

# Publicly Shared Data: A Gap Analysis of Researcher Actions and Institutional Support throughout the Data Life Cycle

March 2024

This work is licensed under a Creative Commons Attribution 4.0 International License.



<https://doi.org/10.29242/report.radsgapanalysis2024>



U.S. National  
Science  
Foundation

## Authors

Jonathan Petters

Virginia Tech

ORCID: [0000-0002-0853-5814](https://orcid.org/0000-0002-0853-5814)

Joel Herndon

Duke University

ORCID: [0000-0001-9995-9040](https://orcid.org/0000-0001-9995-9040)

Shawna Taylor

Association of Research Libraries

ORCID: [0000-0002-9842-7867](https://orcid.org/0000-0002-9842-7867)

Wendy Kozlowski

Cornell University Library

ORCID: [0000-0001-6539-3798](https://orcid.org/0000-0001-6539-3798)

Alicia Hofelich Mohr

University of Minnesota

ORCID: [0000-0002-7644-4105](https://orcid.org/0000-0002-7644-4105)

Jennifer Moore

Washington University in St. Louis

ORCID: [0000-0001-6628-6820](https://orcid.org/0000-0001-6628-6820)

Jake Carlson

University at Buffalo

ORCID: [0000-0003-2733-0969](https://orcid.org/0000-0003-2733-0969)

Cynthia Hudson Vitale

Association of Research Libraries

ORCID: [0000-0001-5581-5678](https://orcid.org/0000-0001-5581-5678)

Lizhao Ge

George Washington University

ORCID: [0009-0005-7862-6016](https://orcid.org/0009-0005-7862-6016)

This material is based upon work supported by the US National Science Foundation grant [#2135874](https://www.nsf.gov/awardsearch/showAward?AWD_NUM=2135874).

Suggested citation: Petters, Jonathan, Shawna Taylor, Alicia Hofelich Mohr, Jake Carlson, Lizhao Ge, Joel Herndon, Wendy Kozlowski, Jennifer Moore, and Cynthia Hudson Vitale. *Publicly Shared Data: A Gap Analysis of Researcher Actions and Institutional Support throughout the Data Life Cycle*. Washington, DC: Association of Research Libraries, March 2024. <https://doi.org/10.29242/report.radsgapanalysis2024>.

# Table of Contents

<b>Executive Summary.....</b>	<b>4</b>
<b>Background and Research Aims .....</b>	<b>8</b>
<b>Reported Institutional Data Sharing Support .....</b>	<b>11</b>
<a href="#">Visualization 1</a> : Data management and sharing activities supported by libraries across the six RADS institutions .....	13
<a href="#">Visualization 2</a> : Data management and sharing activities supported by information technology (IT) offices across the six RADS institutions .....	14
<a href="#">Visualization 3</a> : Data management and sharing activities supported by administrative research offices (for example, Office of Research, Sponsored Projects Office) across the six RADS institutions .....	15
<a href="#">Visualization 4</a> : Data management and sharing activities supported by research institutes, academic departments, institutional centers, and related centers, across the six RADS institutions .....	17
<b>Researcher-Reported Data Management and Sharing Activities .....</b>	<b>19</b>
Researcher Data Sharing Activities .....	20
Researcher Data Sharing Mechanisms .....	26
<b>Discussion: Service Gaps and Institutional Opportunities .....</b>	<b>29</b>
Opportunities for IT Departments .....	30
Opportunities for Central Research Offices.....	31
Opportunities for Research Institutes and Specialized Centers .....	31
Opportunities for Libraries.....	31
Institution-Wide Opportunities .....	33
Coordinating Efforts to Increase Data Sharing Efficiencies .....	36
Limitations.....	37
<b>Assessing the Institutional Landscape Locally .....</b>	<b>38</b>
Version 3 of the RADS Public Access Data Management and Sharing (DMS) Activities .....	38
<b>Conclusion and Recommendations .....</b>	<b>39</b>
<b>Research Instruments.....</b>	<b>41</b>
<b>Data Availability Statement .....</b>	<b>41</b>
<b>Acknowledgements.....</b>	<b>41</b>
<b>For Further Reading .....</b>	<b>42</b>
<b>Appendix A: Administrative Unit Categorization.....</b>	<b>45</b>
<b>Appendix B: Tableau Visualization Links and Labels .....</b>	<b>49</b>
<b>Appendix C: Ranked Factors Influencing Researcher Decisions on Where to Share Data .....</b>	<b>51</b>

## Executive Summary

In recent years, the landscape of research data management and sharing has undergone significant transformation. This transformation, significantly influenced by the 2013 White House Office of Science and Technology Policy (OSTP) Holdren Memo<sup>1</sup> and further reinforced by the 2022 OSTP Nelson Memo,<sup>2</sup> has increased transparency and accessibility in academic research across the United States and has led the largest federal funders of academic research to mandate public access to funded research output. This shift has placed a substantial responsibility on institutions, which are now tasked with not only stewarding research data but also ensuring compliance with the conditions of externally awarded grants to individual researchers at their institutions. To meet these new requirements, institutions have invested in the development and maintenance of robust infrastructure and services for data management and sharing. Understanding how researchers manage and share data, and whether or not they use institutional services or external resources towards these activities, is vital for institutions aiming to make informed decisions in enhancing their data sharing infrastructure.

This report provides the results of research conducted at six academic institutions as part of the US National Science Foundation (NSF) grant “Completing the Lifecycle: Developing Evidence-Based Models of Research Data Sharing” (#2135874), under the Association of Research Libraries (ARL) Realities of Academic Data Sharing (RADS) Initiative. Administrators with expenditure knowledge of their units, and whose units provide data management and sharing (DMS) support to

---

1 John P. Holdren, “Increasing Access to the Results of Federally Funded Scientific Research,” memorandum, Office of Science and Technology Policy, Executive Office of the President, February 22, 2013, [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf).

2 Alondra Nelson, “Ensuring Free, Immediate, and Equitable Access to Federally Funded Research,” memorandum, Office of Science and Technology Policy, Executive Office of the President, August 25, 2022, <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf>.

researchers, were surveyed on precisely what services and activities their units provide to enable data sharing. Funded researchers at these same six institutions were also surveyed on what activities they do or do not do, with or without institutional or external support, to enable sharing of their research data. This report highlights where service gaps may exist between researchers' needs and what services and support institutions provide.

From this analysis, we determined several DMS activities that could use more institutional support, either by specific service providers or as cross-institution efforts, or activities that are underutilized by researchers. Our primary findings are as follows:

- **Nearly half of researchers reported not identifying data management and sharing costs in the planning phase of their grant budgets.** Making data publicly available comes at a cost, and appropriately planning for resource allocation increases the chances of generating well-documented and reusable research data.
- **Researchers could benefit from existing institutional support** for several data management and sharing activities. This includes support for:
  - assessing research data security, from IT offices;
  - ensuring funding agency requirements for data sharing have been met, from central research offices;
  - creating persistent identifiers (PIDs) for shared datasets, from research libraries;
  - making decisions about de-accessioning and removal of research data; and
  - budgeting for data management and sharing costs, from cross-institutional efforts.
- **Researchers completed the majority of data management and sharing activities themselves, and relied much less on internal and external support.** For 20 of the 26 data management and

sharing activities researchers were asked about, 70% or more researchers who did that activity reported that they conducted it themselves or within their research group.

- Administrators from research institutes and specialized centers in **half of the institutions participating in this survey reported offering support for all 27 data management and sharing activities** administrators were asked about, across the research life cycle.

Our recommendations are as follows:

- Previous reports have indicated the importance of forming institution-wide research data management working groups or committees.<sup>3</sup> Once formed, **these working groups should focus on coordinating efforts to increase data sharing services and infrastructure efficiencies across the institution.**
- **Research institutes and specialized centers can provide an opportunity for institutions to pilot or experiment with a wide range of data sharing services and infrastructure** to evaluate cost and impact before implementing these at scale/institution-wide.
- Research libraries can and do play an integral role in institutions' ability to meet requirements for public access to research data. **Data sharing services provided by research libraries may be underutilized and could benefit from increased campus outreach and messaging from funder representatives to seek these services out at their local institutions.**

We encourage other research institutions to conduct similar institutional investigations of data sharing support and researcher data sharing activities. Such studies are instrumental in identifying key support units, pinpointing service gaps in institutional research data

---

3 For instance, for institution strategies to make research data publicly available, see Tobin L. Smith, Kacy Redd, Sarah Nusser, Robert Samors, and Emily R. Miller, *AAU APLU Guide to Accelerate Public Access to Research Data*, Washington, DC: Association of American Universities and Association of Public and Land-grant Universities, 2021, <https://doi.org/10.31219/osf.io/tjybn>.

management and sharing support, and conducting targeted outreach to increase researchers' adoption of this support. These gap analyses not only offer a snapshot of the current state of local DMS support, but also serve as a critical foundation for enhancing the coordination of research data management support across institutions.

