

Research Data Survey Recommendations

Executive Summary

Northwestern researchers have expressed increasing interest in data, data management and data storage due-in part-to Federal mandates requiring data sharing and preservation. The E-Science Working group designed a survey in 2013 to investigate how researchers at Northwestern University manage data and to help determine researchers' needs or requirements for data storage, preservation and sharing.

The 21 question survey looked at several areas involving data and data management, including the types and size of data generated by researchers, current and future needs for data storage, data retention and data sharing, what researchers are doing (or not doing) regarding data management planning, and types of training or assistance needed. There were 831 responses and 788 respondents completed the survey, for a response rate of approximately 6.4%.

The survey results indicate investigators need both short and long term storage and preservation. Key findings include:

- 31% of respondents did not know how much storage they will require, highlighting the difficulty of establishing a correctly sized research storage service.
- Much of Northwestern's research data is stored on local hard drives, departmental servers or equipment hard drives. 31% of respondents use web-based storage services, most notably Dropbox.
- 60% of the respondents indicated that they share or plan to share their research data. Sharing tends to occur within a research group or with collaborators prior to publication, expanding to more public availability after publication.
- Responses on data management plans and support needs indicate a need to provide increased consulting and support services, most notably for data management planning, awareness of regulatory requirements, and use of research software.

The survey results will be used to inform potential services or education opportunities as well as to assist stakeholders in defining future goals regarding data management and infrastructure. Further analysis of the results is also possible upon request.

Next steps and high level recommendations

Northwestern researchers have expressed increasing interest in data, data management and data storage due in part to Federal mandates requiring data sharing and preservation. The E-Science Working group designed a survey in 2013 to investigate how researchers at Northwestern University manage data and to help determine researchers' needs or requirements for data storage, preservation and sharing.

The survey results indicate a need for both short and long term storage and preservation on campus. In addition, respondents highlighted the need for information, training, and other assistance with developing data management plans and complying with funding agency requirements.

We recommend NU Libraries, administration and NUIT consider using these results as a basis for improving NU's infrastructure to support current and future data services. Specific next steps and high level recommendations are listed below. The highest concentration of respondents was found in Feinberg, McCormick, and Weinberg, which may be appropriate units for initial pilots and adoption of new services.

1. Develop additional data management training, targeted at PIs and lab staff (graduate students in particular). Collaboration between Libraries and NUIT
 - a. Carry out at least 4 training sessions in FY2014-2015
 - b. Identify additional venues for publicizing data management plan development support
2. Convene a research computing governance group, and appoint an IT governance level 3 committee with a focus on data management services and support.
3. The Libraries and NUIT provide introductory research data lifecycle consulting services during FY15, using the north campus Research Computing Services office located at Mudd Library.
 - a. publicize availability of consulting services with Northwestern researchers
 - b. Explore additional avenues for expanding consulting services to Chicago in collaboration with Galter Library.
4. The Libraries and NUIT, in consultation with the Office for Research, should scope and pilot a repository to support data deposit, publishing and preservation; outcome of the pilot will be to understand cost and governance models, and possible partnerships with other universities or agencies.
5. Analyze data management plans submitted with NSF proposals to understand what researchers are saying they are doing to preserve and share their data

6. Develop a set of guidelines recommending how Northwestern researchers can best comply with funders data policies; the existing [Data Management Plan](#) support site should be expanded to house this comprehensive advice; also add guidelines as boilerplate to the DMPTool pages for Northwestern researchers
7. Schedule targeted conversations with key stakeholders to present the survey and discuss results and recommendations; determine who from Libraries, NUIT, Office for Research should participate in conversations with:
 - a. Research deans
 - b. School IT management
 - c. Academic department chairs
 - d. Open fora for discussion with faculty and graduate students

Conclusions

The responses to this survey were generally positive. Some humanists and social scientists felt that this survey did not apply to them, even though the ESWG tried to make it clear that the survey applied to all disciplines.

In the survey, a question soliciting open comments received a number of varied responses. The largest percentage of comments indicated respondents' desire for comprehensive University-level policies and guidelines on data management. This can prove difficult, due to the diversity of types and size of data and discipline-specific requirements for data management and sharing. Responses to the survey, especially written responses to open questions, suggest that Northwestern University is similar to institutions studied by Tenopir in her 2011 study [7]. Storage and management of data, especially over the long-term, is an issue of concern for researchers, and no single solution may fit the needs of all disciplines. Also, the survey shows that there is a need for assistance and education regarding data management across all user groups at Northwestern.

Additionally, the results can be employed by stakeholders in defining future goals regarding data management and infrastructure. The report can be shared with heads of university libraries, NUIT, offices for research and graduate education and with school administrators, to guide the ESWG and other interested offices to formulate services in these areas. The survey data will be archived so that additional interrogations, such as school-specific cross-tabulations, can be performed as needed.