Researcher Perspectives: Realities of Academic Data Sharing

Institution

1) Please select your current University affiliation:*  
   ( ) Cornell University  
   ( ) Duke University  
   ( ) University of Michigan  
   ( ) University of Minnesota  
   ( ) Virginia Tech  
   ( ) Washington University in St. Louis

Demographics

2) Name*

3) Email*

4) Institution*

5) Department/Organizational Unit*
Broad Access to Research Data

6) Did you make research data from your most recent grant broadly available (e.g., in a data repository or as part of a research article, whether restricted or open access)?*

( ) Yes
( ) No

7) Why wasn't research data from your most recent grant made broadly available? Please be as specific as possible.

______________________________________________________________________________
______________________________________________________________________________

8) After the grant or project ended, how was the research data stored and maintained? (Select all that apply)

[ ] Maintained and managed indefinitely on a server, hard drive, or cloud
[ ] Maintained to be shared at a later date
[ ] Deposited in a closed (not shared) archive
[ ] Stored privately but shared upon request
[ ] The research data, including any de-identified data, have been destroyed or will be destroyed
[ ] Other - Write In: ____________________________________________________________

Data Sharing Practices

9) Please select where you shared your research data for your most recent grant. (Select all that apply)
10) What online repository (or repositories), including institutional, generalist, or disciplinary, did you use to share your research data? (For example: Dryad, Dataverse, Figshare, dbGAP, Morphosource, ICPSR, Deep Blue, etc.)

11) From the list below please rank the top 5 factors which have the greatest influence on your decisions on where to share your research data. If more than 5 factors influence your decisions, please include them in the ranking as well.

______Personal experience
______Least expensive option
______Easiest/quickest option
______The option to personally control and manage content
______Funder recommendation
______Disciplinary/scholarly society recommendation
______Peer or colleague recommendation
______Journal/publisher recommendation
______Library recommendation
______Research office recommendation
______Campus IT recommendation
______Other
### Data Sharing Activities

The following questions will ask about various activities that you may or may not have engaged in throughout the various phases of your grant-funded project. For each activity, please select the option that best fits if and how you engaged with various support services when completing that activity.

If you completed the activity yourself, or with the assistance of others in your research lab (such as a graduate student), please select "Did it myself or within the lab/research team." If you consulted with someone within your University (such as a librarian or IT consultant), please select "Did it with institutional assistance." If you consulted with someone outside your University and research team (such as a funder Program Officer), please select "Did it with external assistance."

12) How did you primarily address each of the following data sharing activities in the 'Planning, Design and Start Up of Projects' phase of your research? For each activity, please select the option that best fits if and how you engaged with various support services when completing it.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Did not do</th>
<th>Did it myself or within the lab/research team</th>
<th>Did it with institutional assistance</th>
<th>Did it with external assistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparing Data Management Plans (DMPs) or Data Management and Sharing (DMS) Plans</td>
<td>()</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Identifying data management and sharing costs to be included in grant budgets</td>
<td>()</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>
13) How did you primarily address each of the following data sharing activities in the 'Data Collection, Storage, and Management' phase of your research? For each activity, please select the option that best fits if and how you engaged with various support services when completing it.

<table>
<thead>
<tr>
<th>Activity</th>
<th>( )</th>
<th>( )</th>
<th>( )</th>
<th>( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparing IRB protocols and informed consent for data sharing</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Determining storage solutions for active research data</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Selecting an appropriate repository (or repositories) for making research data broadly available</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Evaluating data security needs</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Determining intellectual property and copyright considerations</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Developing Materials Transfer and/or Data Use Agreements</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Data Sharing Activities</td>
<td>Did not do</td>
<td>Did it myself or within the lab/research team</td>
<td>Did it with institutional assistance</td>
<td>Did it with external assistance</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------</td>
<td>---------------------------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Developing documentation of data (e.g., data dictionary, protocols)</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Creating quality control mechanisms or procedures</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Evaluating data analysis tools and processes to support sharing and reproducibility</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Managing active data (e.g., storage, security, backup, lab notebooks)</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
</tbody>
</table>

**Data Sharing Activities**

14) How did you primarily address each of the following data sharing activities in the 'Making Data Broadly Available' phase of your research? For each activity, please select the option that best fits if and how you engaged with various support services when completing it.
<table>
<thead>
<tr>
<th>Activity</th>
<th>Did not do</th>
<th>Did it myself or within the lab/research team</th>
<th>Did it with institutional assistance</th>
<th>Did it with external assistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making decisions about what data to share or host</td>
<td>()</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Submitting data into a data sharing platform (e.g., institutional repository, generalist repository, disciplinary repository)</td>
<td>()</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Preparing data for sharing (e.g., de-identification, selection, curation, data cleaning, validation, and quality control)</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Creating documentation for sharing (e.g., structured metadata, README files)</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Selecting or applying licenses for data reuse</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Migrating data file formats to be more open or accessible</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
</tbody>
</table>
### Data Sharing Activities