## Background and Research Aims

Federally mandated policies requiring public access to funded research data have impacted how funded researchers share their research data and how institutions support researchers in these efforts. Funding agencies' policies created in the wake of the 2013 White House Office of Science and Technology Policy (OSTP) Holdren Memo, "[Increasing Access to the Results of Federally Funded Scientific Research](#)," have transformed data sharing in the United States across academic disciplines over the last decade. Contributing to the upward trend of research data sharing are: increasing journal and publisher data availability policies,<sup>4</sup> internal institutional research data policies,<sup>5</sup> and an increasing recognition of the importance of research data sharing for public access from the scientific community.<sup>6</sup> Research data sharing requirements will only continue to increase, as many, if not all, US federal funding agencies will implement their own public access and/or data sharing policies in the coming years in response to the 2022 OSTP Nelson Memo, "[Ensuring Free, Immediate, and Equitable Access to Federally Funded Research](#)." This expansion will impact all federal agencies with extramural research and development (R&D) budgets, and, subsequently, the institutions and researchers funded by them.<sup>7</sup>

While funded researchers are primarily responsible for the stewardship of their research data, the institution itself is ultimately responsible for complying with the conditions of funded grants and contracts. Thus,

---

4       Iain Hrynaszkiewicz, Natasha Simons, Azhar Hussain, Rebecca Grant, and Simon Goudie, "Developing a Research Data Policy Framework for All Journals and Publishers," *Data Science Journal* 19, no. 1 (2020): 5, <https://doi.org/10.5334/dsj-2020-005>.

5       For more information, see Kristin Briney, Abigail Goben, and Lisa Zilinski, "Do You Have an Institutional Data Policy? A Review of the Current Landscape of Library Data Services and Institutional Data Policies," *Journal of Librarianship and Scholarly Communication* 3, no. 2 (2015): eP1232, <https://doi.org/10.7710/2162-3309.1232>.

6       For example, see Gabriel Popkin, "Data Sharing and How It Can Benefit Your Scientific Career," *Nature* 569, no. 7756 (2019): 445-447, <https://www.nature.com/articles/d41586-019-01506-x>.

7       Association of Research Libraries, "US Office of Science and Technology Policy (OSTP) 2013 & 2022 Public Access Memo Comparison," November 17, 2022, <https://www.arl.org/wp-content/uploads/2022/11/Table-Comparison-Office-of-Science-and-Technology-Policy-2.pdf>.



institutions often provide *infrastructure*<sup>8</sup> to meet funder requirements for data management and sharing (DMS). This same support also addresses DMS requirements from scholarly publishers or other stakeholders. Enabling these support mechanisms requires institutions to invest substantially in the development and maintenance of this data sharing infrastructure. Additionally, since no one institutional unit or group is wholly responsible for research data support services, holistic assessment and coordination of these services is vital.

Conducting an inventory of the institutional infrastructure that supports research data management and sharing is a useful approach towards assessing how well researchers are supported. Until institutions inventory their research data infrastructure, institutional leadership may not know where gaps or overlaps exist, and, therefore, where to strategically place investments. These investments help institutions and their researchers treat research data as an asset, ultimately facilitating compliance with funder mandates and promoting good and open science within their research communities. It is additionally useful for institutions to understand how their funded researchers make data publicly available, and if these researchers are taking advantage of services provided within the institution or externally. With this information, institutions can make informed and effective investments in further data sharing infrastructure.

To this end, completing an institutional infrastructure inventory, or scan, of research data services and researcher data sharing activities is an important first step in understanding the research data management and sharing efforts at institutions. We prioritized this inventorying process in the first stage of research of the Association of Research Libraries (ARL) [Realities of Academic Data Sharing \(RADS\) Initiative](#). The first stage of the RADS Initiative was funded by the US National Science

---

<sup>8</sup> The term “infrastructure” in this report, is used as a singular term to encompass all institutional efforts to support research data sharing and management activities, broadly speaking. This includes: technical infrastructure (such as institutional repository support); data governance, including the development, implementation, and oversight of data policies; one-time efforts or investments to accelerate services; and ongoing service operations.

Foundation ([NSF #2135874](#)),<sup>9</sup> and the six participating institutions included Cornell University, Duke University, University of Michigan, University of Minnesota, Virginia Tech, and Washington University in St. Louis. This report provides the results of data management and sharing institutional services investigations at the aforementioned six institutions, and highlights where service gaps may exist between funded researchers' needs and what institutions provide.

---

9 Providing a scan of institutional data sharing activities was not the only goal in phase one of the RADS Initiative research, which also included gathering information about public access to data sharing expenses, and assessing shared research data metadata quality. For additional project goals and findings, visit the RADS website, "Realities of Academic Data Sharing (RADS) Initiative," Association of Research Libraries, accessed March 3, 2024, <https://www.arl.org/realities-of-academic-data-sharing-rads-initiative/>.

## Reported Institutional Data Sharing Support

As described in the Realities of Academic Data Sharing methodology report,<sup>10</sup> administrators at our six institutions were surveyed between October and December 2022, and asked if their units supported data management and sharing (DMS) activities between 2013 and the present day (Yes or No).<sup>11</sup> Across our six institutions, 138 administrators were identified as those working in units to support research data sharing services and infrastructure. From this participant pool, 69 administrators responded, resulting in a 50% response rate.

In this survey, administrators were specifically asked if their unit provided services or infrastructure across five broad research data life cycle phases:

- Planning, Design, and Start Up of Projects
- Data Collection, Storage, and Management
- Making Data Broadly Available
- Data Retention
- Project Closeout

Based on the phases selected, administrators were then asked if their unit supported corresponding DMS activities, for up to 27 DMS activities total.<sup>12</sup> Responding units were then categorized into one of four service categories (refer to Appendix A), and then visually represented in Tableau (Visualizations 1–4). These four service categories are:

---

<sup>10</sup> See Shawna Taylor, Alicia Hofelich Mohr, Jonathan Petters, Jake Carlson, Lizhao Ge, Joel Herndon, Wendy Kozlowski, Jennifer Moore, and Cynthia Hudson Vitale, *Realities of Academic Data Sharing (RADS) Initiative: Research Methodology 2022–2023 Surveys and Interviews*, Washington, DC: Association of Research Libraries, January 2024, <https://doi.org/10.29242/report.radsmethodology2023>.

<sup>11</sup> The starting year for research, 2013, was selected due to the release of the OSTP Holdren memo and to understand how support activities have developed since then. Survey participants reported on current (2022) activities their unit supports.

<sup>12</sup> For a full list of the activities and their corresponding phases, see Shawna Taylor, with Jake Carlson, Joel Herndon, Alicia Hofelich Mohr, Wendy Kozlowski, Jennifer Moore, Jonathan Petters, and Cynthia Hudson Vitale, *Public Access Data Management and Sharing Activities for Academic Administration and Researchers*, Washington, DC: Association of Research Libraries, November 2022, <https://doi.org/10.29242/report.rads2022>.

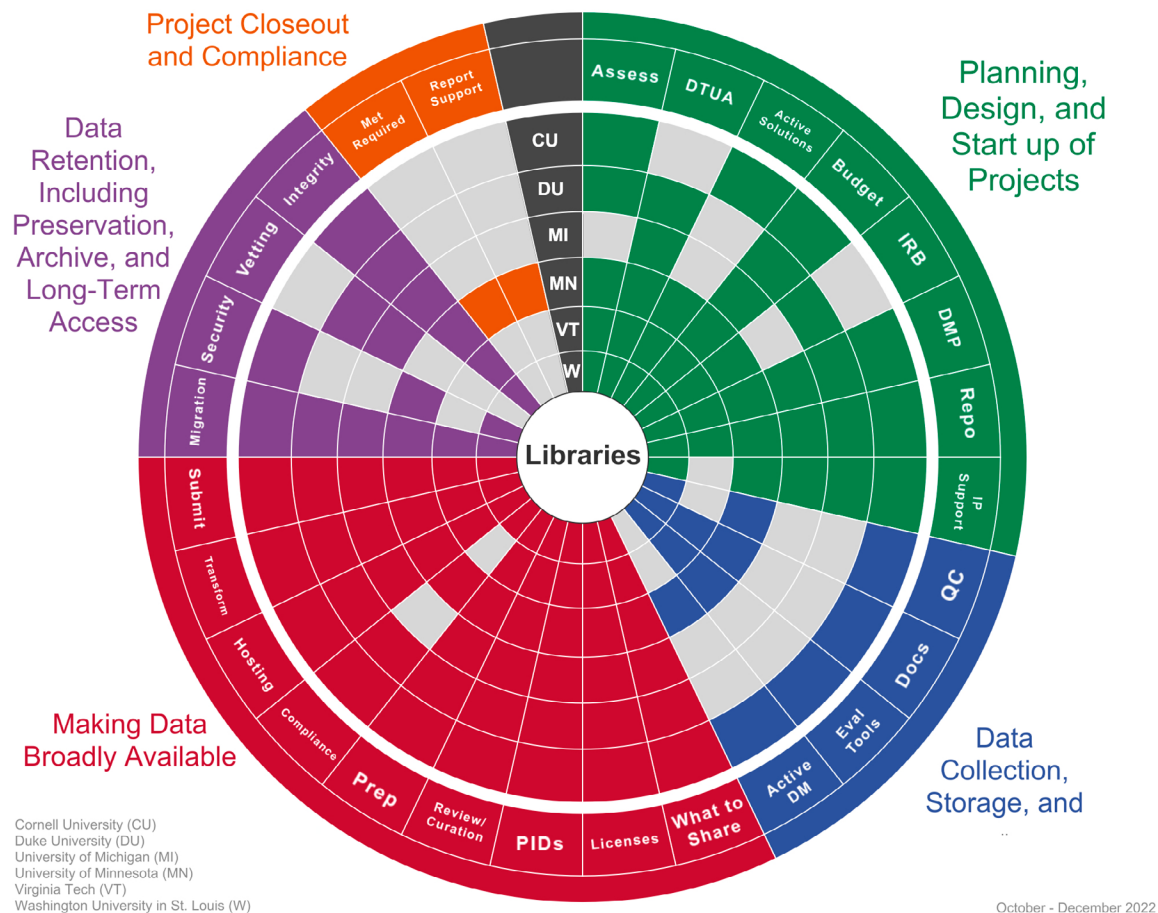
- Campus libraries and archives (Libraries)
- Information technology offices, including sub-units (Information Technology)
- Administrative research offices such as the Office of Research, the Office of the Vice Chancellor for Research, and Sponsored Projects Offices (Research Offices)
- Academic institutes, research centers, departments, or other specialized service areas (Institutes & Research Centers)

We chose to group responding offices into these four broad categories to enable a mechanism for comparison across the six RADS institutions, all of which house different departments and/or offices to support researchers in data management and sharing activities, and provide varying levels of service. Instead of attempting comparisons across institutions at an apples-to-oranges level, the aim of classifying the offices into these four broad categories is to enable apples-to-apples comparisons, as academic institutions typically have similar higher-level administrative units even if the scale differs across institutions.

