriptive information or documentation about data.

Have you produced or do you anticipate producing metadata for this project?

- Yes
 - No
 - I'm not sure
-

Additional comments

Does the metadata you have produced or intend to produce conform to known standards in your discipline?

- Yes
 - No
 - I'm not sure
-

Please specify the standard(s) you are using.

Would you make use of a service to produce metadata for this project?

- Yes, and I would be willing to pay for this service
 - Yes, but I would not be willing to pay for this service
 - No, I would produce metadata myself
-

Block 4

According to the NSF, a data management plan may include a description of "policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements."

Do you anticipate needing to consult with an intellectual property specialist to create a license agreement or usage statement for the data you have produced or intend to produce?

- Yes
 - No
 - I'm not sure
-

When would you be able to share the data you have produced or intend to produce for this project?

- Immediately after collection
 - Immediately after my team has analyzed the data
 - Six months or more after my team has analyzed the data
 - I would not be able to share this data
-

What might prevent you from sharing the data you have produced or intend to produce for this project? Check all that apply. (You may also check no boxes if none apply.)

- Little value to others
 - Confidentiality or privacy issues
 - Commercialization or patent issues
 - Some or all of the data I work with has license or usage restrictions that prevent me from sharing
 - Data requires secure access I am not capable of providing
-

Additional comments

Block 5

According to the NSF, a data management plan may include a description of "policies and provisions for re-use, re-distribution, and the production of derivatives." Furthermore, 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."

Given the NSF expectation to share data with other researchers, how much data would you intend to share?

- I do not plan on sharing data
- No more than 1 GB
- More than 1GB but less than 100 GB

- More than 100 GB but less than 1 TB
- More than 1 TB but less than 100 TB
- More than 100 TB

When you publish your findings from this research project, do you plan on submitting your supporting data to a journal publisher?

- Yes
- No
- I'm not sure

Additional comments

Do you plan on using a custom solution to share the data you have produced or intend to produce? (i.e., Sharing data on a personal or departmental website or FTP server...)

- Yes, and I plan to do this work in-house
- Yes, and I plan to contract all or part of this work
- No
- I'm not sure

Additional comments

Block 6

According to the NSF, a data management plan may include a description of "plans for archiving data, samples, and other research products, and for preservation of access to them."

Do you plan to deposit the data you have produced or intend to produce in Cornell's Institutional Repository, eCommons (<http://ecommons.cornell.edu/about.html>), or would you be interested in doing so to satisfy the NSF requirement?

- Yes
- No
- I'm not sure

Additional comments

Do you plan to deposit the data you have produced or intend to produce in CISER's Data Archive (<http://ciser.cornell.edu/info/about.shtml>), or would you be interested in doing so to satisfy the NSF requirement?

- Yes
 - No
 - I'm not sure
-

Additional comments

Do you plan to utilize the Cornell Restricted Access Data Center http://ciser.cornell.edu/CRADC/What_is_CRADC.shtml) to work with restricted access or limited use licensed data, or would you be interested in doing so to satisfy the NSF requirement?

- Yes
 - No
 - I'm not sure
-

Additional comments

Do you plan to store the data you have produced in the Center for Advanced Computing Disk Farm (<http://www.cac.cornell.edu/services/storage.aspx>), or would you be interested in doing so to satisfy the NSF requirement?

- Yes
 - No
 - I'm not sure
-

Additional comments

Do you plan to deposit the data you have produced or intend to produce in a data center or other non-Cornell repository, or would you be interested in doing so to satisfy the NSF requirement?

- Yes
- No
- I'm not sure

Please specify the repository (or repositories) you plan to deposit your data into.

Block 7

What is your current method of backing up the data you have produced or intend to produce for this project? Check all that apply.

- Own IT infrastructure (e.g., external hard drives)
- EZBackup or other campus-based solution
- Commercial solution (i.e., Google Docs, Amazon S3)
- No backup

Approximately how much data needs to be backed up?

- No more than 1 GB
- More than 1 GB but less than 100 GB
- More than 100 GB but less than 1 TB
- More than 1 TB but less than 100 TB
- More than 100 TB

Block 8

The NSF specifies that if "any PI or co-PI identified on the project has received NSF funding in the past five years, information on the award(s) is required." Specifically, applicants must indicate "evidence of research products and their availability, including, but not limited to: data, publications, samples, physical collections, software, and models, as described in any Data Management Plan."

Do you currently keep track of research outputs and their availability?

- Yes
- No

Additional comments

If there was a service offered where you could enter in basic information about your data (including the description, where it was available on the web) to demonstrate compliance with NSF's policy, would you make use of it?

- Yes
- No
- I'm not sure

Additional comments

Do you anticipate or would you be interested in any sort of guidance, including consultation or instruction, for any of the data management plan components mentioned above?

- Yes
- No
- I'm not sure

Additional comments

Which components are you interested in receiving consultation or instruction for?

A review of the datamangement components:

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.

Block 9

A review of the datamangement requirements:

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.
-

Please share any additional thoughts or concerns you have regarding campus support for complying with funders' data management policies.

Would you like to be contacted in the future about your response or participate in focus groups about Data Management Plans? (Please include your contact information if yes.)