Federal Data Sharing

Sunday, February 5, 2017

Author(s)

Danielle Friend

This post accompanies the Sci on the Fly podcast "[Data Science (https://pfellowsqaweb02.aaas.org/blog/podcast-what-makes-data-scientist)](https://pfellowsqaweb02.aaas.org/blog/podcast-what-makes-data-scientist)."

In the last Sci on the Fly podcast on data science, several data scientists were asked "What is data science?" Although each data scientist thinks of data science differently, one conversation that continues to be discussed is how the data science community should share data.

In the podcast, Amy Nurnberger, a research data manager at the Center for Digital Research and Scholarship at Columbia University, discussed the importance of changing the culture around data sharing in order for data sharing to become the norm. Additionally, Sarah Callaghan at the Center for

Environmental Data Analysis mentioned the role incentives for data sharing play in encouraging researchers to share their data.

Within the last few years, the discussion of "open data" or "public access" to data has increased drastically within the data science community and this is likely due in part to Federal initiatives calling for increased data sharing. On February 22, 2013, the White House Office of Science and Technology Policy (OSTP) released a memorandum, *Increasing Access to the Results of Federally Funded Scientific Research* (https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf), directing federal agencies and offices to develop plans indicating how they will make peer-reviewed, federally-funded publications and scientific data accessible to the public, industry, and the scientific community. That same year, President Obama also released an executive order entitled *Open Data Policy-Managing Information as an Asset* (https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf) which calls Federal agencies to collect and create information in a way that supports downstream information processing and dissemination. This mandate encourages the use of machine-readable and open formats, and the development of data standards and metadata for all new information creation and collection efforts.

In response to these federal initiatives, several Federal agencies have since developed and made public, plans that outline their efforts for complying with these directives. To get a better idea of how federal agencies are implementing these directives, let's take a look at what the National Institutes of Health (NIH) has done to move public access forward.

Data Management Activities at the National Institutes of Health

In response to the Federal Initiatives, the NIH developed the *National Institutes of Health Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research* (https://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf). While the NIH has already met federal standards governing public access to publications, the NIH plan focuses on what the NIH will do in regards to data. The NIH plan indicates that the NIH is considering:

- Developing policies that would require NIH-funded researchers to make the data underlying the conclusions of peer-reviewed scientific research publications available to the public at the time of publication.
- Ensuring that all NIH-funded researchers prepare data management plans that will be evaluated during peer review. The plans may be required to include, for example, the data that will be produced in the study, mechanisms for providing access to the data, and plans for long-term preservation of the data.
- Encouraging supported researchers to deposit data in established public repositories and encouraging researchers to use data standards relevant to their research community.
- Exploring approaches for improving the discoverability of NIH-funded data and exploring ways to advance data as a form of scholarship, for example by standardizing data citation.
- Developing shared space for storage of basic and clinical research data and software.

What has the NIH done since the development of the plan?

Since the 2015 release of the NIH Plan, the NIH has sought feedback from stakeholders on several topics associated with data sharing through the release of several public comments periods. For example, in August 2016 the Request for Information on Metrics to Assess Value of Biomedical

Digital Repositories (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-133.html) was issued to obtain feedback on approaches to measure and evaluate biomedical data repositories. The information gathered from this request for information will provide significant insight into determining the usage and value of data repositories and the prioritization of activities related to data storage. In October 2016 the Request for Information on Including Preprints and Interim Research Products in NIH Applications and Reports (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-006.html) sought public feedback on the inclusion of preprints, or a complete and public draft of a scientific document, intended to be submitted to a journal for peer-review and subsequent publication, and other non-publication products in annual grant progress reports and grant applications.

Finally, in November 2016 the NIH released the Request for Information on Strategies for Data Management, Sharing, and Citation (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-015.html) in order to obtain feedback on the highest-priority data to be shared, the length of time data should be made publicly available, and barriers for sharing data as well as mechanism to overcome those barriers.

The information from all three requests for information will assist with decision making related to funding data repositories, developing mechanisms for citing shared products, and shaping policies regulating what should be shared. Although specific plans for moving forward based on the results of these requests for information have yet to be made public, the gathered information will provide valuable information to the NIH about how to move forward and improve sharing federally funded research.

Source URL: *https://www.aaaspolicyfellowships.org/blog/federal-data-sharing*

List of links present in page

- https://pfellowsqaweb02.aaas.org/blog/podcast-what-makes-data-scientist
- https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf
- https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf
- https://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf
- https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-133.html
- https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-006.html
- https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-015.html