The visualizations below (Visualizations 1–4) represent reported engagement in and support of the 27 DMS activities across the six RADS institutions. Note that services may exist within an institution for a particular data sharing activity where that service unit did not respond to the survey, or even in offices that were not identified by our institutional scans. Furthermore, services seen in the visualizations may not be offered to all researchers at an institution (e.g., may only be available within a college or research institute). For more detail, including information on which institutional offices support a particular activity, view the online versions of the visualizations and hover over each rectangle for the full activity name and the supporting office (use the link for each visualization title). See Appendix B for the list of Tableau labels with their corresponding activity.

## Visualization 1: Data management and sharing activities supported by libraries across the six RADS institutions

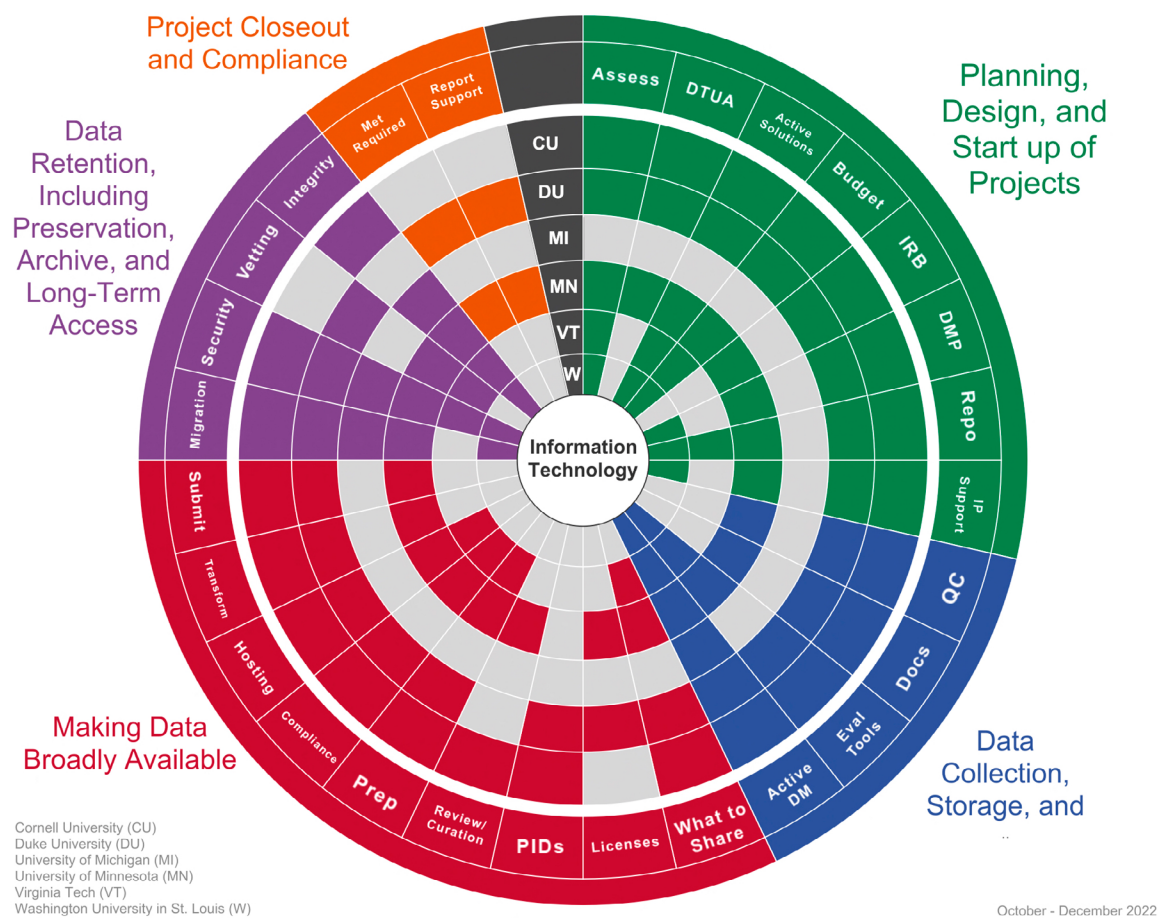
This visualization only shows support of DMS activities from library administrators who responded to the RADS Institutional Infrastructure Survey, and may not include support that exists at a particular institution.



From the Libraries visualization (Visualization 1), we can clearly see that libraries support data sharing activities in some capacity throughout the entire research life cycle. As key partners at academic institutions, libraries may offer services supporting upwards of two dozen activities to enable data sharing, including: consulting on data use agreements, supporting intellectual property questions, fielding copyright considerations, managing active data, providing data curation

and preservation services, and making data broadly available through technical infrastructure, such as institutional repositories (or providing consultation on alternative repositories). In contrast, libraries showed the least amount of support in the Closeout and Compliance phase of the research life cycle; however, one of our six institutions (the University of Minnesota) indicated support in this area as well.

## **Visualization 2: Data management and sharing activities supported by information technology (IT) offices across the six RADS institutions**



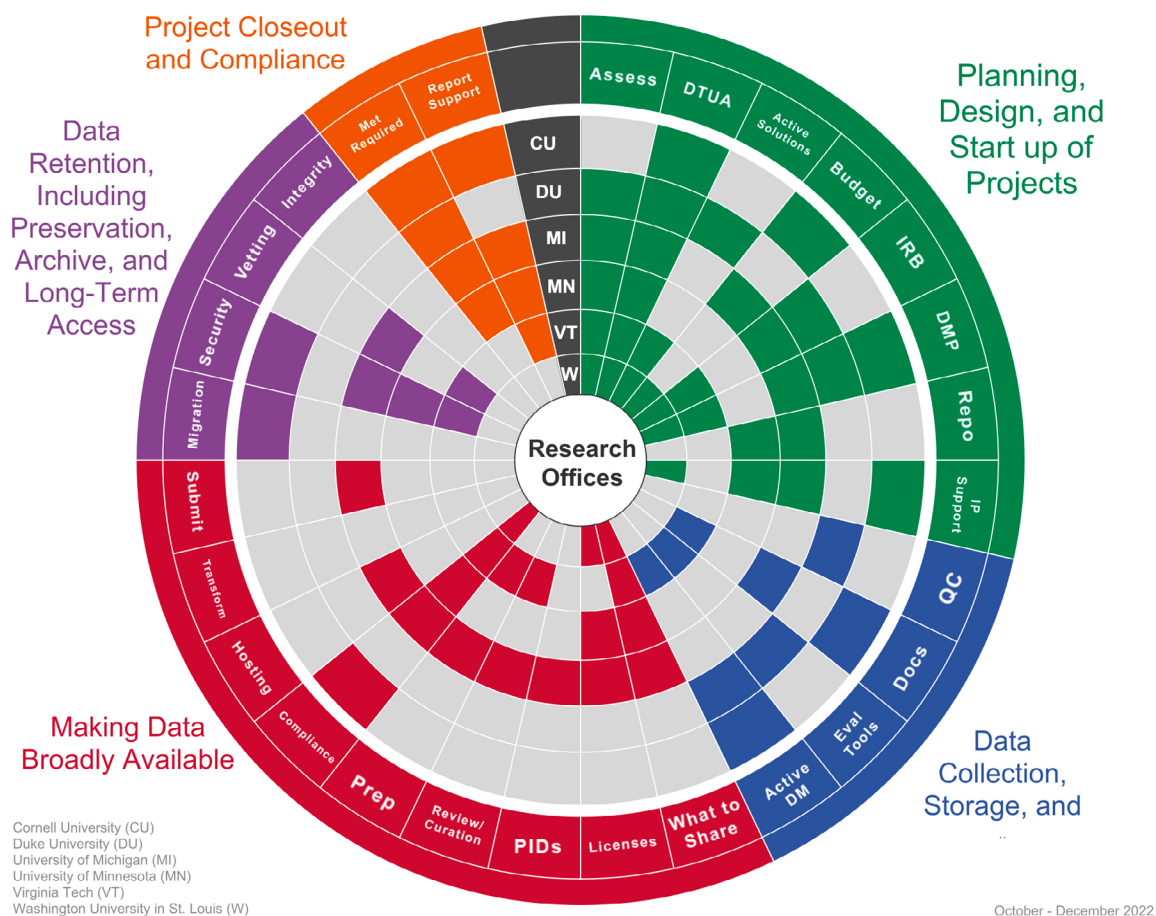
This visualization only shows support of DMS activities from IT administrators who responded to the RADS Institutional Infrastructure Survey, and may not include support that exists at a particular institution.



