Institutional Infrastructure: Realities of Academic Data Sharing

Select 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) Department/Organizational Unit*

_____

5) Position/Title:

_____

Broad Access to Research Data

6) In order to share data or research outputs broadly at the end of a study, researchers need to plan and manage their data across the entire course of their project. This often involves working with institutional offices at various phases in this process. Below are phases of a general research lifecycle with example activities that researchers may need to complete to make data or research products available.

In which of the following phases does your unit or department provide support for these or associated activities? (Select all that apply) *

[ ] **Planning, Design, and Start Up of Projects** (e.g., developing Data Management Plans, complying with IRB protocols, ensuring grant compliance)

[ ] **Data Collection, Storage, and Management** (e.g., developing documentation of data, creating quality control mechanisms, managing active data)

[ ] **Making Data Broadly Available** (e.g., making decisions about what data to share or host, selecting or applying licenses to data, creating persistent identifiers)

[ ] **Data Retention Including Preservation, Archive, and Long-Term Access** (e.g., submitting data to a shared platform, migrating files to new formats, ensuring data security)

[ ] **Project Closeout and Compliance** (e.g., ensuring funding agency requirements for data sharing have been met, ensuring grant compliance)

[ ] Other - Write In: _____

[ ] None

Planning, Design, and Start Up of Projects

This phase of the research data lifecycle includes all data-related activities associated with planning, design, and startup of research projects.

7) In the 'Planning, Design, and Start Up of Projects' phase, which of the following activities does your unit/department support or offer services for? (Select all that apply)

[ ] Reviewing or preparing Data Management Plans (DMPs) or Data Management and Sharing (DMS) Plans

[ ] Reviewing data management and sharing costs to be included in grant budgets [ ] Reviewing IRB protocols and informed consent for data sharing

[ ] Developing, building, or recommending storage solutions for active research data

[ ] Supporting an appropriate repository (or repositories) for making research data broadly available

[ ] Assessing data security needs and recommending solutions

[ ] Supporting intellectual property and copyright considerations

[ ] Developing or reviewing Materials Transfer and/or Data Use Agreements

[ ] Other - Write In: _____

Data Collection, Storage, and Management

This phase of the research data lifecycle includes all data-related activities associated with the collection, storage, and management of active data.

8) In the 'Data Collection, Storage, and Management' phase, which of the following activities does your unit/department support or offer services for? (Select all that apply)

[ ] Developing or reviewing documentation of data (e.g., data dictionary, protocols)

[ ] Creating quality control mechanisms or procedures

[ ] Evaluating or recommending data analysis tools and processes to support sharing and reproducibility

[ ] Managing active data (e.g., storage, security, backup, lab notebooks)

[ ] Other - Write In: _____

Making Data Broadly Available

This phase of the research data lifecycle includes all data-related activities associated with making data broadly available.

9) In the 'Making Data Broadly Available' phase, which of the following activities does your unit/department support or offer services for? (Select all that apply)

[ ] Consulting on decisions about what data to share or host

[ ] Providing or hosting repositories for making data available

[ ] Preparing or consulting on preparing data for sharing (e.g., de-identification, selection, curation, data cleaning, validation, quality control)

[ ] Submitting data into a data sharing platform (e.g., institutional repository, generalist repository, disciplinary repository)

[ ] Creating or reviewing documentation for sharing (e.g., structured metadata, README files)

[ ] Consulting, selecting, or applying licenses to data

[ ] Recommending or migrating data file formats be open or more accessible

[ ] Creating or recommending persistent identifiers (PIDs; e.g., DOIs)

[ ] Checking for compliance with existing Data Use Agreements (DUAs)

[ ] Other - Write In: _____

Data Retention, Including Preservation, Archive, and Long-Term Access

This phase of the research data lifecycle includes all data-related activities associated with data retention, including preservation, archive, and long-term access.

10) In the 'Data Retention, Including Preservation, Archive, and Long-Term Access' phase, which of the following activities does your unit/department support or offer services for? (Select all that apply)

[ ] Consulting on or migrating files to new formats or systems as needed

[ ] Ensuring data security when appropriate (e.g., PHI/HIPAA, Export Controls, FISMA, student data, and Intellectual Property)

[ ] Monitoring integrity of preserved data

[ ] Making decisions about de-accessioning and removal of research data

[ ] Other - Write In: _____

Project Closeout and Compliance

This phase of the research data lifecycle includes all data-related activities associated with project closeout and compliance.

