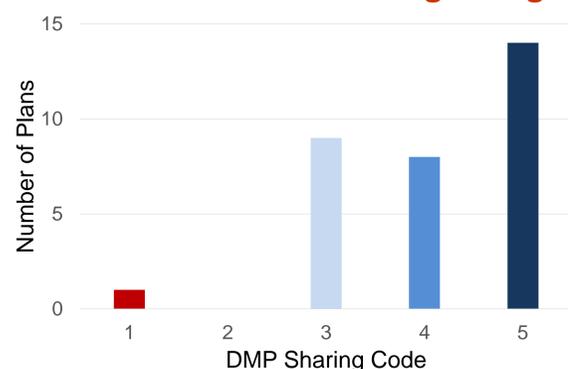# Shares Well With Others: Current Trends in DMP Data Sharing at UIUC

Cassia Smith, Elizabeth Wickes, Elise Dunham, and Heidi Imker
Research Data Service, University Library, University of Illinois at Urbana-Champaign

School of
Information Sciences
The iSchool at Illinois

RESEARCH DATA SERVICE
RESEARCHDATASERVICE.ILLINOIS.EDU

UNIVERSITY LIBRARY

Total DMPs in Each Sharing Category

Number of Plans / DMP Sharing Code

DMP Sharing Correctness Scale

1 – **Not planning to share the data**, where that is not due to a necessary legal or ethical restriction.

2 – Plans to **share the data by request only**, when that is not required by the nature of the data.

3 – Wants to **share the data, but not through a repository**. (For example, through a project website.)

4 – Wants to **share the data in a repository, but may need help** choosing the best option.

5 – **Choice of sharing option is ideal**, including projects where sharing is not possible due to the nature of the data. Wording may need to be tweaked, but that is the only modification needed.

Sharing Correctness by Fiscal Year

Total Reviews by Code / Fiscal Year

Introduction

The Research Data Service (RDS) offers Data Management Plan (DMP) reviews for UIUC researchers. **We evaluated these DMPs and our feedback to get a feel for the most common mistakes and issues found in the DMPs.** The goal was to improve our standard guidance and resource materials, and to get an understanding of how DMP quality affects workflow and staff time requirements.

One focus of this evaluation was data sharing. Most federal agencies that require a DMP also require that the project data be shared. **The accepted wisdom in the data curation and open science communities is that researchers are adamantly against sharing data.** There are frequent conversations and presentations dedicated to helping university research data services convince recalcitrant researchers to release their data. We wanted to know how researchers at our institution typically handled this requirement.

Method and Limitations

The review of DMPs and RDS feedback collected the following data:

- Basic DMP metrics
  - funding agency, date received
- Basic workflow metrics
  - turnaround time, referral sources
- **Qualitative coding of DMP content**
  - problems identified, including severity scales
  - **Correctness of DMP sharing option**

The project only evaluated 32 reviews (the total number of reviews received between 2015-05-15 and 2016-11-11). The DMPs are all for federal grants, and they only represent projects that were interested in receiving DMP feedback from the RDS.

Preliminary Results

One finding that has emerged quickly from the initial coding and early analysis is that **DMPs are typically much more willing to share data than accepted knowledge would suggest**. Almost all of the DMPs in our project intended to share results of their research on a public platform. DMP review guidance provided by RDS **tended to involve suggesting a repository to share the data** (rather than a website), or **selecting a more appropriate repository** for the project's research domain. Those DMPs that did not intend to share all the data usually had important legal considerations, such as privacy or proprietary data restrictions, influencing their decisions. Only one DMP did not initially intend to share the data when there were no legal restrictions guiding that choice.
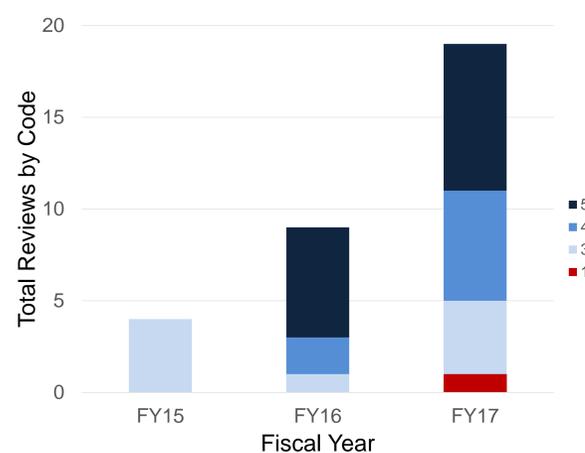
Conclusions

These findings suggest that **data sharing in DMPs is actually fairly widespread**, at least among UIUC researchers who seek our feedback. They also indicate that there is still confusion over what data sharing entails. **Academic data curation professionals may want to shift their focus from advocating for data sharing to educating researchers on the most robust and stable options for sharing their data.**

How Should You Share Your Data?

Many DMP authors seemed uncertain how to go about sharing their data.

Through a domain repository – The ideal way to share a dataset is through an established data repository dedicated to the researcher's field of interest, such as GenBank for genomics research.

Through a data repository – Any reputable data repository should have basic digital preservation practices and robust backups in place. This could be a commercial repository like figshare or Zenodo or an institutional repository like Illinois Data Bank.

Through a project website – Project websites can be a great way to promote a project, but your data is more likely to be preserved for the long term if it's also deposited in a data repository.

By request – This may be the most appropriate option for certain sensitive or otherwise restricted datasets. For the purposes of backups and preservation, it would still be wise to consider a repository that allows granular access controls, such as IDEALS.

On a department server – Storing your data long-term on a departmental server could make it hard to achieve the recommended three backups, and it places the burden of digital preservation on your department's IT staff. This option also unnecessarily restricts access without guaranteeing security.

Carrier pigeon – While a very cute data sharing option, delivery becomes less reliable during hunting season. There are also bandwidth restrictions and the need for a grad student to clean up droppings.

Acknowledgments