After libraries, information technology (IT) units across our six institutions supported the greatest number of DMS activities. IT generally supports research data sharing activities to some extent across all phases, but generally provides less support within Project Closeout and Compliance.

As seen in this visualization, there is substantial variability with what data sharing activities IT offices from each institution support. Once again, as described in our limitations below, we recognize our visualizations may not show all institutional support owing to lack of survey responses from units known to support data sharing activities.

**Visualization 3: Data management and sharing activities supported by administrative research offices (for example, Office of Research, Sponsored Projects Office) across the six RADS institutions**



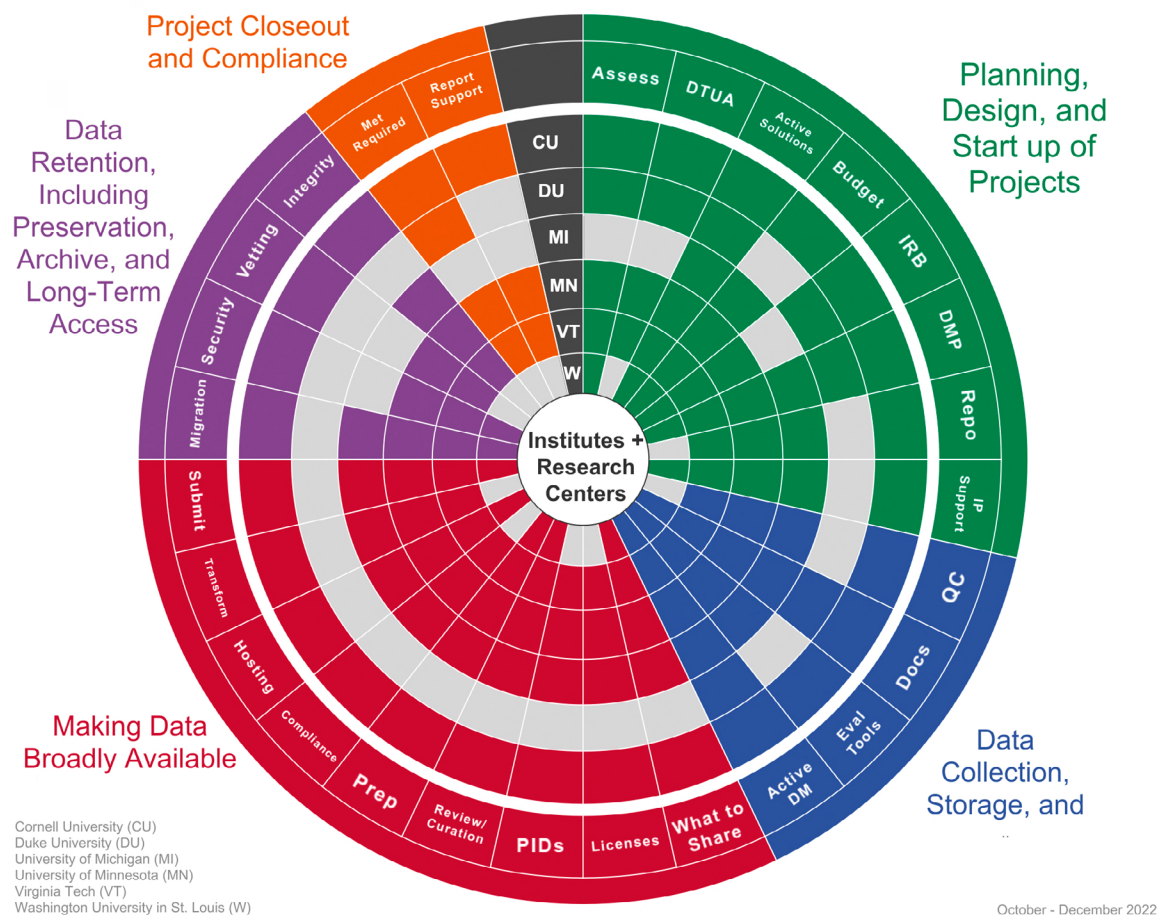
Central research offices typically reported offering the most support in the data sharing activity phases of Project Planning, Design, and Startup

and of Project Closeout and Compliance. However, the variability from institution to institution of data sharing activities supported by these research offices is striking. No research office reported supporting two activities “Monitoring integrity of preserved data” and “Migrating data file formats to be more open or accessible”, yet five or more institutions reported supporting three activities: “Evaluating data security needs,” “Developing Materials Transfer and/or Data Use Agreements,” and “Preparing Data Management Plans (DMPs) or Data Management and Sharing (DMS) Plans.” Every other activity is supported by one to four central institution research offices.



## Visualization 4: Data management and sharing activities supported by research institutes, academic departments, institutional centers, and related centers, across the six RADS institutions

This visualization only shows support of DMS activities from research institute and center administrators who responded to the RADS Institutional Infrastructure Survey, and may not include support that exists at a particular institution.



Three out of six of our institutions had responses from specialized service areas, such as institutional research centers, research institutes, academic departments, or related centers, indicating they supported all or nearly all of the 27 DMS activities across every phase of the research life cycle (refer to Institutes and Centers—Visualization 4). For

example, at Virginia Tech, the Virginia Tech Transportation Institute (VTTI) reported supporting all 27 activities. However, in a subsequent interview with VTTI administrators it was clarified that VTTI would only provide this support to its research affiliates, and only with external funding in most cases.

## Researcher-Reported Data Management and Sharing Activities

In addition to surveying institutional administrators, funded researchers at our six institutions were also surveyed between October and December 2022.<sup>13</sup> Five discipline areas were selected based on strong institutional research areas and on datasets submitted to the Data Curation Network (DCN) up until 2021.<sup>14</sup> These disciplines are: environmental science, materials science, psychology, biomedical sciences, and physics. Funders were then selected based on these discipline areas, overall awards amounts from funders at each of our six institutions, and publicly available grant information in funder award databases. As a result, researchers with awards from the Department of Energy (DOE), the National Institutes of Health (NIH), and the National Science Foundation (NSF), with awards between 2013 and 2022, in the five discipline areas above, were asked to complete the RADS Researcher Survey regarding grant-specific data management and sharing practices. Our participant pool included 3,467 possible participants, with 255 total responses, for an 8% response rate.

---

<sup>13</sup> For more detail on the methodology of the RADS funded-researcher investigative stream, see Taylor et al., *RADS Initiative: Research Methodology*.

<sup>14</sup> The members of the Data Curation Network (DCN) are part of a shared staffing model where, if there is no expertise or capacity at one institution to curate a particular type of dataset, datasets can be submitted to the entire DCN for curation. DCN data from 2019-2021 informed the selection of discipline areas. The [DCN data on datasets submitted to the DCN by discipline](#) reflect current shared curation data.

## Researcher Data Sharing Activities

The Researcher Survey included a similar set of data management and sharing activities presented in the Administrator Survey. The data sharing activities in the Researcher Survey were the same as the Institutional Infrastructure Survey, except for the actions taken (asking how they “did” rather than “supported” each activity) and the removal of one question, “Providing or hosting repositories for making data available.”<sup>15</sup> Researchers were asked which of these activities they: did **Themselves/within the Research Lab/with Research Team**; did with **Institutional Assistance**; did with **External Assistance**; or **Did Not Do**. Tables 1–4 show the top 10 aggregated (all institutions) researcher responses to each of these four options. In the survey, there was no “Not Applicable” option, so it should be assumed that where the activity did not apply, researchers selected “Did Not Do.” Percentages within Tables 1, 2, and 3 were computed using the total number of researchers who reported doing the activity.

Overall, researchers are doing the majority of data sharing activities on their own or within their research labs. The top 10 of these activities are reported in Table 1, and show that 90% of researchers/lab groups made their own decisions about what data to share (Activity #1.1), and prepared their own data for sharing (Activity #1.2). The rest of the activities on this top 10 list were handled by 85% or more of researchers and their groups on their own.

---

<sup>15</sup> For a comparison of the activities for administrators/research support vs. researchers, see Shawna Taylor, [“Realities of Academic Data Sharing \(RADS\) Initiative: Research Update #2—Activities for making Research Data Publicly Accessible,”](#) ARL Views (blog), November 28, 2022.

**Table 1: Top 10 data management and sharing activities funded-researchers reported as “Did Myself/within the Lab/with Research Team” during their award period**

<b>Activity Number</b>	<b>Activity</b>	<b>Percent Reported</b>
1.1	Making decisions about what data to share or host	91.8%
1.2	Preparing data for sharing (e.g., de-identification, selection, curation, data cleaning, validation, and quality control)	91.4%
1.3	Creating quality control mechanisms or procedures	90.4%
1.4	Developing documentation of data (e.g., data dictionary, protocols)	90.3%
1.5	Monitoring integrity of preserved data	90.0%
1.6	Evaluating data analysis tools and processes to support sharing and reproducibility	89.5%
1.7	Creating documentation for sharing (e.g., structured metadata, README files)	88.2%
1.8	Ensuring funding agency requirements for data sharing have been met	87.2%
1.9	Preparing Data Management Plans (DMPs) or Data Management and Sharing (DMS) Plans	84.8%
1.10	Migrating files to new formats or systems as needed	84.7%

In analyzing the data presented in Table 1, it’s beneficial to determine when it’s suitable for researchers to leverage centralized services (like those provided by central IT, the library, etc., as illustrated in Visualizations 1–3 above), as opposed to situations where specialized or discipline-specific expertise is necessary.