15) How did you primarily address each of the following data sharing activities in the 'Data Retention, Including Preservation, Archive, and Long-Term Access' phase of your research? For each activity, please select the option that best fits if and how you engaged with various support services when completing it.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Did not do</th>
<th>Did it myself or within the lab/research team</th>
<th>Did it with institutional assistance</th>
<th>Did it with external assistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migrating files to new formats or systems as needed</td>
<td>()</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Monitoring integrity of preserved data</td>
<td>()</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Making decisions about de-accessioning and removal of research data</td>
<td>()</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Ensuring data security when</td>
<td>()</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>
appropriate (e.g., PHI/HIPAA, Export Controls, FISMA, student data, and Intellectual Property)

<table>
<thead>
<tr>
<th>Did not do</th>
<th>Did it myself or within the lab/research team</th>
<th>Did it with institutional assistance</th>
<th>Did it with external assistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ensuring funding agency requirements for data sharing have been met</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>Compiling reports for project closeout</td>
<td>()</td>
<td>()</td>
<td>()</td>
</tr>
</tbody>
</table>

16) Which of the following best describes how you primarily addressed each of the following data sharing activities in the 'Project Closeout and Compliance' phase of your research? For each activity, please select the option that best fits if and how you engaged with various support services when completing it.

Data Sharing Expenses
17) Did you hire or rely on staff (including graduate and undergraduate students) for making the research outputs of this grant broadly available?

( ) Yes
( ) No

**Staffing Expenses**

18) Approximately how many staff did you hire or rely on to make the research outputs of this grant broadly available? (Please enter a whole number only, e.g., 8)

19) Of all the staff members that you hired or relied on for making research outputs of your grant broadly available, please consider the five positions most responsible for this work. If there were less than five positions responsible for this work, please fill out the information for the number of positions accordingly.

What percentage of time did each of these five positions dedicate to data sharing and what was their approximate salary?

<table>
<thead>
<tr>
<th>Employee 1</th>
<th>Percentage of time dedicated to data sharing (e.g., 50%)</th>
<th>Approximate Salary (e.g., $70,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

20) If more than five positions were responsible for making research outputs of your grant broadly available, please approximate the total cost of all salaries (annually) for the remaining positions.
Average percentage of time dedicated to data sharing (e.g., 10%): __________
Average annual salaries (e.g., $75,000): __________

21) How did you pay for the staff who made the research data broadly available? (Select all that apply)

[ ] Grant funded
[ ] Department funded
[ ] Lab funded
[ ] Other - Write In: __________________________________________________________

Infrastructure Expenses

22) Approximately how much money did you spend on infrastructure (such as storage, hardware, software, etc.) to provide broad access to your grant’s research data?

( ) $0
( ) $1 - $4,999
( ) $5,000 - $9,999
( ) $10,000 - $19,999
( ) $20,000 - $29,999
( ) $30,000 or greater, please specify amount ($):
________________________________________________________

23) How did you pay for infrastructure to make the grant's research data broadly available? (Select all that apply)

[ ] Grant funded
[ ] Department funded
[ ] Lab funded
[ ] Other - Write In: ________________________________________________

---

Data Sharing Expenses

24) Thinking about your grant, which activities or phase(s) of preparing your data for sharing cost the most?

Please rank each step below in order of greatest cost to least cost. (If you did not spend any money on activities for a phase, you can omit it from the ranking)

_______ Planning, Design and Start Up of Projects (e.g., preparing a data management plan, determining intellectual property and ownership)

_______ Data Collection, Storage, and Management (e.g., documenting data, managing active data)

_______ Making Data Broadly Available (e.g., preparing data for sharing, selecting or applying licenses, migrating formats of data to be open or more accessible)

_______ Data Retention and Preservation (e.g., migrating files to new formats or systems as needed, monitoring integrity of preserved data, making decisions about de-accessioning and removal)

_______ Project Closeout and Compliance (e.g., ensuring funding agency requirements for data sharing have been met, compiling reports for project closeout)

---

Additional Considerations

25) How likely are you to make your research data broadly available if you were not required to do so (by publishers or funding agencies)?

( ) Very likely
26) We are interested in collecting and sharing more detailed case studies about researchers’ experiences in sharing data.

Would you be willing to participate in a follow up 45-60 minute group interview with your grant’s research team to discuss your data sharing practices and their requisite costs?

( ) Yes
( ) Not at this time

Thank You!

We appreciate the time you took to take this survey!

To find out more about the Realities of Academic Data Sharing Initiative and project outcomes, please visit our website.