11) In the 'Project Closeout and Compliance' phase, which of the following activities does your unit/department support or offer services for? (Select all that apply)

[ ] Ensuring funding agency requirements for data sharing have been met

[ ] Providing compliance support around research project reports

[ ] Other - Write In: _____

Expenses

12) Approximately how many staff did you hire or rely on to support public access to research data? (Please enter whole numbers only, e.g., 8)

_____

13) Of the staff members hired or responsible for making research outputs 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?

	Percentage of time dedicated to data sharing (e.g., 50%)	Approximate Salary (e.g., $55,000)
Employee 1		
Employee 2		
Employee 3		

Employee 4		
Employee 5		

14) If more than five positions contributed to making research outputs broadly available, please approximate the average percent of time and average cost of all salaries (annually), per person, for the remaining positions.

	Average percentage of time dedicated to data sharing (e.g., 10%)	Average annual salaries (e.g., $75,000)
Remaining Positions (more than 5)		

15) For services, infrastructure, or staffing costs related to data sharing across the research lifecycle, what was your approximate budget (e.g., software, contracts, fees) for 2021-2022?

Recurring annual expenses ($): _____

One time expenses ($): _____

16) Please rank each research lifecycle phase below, in order of greatest to least support cost during the last fiscal year. If your unit/department did not have expenditures for one or more phases, you can omit them from the ranking.

_____ **Planning, Design and Start Up of Projects** (e.g., preparing data management plans, 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)

Investments

17) In 2013, the Office of Science and Technology Policy (OSTP) required federal agencies to create a plan for requiring public access to research data and other outputs from funded projects. This led to Data Management and Sharing Plan requirements for the Department of Energy (DOE), the National Institutes of Health (NIH), the National Science Foundation (NSF), and other agencies.

Since this time (2013), has your unit increased investments (staff, infrastructure, or services) to support activities related to data sharing due to this OSTP policy, other federal policies, or other local policies or requirements?

Please do not consider activities, infrastructure, or staffing costs paid for by PIs/investigators by their project funds, as they will be surveyed separately for their expenditures.

( ) Yes

( ) No

18) Which areas has your unit increased investments in to support the activities related to data sharing? (Select all that apply)

[ ] Additional staffing
[ ] Additional technical infrastructure
[ ] Additional services
[ ] Other – Write In: _____

19) What additional level of investment (staff, infrastructure, or services) does your unit anticipate making within the next 5 years in each of the following areas?

	None (0% change to current investment)	Minimal (<50% addition to current investment)	Moderate (50-100% addition to current investment)	Substantial (>100% addition to current investment)
Planning, Design, and Start Up of Projects	( )	( )	( )	( )
Data Collection, Use, and Management	( )	( )	( )	( )
Making Data Broadly Available	( )	( )	( )	( )
Data Retention, Including Preservation, Archive, and Long-Term Access	( )	( )	( )	( )
Project Closeout and Compliance	( )	( )	( )	( )

Additional Expenses and Considerations

20) Has your department/unit incurred any additional expenses to support activities related to sharing research data that were not reflected in your previous responses?

( ) Yes

( ) No

21) Please tell us more about these additional expenses.

Can you describe the additional expense(s) - was it to support a service, increase staffing or infrastructure, or something else?

_____

_____

What is the estimated dollar amount of the additional expense(s)? (Please enter numbers only):

_____

22) Did your unit or department partner with other units or departments to make federally funded research data broadly available?

( ) Yes

( ) No

23) Please state which unit(s) and/or department(s) your office or unit partnered with to make federally funded research data broadly available.

_____

24) Please describe any additional considerations or context you'd like to provide us regarding your services and expenses for broad access to research data that have not been discussed elsewhere.

_____

_____

25) If needed, may we contact you for an additional 20-45 minute follow up interview? We may need clarification on your survey responses or request additional information on your unit's expenditures.

( ) Yes

( ) No thank you

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.