Reported activities done with institutional assistance, as reported in Table 2, show that surveyed researchers most heavily relied on their institutions for activities involving legal liability; for example, 64% selected development for materials transfer/data use agreements (Activity #2.1), and 40% selected data security related to legal requirements (Activity #2.2). For the next seven data sharing activities (#2.3–#2.9) surveyed researchers reported utilizing institutional assistance between 30% and 20% of the time. Notably, 20% of researchers used institutional support for activity #2.9, “Creating persistent identifiers (PIDs; e.g., DOIs).”

**Table 2: Top 10 data management and sharing activities funded-researchers reported as “Did with Institutional Assistance” during their award period**

Activity Number	Activity	Percent Reported
2.1	Developing Materials Transfer and/or Data Use Agreements	64.4%
2.2	Ensuring data security when appropriate (e.g., PHI/HIPAA, Export Controls, FISMA, student data, and Intellectual Property)	39.5%
2.3	Determining intellectual property and copyright considerations	29.9%
2.4	Evaluating data security needs	25.7%
2.5	Preparing IRB protocols and informed consent for data sharing	25.4%
2.6	Selecting or applying licenses for data reuse	22.2%
2.7	Checking for compliance with existing Data Use Agreements (DUAs)	21.7%
2.8	Compiling reports for project closeout	21.2%
2.9	Creating persistent identifiers (PIDs; e.g., DOIs)	20.0%
2.10	Identifying data management and sharing costs to be included in grant budgets	17.9%

Institutions do provide support for many of these and other research data sharing activities, and as previously described above, there are clear areas where researchers may want to leverage institutional support. Activities #2.3, #2.4, and #2.7, which are associated with legal requirements and data security, are typically supported by units in central research and IT offices, while libraries offer specific support for copyright considerations (Activity #2.3). Additionally, the research libraries of the six participating institutions (Visualization 1 above) provide support for activities #2.5, #2.6, and #2.9, with the central research office playing a more prominent role in IRB protocol development. The alignment of these support services with researcher needs is discussed below in the “Discussion” section.

Table 3 shows us that surveyed researchers used external support for data sharing activities the least frequently. Approximately a

quarter of the surveyed researchers received external assistance for, “Creating persistent identifiers (PIDs; e.g., DOIs)” (Activity #3.1), a service commonly provided by many institutional (data) repositories in their standard deposit and access services. Activity #3.2, “Checking for compliance with existing Data Use Agreements (DUAs),” was completed with external assistance by 10% of the surveyed researchers. The remaining activities in Table 3 were reported by less than 8% of surveyed researchers.

**Table 3: Top 10 data management and sharing activities funded-researchers reported as “Did with External Assistance” during their award period**

Activity Number	Activity	Percent Reported
3.1	Creating persistent identifiers (PIDs; e.g., DOIs)	25.9%
3.2	Checking for compliance with existing Data Use Agreements (DUAs)	10.1%
3.3	Selecting an appropriate repository (or repositories) for making research data broadly available	7.9%
3.4	Selecting or applying licenses for data reuse	7.4%
3.5	Migrating data file formats to be more open or accessible	5.6%
3.6	Creating documentation for sharing (e.g., structured metadata, README files)	5.0%
3.7	Ensuring data security when appropriate (e.g., PHI/HIPAA, Export Controls, FISMA, student data, and Intellectual Property)	4.9%
3.8	Identifying data management and sharing costs to be included in grant budgets	4.8%
3.9	Submitting data into a data sharing platform (e.g., institutional repository, generalist repository, disciplinary repository)	4.6%
3.10	Making decisions about de-accessioning and removal of research data	4.6%

Table 4 shows the top 10 activities researchers reported as “Did Not Do” during their grant period. Half of these activities (#4.2, #4.4, #4.5, #4.7, #4.10) may be considered as “not applicable” to a substantial subset of funded research projects. For instance, 61% of respondents selected “Did Not Do” for “Preparing IRB protocols and informed

consent for data sharing” (Activity #4.2), suggesting many of these projects did not involve human participants.

**Table 4: Top 10 data management and sharing activities funded-researchers reported as “Did Not Do” during their award period**

Activity Number	Activity	Percent Reported
4.1	Selecting or applying licenses for data reuse	66.3%
4.2	Preparing IRB protocols and informed consent for data sharing	61.1%
4.3	Making decisions about de-accessioning and removal of research data	58.6%
4.4	Checking for compliance with existing Data Use Agreements (DUAs)	56.6%
4.5	Ensuring data security when appropriate (e.g., PHI/HIPAA, Export Controls, FISMA, student data, and Intellectual Property)	48.7%
4.6	Identifying data management and sharing costs to be included in grant budgets	48.5%
4.7	Developing Materials Transfer and/or Data Use Agreements	46.6%
4.8	Creating persistent identifiers (PIDs; e.g., DOIs)	46.5%
4.9	Migrating data file formats to be more open or accessible	44.1%
4.10	Determining intellectual property and copyright considerations	40.5%

The other activities (#4.1, #4.3, #4.6, #4.8, #4.9) reported as “Did Not Do” may be considered opportune areas for researchers to take advantage of support and advocacy from their institutions in undertaking. Two-thirds of surveyed researchers reported not doing activity #4.1, “Selecting or applying licenses for data reuse,” yet this activity is relevant to all data made publicly available. Dataset creators should clearly consider how their data may or may not be used by others, along with any other ethical or legal issues surrounding the reuse of this data. Activity #4.6, “Identifying data management and sharing costs to be included in grant budgets,” will be a requirement from researchers by all federal funders after the policies resulting from



the 2022 OSTP Nelson Memo go into effect, and, as such, this activity should be done more frequently.

Furthermore, “Making decisions about de-accessioning and removal of research data” (Activity #4.3) and “Migrating data file formats to be more open or accessible” (Activity #4.9) are both critical curation-specific data management and sharing activities supported by research libraries. Researchers stand to benefit by utilizing services offered by libraries. Surprisingly, almost half of the researchers reported not “creating persistent identifiers (PIDs; e.g., DOIs)” (Activity #4.8). Yet, as we’ll discuss in the section below, creating PIDs is a crucial step for nearly all shared datasets, as they link research data, publications, other research outputs, and their creators and authors together.

## Researcher Data Sharing Mechanisms

In addition to asking about the specific data management and sharing activities, the RADS Researcher Survey also asked researchers to report if and how they share their data (not necessarily public-facing data sharing). Out of the 255 researchers who completed the survey, 178 (70%) reported sharing their data, while 77 (30%) reported not sharing their data. Of those who “did not” share their data, the most common reasons were: data will be shared with publication or data is not yet published/shared (32%), data sharing not required or applicable to the grant-type (27%), ownership restrictions (e.g., proprietary) (12%), and the restrictive cost of preparing and/or sharing data (7%).

Respondents who reported sharing their data were then asked how their research data was shared. This question allowed for multiple selections, asking researchers to select all that apply and/or provide a write-in response. From the 178 respondents who ‘did’ share their data, 282 responses were provided on where they shared their data. Table 5 below shows that sharing data as part of a publication was the most common data sharing mechanism (34%). However, when all the repository responses (disciplinary, institutional, generalist, and code repositories) were added together, data sharing in repositories account for 40% of the responses.

**Table 5: Mechanisms used by researchers to share their data, including type of technical infrastructure**

<b>Mechanism for Data Sharing</b>	<b>Count</b>	<b>Percent</b>
In a paper or as supplemental materials with a publication	96	34.0%
Disciplinary repository	51	18.1%
Shared only on request	49	17.4%
Institutional repository	32	11.4%
Generalist repository	26	9.2%
Personal website	18	6.4%
Not a repository	4	1.4%

<b>Mechanism for Data Sharing</b>	<b>Count</b>	<b>Percent</b>
Code repository	3	1.1%
Unclear	3	1.1%

Furthermore, from survey write-in responses, 53 unique repositories were identified, 13 of which are NIH repositories or NIH-supported repositories. The high reporting of NIH repositories is likely due to the discipline areas of the grants we surveyed, with biomedical and psychology as two fields that would use NIH repository infrastructure.

Finally, researchers were asked to rank the top five factors that influenced their decisions on where to share their research data. These responses, shown in Table 6, show that cost and convenience were the most important factors, followed by external institutional recommendations, such as funder, publisher, or peer recommendations. The factors with the least influence on researcher data sharing decisions were internal campus recommendations, such as library, research office, or IT recommendations. See Appendix C for a graph with additional response details and full rankings.

**Table 6: Ranked factors with the greatest influence on researchers' decisions on where to share their research data**

<b>Factors—Researcher Decisions on Where to Share Data</b>	<b>Count—Ranked in the Top Five Factors</b>	<b>Percent</b>
Easiest/quickest option	120	17.0%
Personal experience	114	16.0%
Least expensive option	94	13.3%
Funder recommendation	74	10.5%
Journal/publisher recommendation	73	10.3%
Peer or colleague recommendation	62	8.8%

<b>Factors—Researcher Decisions on Where to Share Data</b>	<b>Count—Ranked in the Top Five Factors</b>	<b>Percent</b>
The option to personally control and manage content	59	8.3%
Disciplinary/scholarly society recommendation	49	6.9%
Other	24	3.4%
Library recommendation	18	2.5%
Research office recommendation	16	2.3%
Campus IT recommendation	5	0.7%

## Discussion: Service Gaps and Institutional Opportunities

Our findings show that research libraries and IT offices offer the broadest support, in terms of service, infrastructure, and personnel support, for data sharing activities (Visualizations 1–4). Although RADS research is retrospective, these findings align with COGR’s May 2023 prospective report, “[Data Management and Sharing \(DMS\) and the Cost of Compliance](#).”<sup>16</sup> COGR found that campus libraries and IT offices provide the largest share of institutional support in implementing the 2023 NIH Data Management and Sharing Policy at the administrative level, and thus they incur the highest costs for compliance at the administrative level.

