Data Management: Streamlining Your Research and Publication Pipeline

Veronica Ikeshoji-Orlati, CLIR Postdoctoral Fellow
Elisabeth Shook, Scholarly Communications Librarian

In February 2013, the White House Office of Science and Technology Policy released a memorandum stating that United States' taxpayers deserved to have access to all publicly funded research, including the underlying data. How can you respond to federal mandates - and increasing calls for research reproducibility - in your research and data management practices?

Craft a Data Management Plan

A data management plan (DMP) is a roadmap for your research data throughout the lifecycle of your project. In addition to being required to win federal research funding, a DMP can ensure the sustainability and viability of your project through every phase of its development. To write a useful DMP, be explicit about each aspect, including who your collaborators are; what types of data you plan to collect; the file formats, technological infrastructure, and PHI/PII de-identification practices you will use; and the lifespan of your data. Check out the DMPTool for guidance in crafting a DMP, and ask the Data Management Librarian at your institution to review your DMP before submission.

Streamline Your Research Workflow

Once you have written a DMP, follow it to maximize your research efficiency and streamline your publication pipeline. The Open Science Framework (OSF) is a free, open-source platform for doing so. It ties together the myriad tools which facilitate collaborative research such as citation managers (e.g. Mendeley, Zotero), cloud storage services (e.g. Dropbox, Google Drive, Amazon S3), and data repositories (e.g. figshare, Dataverse). By integrating the OSF into your workflow, you ensure that you and all of your collaborators can transition from idea incubation to final publications with ease.

Publish Your Data

The next step in following your DMP is to publish your data in a data repository. When searching for a suitable data repository, consider the facets of your data outlined in your DMP. Two great tools for identifying suitable repositories based on the specific needs of your data are the Registry of Research Data Repositories (re3data) and the PLOS Recommended Repositories. Another option for publishing your data may be your library's Institutional Repository (IR), an openly available database of work generated by your institution.

Get Cited

There are two ways to increase your citations after you publish your dataset. A Digital Object Identifier (DOI) is a unique, permanent link to published materials, from journal articles to datasets. For help identifying repositories that assign DOIs, contact your institution's Scholarly Communications Librarian. Another method of ensuring credit for your data is by creating an ORCID, a unique, 16-digit persistent identifier for *you* as a researcher. When you publish, add the DOI to your ORCID profile to guarantee you receive credit for your work!