However, the researchers surveyed did not rely on either internal or external support towards many data sharing activities. In 20 of the total 26 activities, 70% or more researchers reported that they conducted these activities themselves or within their research group. By integrating our visualizations of existing institutional support for data sharing activities with researchers’ responses on which data sharing activities they do by themselves or with institutional/external support, and where they are sharing their research data (Tables 1–6), we can identify opportunities that are ripe for further institutional support. In doing so, we should keep in mind that, while the activities given to both researchers and administrators are almost identical, each group may conceive of the activities differently.

Research data outputs are different for each research project and clearly there are activities best-suited for the researchers to do themselves or within their research labs. These activities include creating the documentation associated with their data or research output (Activity

---

<sup>16</sup> COGR, “Data Management and Sharing (DMS) and the Cost of Compliance: Results from the COGR Survey on the Cost of Complying with the New NIH DMS Policy,” May 11, 2023, [https://www.cogr.edu/sites/default/files/DMS\\_Cost\\_of\\_Compl\\_May11\\_2023\\_FINAL%20%281%29.pdf](https://www.cogr.edu/sites/default/files/DMS_Cost_of_Compl_May11_2023_FINAL%20%281%29.pdf).

#1.4) and making decisions about what data to share (Activity #1.1). Although services exist within research libraries to provide high-level guidance for these activities, researchers are ultimately responsible for these activities as they understand the nuance and complexity of their data.

## **Opportunities for IT Departments**

Information technology (IT) units in five out of six of our participating institutions indicated providing assistance for “ensuring data security when appropriate” in the Data Retention phase (Visualization 2), whereas nearly half of responding researchers reported not doing this activity (#4.5). Although researchers were not asked to specify why they did not do each activity, we can presume that many in this case did not have research data security concerns. However, of the researchers who reported they “[ensured] data security when appropriate,” only 39.5% reported utilizing institutional assistance (Activity #2.2), and 4.9% percent sought external assistance (Activity #3.7). Taken together, these percentages indicate that more researchers could benefit from discussion with IT personnel regarding research data security. Intentional collaborations between IT personnel and researchers is one strategy institutions could use to lower the risk of inappropriate data sharing (e.g., breaches of human participant confidentiality).

On the other hand, 90% of researchers reported “Creating quality control mechanisms or procedures” (Activity #1.3) as an activity they did themselves or within their research groups. Visualization 2 shows us that three out of six participating institutions provide support for this activity through their IT offices, suggesting that there are further opportunities for researchers to take advantage of IT support in creating quality control mechanisms or procedures, especially if the researcher’s work involves complex technical infrastructure.

## **Opportunities for Central Research Offices**

Units under central research offices in four out of six of our participating institutions reported providing support for “ensuring funding agency requirements for data sharing have been met” (Visualization 3), whereas 87% of researchers reported conducting this activity themselves or within their labs (Activity #1.8). Taken together, these results suggest that researchers were either unaware of this institutional support for project closeout, or were aware of it but did not find this support was meeting their needs. This may be of particular importance in the coming years, as award closeout activities at institutions will include the data and publication sharing requirements mandated by the 2022 Nelson Memo.

## **Opportunities for Research Institutes and Specialized Centers**

Somewhat surprisingly, research institutes and specialized centers at three institutions reported offering support for all 27 data management and sharing activities, across the research life cycle. Other specialized centers reported support for a majority of the activities (Visualization 4). However, this support is generally only available to affiliates of the institutes or centers, and at least one of our subsequent interviews found that the extent of this support depends on the availability of extramural funding within the project. Understanding the distribution of services offered by these specialized areas may give administrators insight when considering how to support data-sharing efforts at their respective institutions. Research institutes and specialized centers may also provide an opportunity for institutions to experiment with a wide range of data sharing efforts and evaluate cost and impact before implementing these services at scale/campus-wide.

## **Opportunities for Libraries**

From Visualization 1, we see that all six participating institutions have research libraries that provide support for “selecting or applying

licenses for data reuse,” which 66% of researchers reported not doing (Activity #4.1). The provision of licenses for data reuse is important for all researchers, as it helps future users understand how the data can and should be used in accordance with the interests of the data creator, institution, and possibly the funder. Also, 22% percent of researchers reported that when they do select or apply for a license for data reuse, they do it with institutional assistance (Activity #2.6), and 7% reported doing this activity with external assistance (Activity #3.4). This indicates that libraries could increase researcher awareness of these support services. Furthermore, libraries should consider forming partnerships with research office units or central units offering intellectual property support, as they typically address licensing and intellectual property concerns associated with research data.

Research libraries have established support services for several of the data sharing activities (Visualization 1) that a majority of surveyed researchers reported doing themselves or within their research groups. For example, five out of six surveyed research library administrators indicated providing assistance with “Making decisions about what data to share or host” and “Preparing data for sharing,” the top two data sharing activities (#1.1 and #1.2) that surveyed researchers reported doing themselves.

Researchers might also benefit from library support in preparing data management plans (Activity #1.9). Surveyed researchers reported preparing these plans themselves at a rate of 85% (Activity #1.9). Additionally, research libraries from five of our six institutions provide assistance in selecting an appropriate repository (or repositories) for making research data broadly available (Activity #3.3). The majority of researchers did not report doing this activity themselves or with institutional assistance, and only 8% of researchers reported doing this activity with external assistance.



Furthermore, research libraries often provide repository services that can help researchers get persistent identifiers (PIDs, e.g., DOIs, RORs, ORCIDs) for their publicly available datasets (Activity #4.8). PID assignment is a crucial data sharing activity, as PIDs enable linkages between dataset metadata, publications, other research products and outputs, authors, and institutional affiliation. When internal repositories are not used, library personnel may offer consultation to researchers in finding appropriate external repositories, and how to obtain PIDs from external repositories. Of those researchers who do assign PIDs to their shared data (Activity #4.8), only 20% reported doing so with internal assistance (Activity #2.9). These low percentages suggest that while some researchers have taken advantage of these services, clearly there is an opportunity for more researchers to do so. Senior leaders in libraries should collaborate with leadership in other institutional research-support offices to develop strategies for wider PID adoption in shared research products, including data.

These data sharing services provided by research libraries may thus be underutilized and could benefit from (a) increased institutional resourcing towards building researcher awareness of these services and/or (b) messaging from funder representatives to seek these services out at their local institutions.

## **Institution-Wide Opportunities**

There is a clear opportunity for units across the institution to coordinate support for data de-accessioning. Over half of researchers (59%) reported not “making decisions about de-accessioning and removal of research data” (Activity #4.3). Considering this result with the units that support DMS activities, represented in Visualizations 1–4, we see that planning for data de-accessioning is supported by more than one unit. One third of participating research libraries provide such support, and half of IT offices report that they provide this support. Developing guidelines or policies for this activity, in collaboration with libraries, IT, specialized institutes, and central research offices, can result in tangible benefits to the institution and its

researchers. Coordinated de-accessioning guidelines could facilitate the following: resource optimization, as storage will be freed up and costs reduced; data integrity, as keeping outdated or irrelevant data in long-term storage can lead to mistakes in research, decision-making, and reporting; and reduced risk in terms of compliance around human participants or other sensitive data.

In addition to the reported 59% of researchers not “making decisions about de-accessioning and removal of research data” (Activity #4.3), 44% of researchers also reported not “migrating data file formats to be more open or accessible” (Activity #4.9). Migrating to open data formats is in alignment with the FAIR (Findable, Accessible, Interoperable, and Reusable) Principles,<sup>17</sup> as open formats ensure that data can be used in different technological environments over a longer period of time. Open formats enhance the interoperability of data and, with machine-actionable connections increasing within the research ecosystem, migrating data to open formats will become an essential step in data sharing practices. Finally, open and accessible data formats make it easier for other researchers to reuse and analyze data for replication of research findings as well as for novel purposes.

Furthermore, nearly half of researchers (49%) reported not “identifying data management and sharing costs” in the planning phase of their grant budgets (Activity #4.6). Of those who did do this activity, only a quarter reported doing so with internal or external assistance (Activities #2.10 and #3.8). As shown in Visualizations 1–4, research libraries at five of six participating institutions reported offering support for this activity, with the other three service areas (IT, central research offices, and specialized institutes and centers) offering support for this activity in at least half of our participating institutions. Making data publicly available incurs costs,<sup>18</sup> which may be budgeted

---

17 GO FAIR, “FAIR Principles,” accessed March 5, 2024, <https://www.go-fair.org/fair-principles/>.

18 For project results on the costs to share research data, see Alicia Hofelich Mohr, Jake Carlson, Lizhao Ge, Joel Herndon, Wendy Kozlowski, Jennifer Moore, Jonathan Petters, Shawna Taylor, and Cynthia Hudson Vitale, *Making Research Data Publicly Accessible: Estimates of Institutional & Researcher Expenses*, Washington, DC: Association of Research Libraries, February

for directly by the researcher (e.g., migrating data file formats to be more open or accessible) either in terms of time or infrastructure, or absorbed indirectly by the institution (e.g., supporting an institutional repository). When data sharing costs are not accounted for in the planning phase, it is less likely budgets will have room to cover the successful implementation of late-stage activities. Budgeting for these expenses should aid researchers in appropriately planning for resource allocation towards public data sharing in their projects, and thus increase the chances of achieving effective public data sharing results.

In the RADS expense report, *Making Research Data Publicly Accessible: Estimates of Institutional & Researcher Expenses*, we report that the “DMS expense for a researcher who shared some or all of their data using an institutional repository or institutional data repository averaged \$7,200 compared with \$35,000 for those researchers who used [sharing] methods other than an institutional repository.”<sup>19</sup> Additionally, researchers who completed more DMS activities on their own, as opposed to with the support of their institutions or external repositories, had higher overall DMS costs.<sup>20</sup> Given that institutions are already investing in DMS resources and services, and this investment may offset costs to individual researchers, units across the institution may want to encourage researchers to use local infrastructure and services when appropriate.

As an important aside, all data sharing support units (e.g., research libraries, IT, units under central research offices, and specialized institutes and centers) provide education and training on some of these data sharing activities. These training efforts may be a contributing factor as to why researchers are doing some activities themselves (Table 1) and not seeking direct institutional support. For example, research libraries have provided data management planning training for many years, and some researchers have learned to plan for effective data management without further assistance. We should not assume that just because researchers are conducting data sharing activities

---

2024, <https://doi.org/10.29242/report.radsexpense2024>.

19 Hofelich Mohr et al., *Making Research Data Publicly Accessible: Estimates*, 21–22.

20 Hofelich Mohr et al., *Making Research Data Publicly Accessible: Estimates*, 20.

themselves, it does not mean that they do not require further support, and an increased investment in education and training of researchers in these activities is a laudable strategy.

## **Coordinating Efforts to Increase Data Sharing Efficiencies**

Establishing institution-wide groups, such as [Duke University's Research Data Initiative & Research Data Policy](#) or [Cornell University's Research Data Management Support Group](#), is essential for effectively creating data sharing services efficiencies across the institution. As discussed in the previous section “Institution-Wide Opportunities,” many DMS activities are not just supported by one group (refer to Visualizations 1–4). Institution-wide groups can serve as a coordination mechanism for such activities and lead to streamlining workflows and eliminating redundancies.<sup>21</sup> If institutions have not yet established these institution-wide groups, they are encouraged to do so. Libraries should play a key role in these groups.<sup>22</sup> Moreover, institutions are already investing in internal research data management and sharing resources and services like institutional repositories, which can reduce overall expenses and increase resource use efficiency for individual researchers.<sup>23</sup>

---

21 For more information, see Part 2, “Making Priorities Visible and Establishing a Plan,” of Smith et al., *AAU APLU Guide*.

22 Smith et al., *AAU APLU Guide*, 11.

23 Hofelich Mohr et al., *Making Research Data Publicly Accessible: Estimates*, 24.

## Limitations

The following are known limitations of the Institutional Infrastructure Survey for administrators and the Researcher Survey for funded researchers that should be kept in mind when interpreting results:

- The RADS project team identified administrative units thought to support data sharing activities at our respective institutions. However, we may have overlooked some units in our institutional assessments of service providers. Any units overlooked would not be represented in our data.
- When developing the administrator participant pool, determining at what level to survey administrative offices was occasionally challenging (e.g., identifying which units had their own budgets within the Office of the Vice Provost for Research), and unit levels varied between institutions.
- Known units/departments that support data sharing did not respond to the survey; therefore, we know our data, including the resulting interactive visualizations, are incomplete.
- Question 6 asked administrators which data sharing phases their unit/department supports or offers services for, and these phases were listed with sample activities as examples. Based on the phase selections in question 6, respondents were then only asked about the full list of activities from that phase (questions 7–11). As a result, units/departments may support activities not shown to them on the questionnaire. This is a possible area of underreporting in our data.
- Researchers who reported on awards from early in our nine-year timeframe (e.g., 2013–2017) may have underreported which data sharing activities they conducted. Surveying researchers on recently completed projects would yield more accurate results.

## Assessing the Institutional Landscape Locally

The results and discussion shown in this report so far are those aggregated from the six participating institutions. As a whole, they may be considered a reasonable sample of trends in researchers' data sharing activities and institutional support for public data sharing at R1 institutions. However, to best match institutional challenges in public data sharing support with the needs of researchers for a particular institution, the interpretation and analysis of results should be conducted with information specific to that institution only.

### Version 3 of the RADS Public Access Data Management and Sharing (DMS) Activities

Our project team encourages institutions to make use of the RADS Public-Access DMS Activities ([refer v3 of the RADS DMS Activities](#) for the most recent version, released in December 2023). These activities are broad in scope, and administrators can utilize them to conduct an institutional environmental scan on research data management and sharing services at their institutions. A scan based around these activities can aid administrators in identifying areas for service collaborations, and in identifying support gaps and opportunities. This will help institutions maximize their resources by amplifying current services, removing duplication of effort, and providing end-to-end support for researchers in managing and sharing their data.<sup>24</sup> Researchers at individual institutions can also make use of these activities and assess which activities they are either not doing during their projects that should be done, or identify activities that are being done within their research teams that could benefit from local support opportunities offered by campus service providers.<sup>25</sup>

---

24 Shawna Taylor and Mikala Narlock, "Navigating the Complex Landscape of Research Data Management and Sharing (DMS): DMS Activities from the RADS Initiative," ARL Views (blog), December 19, 2023, <https://www.arl.org/blog/navigating-the-complex-landscape-of-research-data-management-and-sharing-dms-dms-activities-from-the-rads-initiative/>.

25 Taylor and Narlock, "Navigating the Complex Landscape."

## Conclusion and Recommendations

From our analysis of our six institutions' reported researcher data sharing activities and institutional support for these activities, we determined that **researchers should take further advantage of institutional support for many data management and sharing activities**. For many of these activities such institution support exists in one or more units. Specifically, these activities are:

- Preparing data management plans (DMPs) or data management and sharing plans (DMSPs)
- Developing materials transfer agreements (MTAs) and/or data use agreements (DUAs)
- Identifying data management and sharing costs to be included in grant budgets
- Ensuring data security when appropriate
- Creating quality control mechanisms or procedures
- Making decisions about what data to share or host
- Creating persistent identifiers (e.g., DOIs)
- Selecting or applying licenses for data reuse
- Preparing data for sharing (e.g., de-identification, selection, curation, data cleaning, validation, and quality control)
- Making decisions about de-accessioning and removal of research data
- Ensuring funding agency requirements for data sharing have been met

Nearly half of researchers reported not identifying data management and sharing costs in the planning phase of their grant budgets. Making data publicly available comes at a cost, and appropriately planning for resource allocation increases the chances of effective public data sharing results. Some of these activities that researchers are doing themselves or within their research group would benefit from coordinated support across institutional units. Institutions should

consider forming cross-campus research data management working groups or committees to coordinate their support of public data sharing. Research libraries are well positioned to lead such coordination activities.

Our findings show that research libraries can and do play an integral role in institutions' ability to meet public access to research data requirements, especially in the phases of (1) Planning, Design, and Start Up of Projects; (2) Data Retention, Including Preservation, Archive, and Long-Term Access; and (3) Making Data Broadly Available. Data sharing services provided by research libraries may be underutilized and could benefit from increased campus outreach and messaging from funder representatives to seek these services out at their local institutions.

Additionally, administrators from research institutes and specialized centers in half of the institutions participating in this survey reported offering support for all 27 data management and sharing activities, across the research life cycle. Research institutes and specialized centers can provide an opportunity for institutions to experiment with a wide range of data sharing efforts and evaluate cost and impact before implementing these services at scale/campus-wide.

In light of these insights, we encourage other research organizations to conduct similar institutional investigations and coordination of data sharing support and researcher data sharing activities in order to create institutional efficiencies. Such studies are instrumental in identifying key support units, pinpointing service gaps in institutional research data management and sharing support, and conducting targeted outreach to increase researchers' adoption of this support. These gap analyses not only offer a snapshot of the current state of local data management and sharing support, but also serve as a critical foundation for enhancing the coordination of research data management support across institutions.



## Research Instruments

Research instruments used for data collection are:

- [Institutional Infrastructure \(Administrator\) Survey](#)
- [Researcher Perspectives Survey](#)

## Data Availability Statement

Primary data used in this paper are from the RADS Researcher and Institutional Infrastructure Surveys. De-identified response data and data dictionaries for both of these surveys are located in the Washington University in St. Louis WashU Research Data (WURD) repository, at <https://doi.org/10.7936/6RXS-103654>.

## Acknowledgements

The authors of this report would like to thank Mollie Webb, data and GIS developer, University Libraries at Washington University in St. Louis, for the development of the Tableau visualizations. We would also like to thank Lisa Johnston, director, Data Governance at the University of Wisconsin–Madison, and member of the RADS metadata team for her conceptual design of the data visualizations. The authors would also like to thank Mikala Narlock, director of the Data Curation Network, for her review of this report, as well as Kaylyn Groves, director of Member Communications, and Angela Pappalardo, senior program manager, Information and Services, both at the Association of Research Libraries, for their editorial contributions to this report. Finally, we are most appreciative to the National Science Foundation, with special thanks to Martin Halbert, science advisor for public access, and to all the participants in our studies, without whom this work would not be possible.

## For Further Reading

Association of Research Libraries. “US Office of Science and Technology Policy (OSTP) 2013 & 2022 Public Access Memo Comparison.” November 17, 2022. <https://www.arl.org/wp-content/uploads/2022/11/Table-Comparison-Office-of-Science-and-Technology-Policy-2.pdf>.

Briney, Kristin, Abigail Goben, and Lisa Zilinski. “Do You Have an Institutional Data Policy? A Review of the Current Landscape of Library Data Services and Institutional Data Policies.” *Journal of Librarianship and Scholarly Communication* 3, no. 2 (2015): eP1232. <https://doi.org/10.7710/2162-3309.1232>.

Hanisch, Robert J., Debra L. Kaiser, Alda Yuan, Andrea Medina-Smith, Bonnie C. Carroll, and Eva M. Campo. NIST Research Data Framework (RDaF). Version 1.5. NIST Special Publication 1500 (NIST SP 1500-18r1). Gaithersburg, MD: National Institute of Standards and Technology, May 2023. <https://doi.org/10.6028/NIST.SP.1500-18r1>.

Note: At the time of publication, the State University of New York (SUNY) released an interactive tool to assist institutions with using the RDaF: <https://github.com/SUNYORIED/rdaf>; <https://sunyoried.github.io/rdaf/>.

Hofelich Mohr, Alicia, Jake Carslon, Lizhao Ge, Joel Herndon, Wendy Kozlowski, Jennifer Moore, Jonathan Petters, Shawna Taylor, and Cynthia Hudson Vitale. *Making Research Data Publicly Accessible: Estimates of Institutional & Researcher Expenses*. Washington, DC: Association of Research Libraries, February 2024. <https://doi.org/10.29242/report.radsexpense2024>.

Kozlowski, Wendy, Joel Herndon, Shawna Taylor, Cynthia Hudson Vitale, Jake Carlson, Alicia Hofelich Mohr, Jennifer Moore, Mikala Narlock, and Jonathan Petters. “Realities of Academic Data Sharing (RADS) Initiative Public Access Data Management and

Sharing (DMS) Activities,” v3. Washington, DC: Association of Research Libraries, November 6, 2023. <https://doi.org/10.29242/radsdmsactivities2023>.

Popkin, Gabriel. “Data Sharing and How It Can Benefit Your Scientific Career.” *Nature* 569, no. 7756 (2019): 445-447. <https://www.nature.com/articles/d41586-019-01506-x>.

“Realities of Academic Data Sharing (RADS) Initiative.” Association of Research Libraries. Accessed March 3, 2024. <https://www.arl.org/realities-of-academic-data-sharing-rads-initiative/>.

Smith, Tobin L., Kacy Redd, Sarah Nusser, Robert Samors, and Emily R. Miller. AAU APLU Guide to Accelerate Public Access to Research Data. Washington, DC: Association of American Universities and Association of Public and Land-grant Universities, 2021. <https://doi.org/10.31219/osf.io/tjybn>.

Steinhart, Gail, and Katherine Skinner. “The Cost and Price of Public Access to Research Data: A Synthesis.” *Invest in Open Infrastructure*, February 29, 2024. <https://doi.org/10.5281/zenodo.10729575>.

Taylor, Shawna. “Realities of Academic Data Sharing (RADS) Initiative: Research Update #2—Activities for Making Research Data Publicly Accessible.” *ARL Views* (blog). November 28, 2022. <https://www.arl.org/blog/realities-of-academic-data-sharing-rads-initiative-research-update-2-activities-for-making-research-data-publicly-accessible/>.

Taylor, Shawna, Alicia Hofelich Mohr, Jonathan Petters, Jake Carlson, Lizhao Ge, Joel Herndon, Wendy Kozlowski, Jennifer Moore, and Cynthia Hudson Vitale. *Realities of Academic Data Sharing (RADS) Initiative: Research Methodology 2022–2023 Surveys and Interviews*. Washington, DC: Association of Research Libraries, January 2024. <https://doi.org/10.29242/report.radsmethodology2023>.

Taylor, Shawna, and Mikala Narlock. “Navigating the Complex Landscape of Research Data Management and Sharing (DMS): DMS Activities from the RADS Initiative.” ARL Views (blog). December 19, 2023. <https://www.arl.org/blog/navigating-the-complex-landscape-of-research-data-management-and-sharing-dms-dms-activities-from-the-rads-initiative/>.

Taylor, Shawna, with Jake Carlson, Joel Herndon, Alicia Hofelich Mohr, Wendy Kozlowski, Jennifer Moore, Jonathan Petters, and Cynthia Hudson Vitale. Public Access Data Management and Sharing Activities for Academic Administration and Researchers. Washington, DC: Association of Research Libraries, November 2022. <https://doi.org/10.29242/report.rads2022>.

## Appendix A: Administrative Unit Categorization

The following is a list of all responding units/departments from the Institutional Infrastructure Survey, and their categorization, used in the [project Tableau visualizations](#).

Institution	Responsive Department/Office	Service-Area Categorization
Cornell University	Center for Advanced Computing	IT
	Center for Technology Licensing	RSCH
	College of Engineering/Bowers College of Computing and Information Science (CIS)/Tech/IT Service Group (ITSG)	RSCH
	College Research Office	RSCH
	Cornell Center for Materials Research (CCMR)	IC
	Cornell Center for Social Sciences	IC
	Cornell Institute of Biotechnology	IC
	Cornell IT (CIT)	IT
	Cornell University Library	LIB
	Information Security Office	IT
	Research Development/Dean's Office	RSCH
	Sponsored Programs in the College of Life Sciences	RSCH
Duke University	Campus IRB	RSCH
	Duke Office of Research Initiatives	RSCH
	Medical Center Library	LIB
	Office of Campus Research Development (OCD)	RSCH
	Office of Information Technology - Central IT Financial	IT

<b>Institution</b>	<b>Responsive Department/Office</b>	<b>Service-Area Categorization</b>
University of Michigan	Office of Information Technology - Central IT Operations	IT
	Office of Science Integrity	RSCH
	School of Nursing	IC
	University Libraries	LIB
	Innovation Partnerships	RSCH
	Information and Technology Services (ITS) - Advanced Research Computing (ARC)	IT
	Medical School Office of Research	RSCH
	Michigan Institute for Data Science (MIDAS)	IC
	Office of General Counsel	RSCH
	Office of Regulatory Affairs, Medical School	RSCH
	Office of Research - Innovation Partnerships	RSCH
	Office of Research UM - Flint	RSCH
	Office of Research and Sponsored Projects	RSCH
	Office of the Vice President for Research (OVPR) - Research Data Stewardship Initiative	RSCH
	Office of the Vice President for Research (OVPR) - Research Integrity	RSCH
	University of Michigan Biological Station (UMBS)	IC
	University of Michigan Library	LIB
University of Minnesota	Center for Transportation Studies	IC

<b>Institution</b>	<b>Responsive Department/Office</b>	<b>Service-Area Categorization</b>
	Chemical Engineering and Materials Science	IT
	Clinical and Translational Science Institute (CTSI)	IC
	College of Liberal Arts	IC
	Export Controls Office	RSCH
	Genomics Center (UMGC)	IC
	Health Sciences Technology	IT
	Masonic Cancer Center	IC
	Neuroscience/University Imaging Centers	IC
	Office of General Counsel	RSCH
	Office of Information Technology (OIT)	IT
	Office of Information Technology - University Information Security (OIT-UIS)	IT
	Office of the Vice President for Research - Risk Intelligence & Compliance Team (OVPR/RIACT)	RSCH
	Office of the Vice President for Research (OVPR) - Office of Biotechnology Activities Oversight	RSCH
	Office of the Vice President for Research (OVPR) - Technology Commercialization	RSCH
	Research Computing	IT
	University Archives-University of Minnesota Libraries	LIB
	University of Minnesota Libraries	LIB
	Advanced Research Computing	IT
	Data Services - University Libraries	LIB
Virginia Tech		

<b>Institution</b>	<b>Responsive Department/Office</b>	<b>Service-Area Categorization</b>
Washington University in St. Louis	Fralin Biomedical Research Institute at VTC	IC
	Information Technology Security Office and Lab	IT
	Office of Sponsored Programs	RSCH
	Research and Innovation	RSCH
	Virginia Tech Transportation Institute (VTTI)	IC
	Bernard Becker Medical Library	LIB
	Institute for Informatics	IC
	Office of the Chief Information Officer (OCIO) - Research Infrastructure Services	IT
	Office of the Vice Chancellor for Research	RSCH
	Office of the Vice Chancellor for Research - Joint Contracts and Research Development (JCRD)	RSCH
	Sponsored Projects Accounting & Office of Sponsored Research Services	RSCH
	University Libraries	LIB



## Appendix B: Tableau Visualization Links and Labels

Visualizations 1–4 in this paper, as well as six institution-specific visualizations, can be found at: <https://public.tableau.com/app/profile/cynthia.vitale8121>. The following is a list of all responding units/departments from the Institutional Infrastructure Survey, and their service area categorization used in Visualizations 1–4.

Phase	Activity	Activity Label on Tableau Visualization
Planning, Design, and Start Up of Projects	Assessing data security needs and recommending solutions	Assess
	Developing or reviewing Materials Transfer and/or Data Use Agreements	DTUA
	Developing, building, or recommending storage solutions for active research data	Active Solutions
	Reviewing data management and sharing costs and expenses to be included in grant budgets	Budget
	Reviewing of IRB protocols and informed consent for data sharing	IRB
	Reviewing or preparing data management plans (DMPs) or data management and sharing (DMS) plans	DMP
	Supporting an appropriate repository (or repositories) for making research data broadly available	Repo
	Supporting intellectual property and copyright considerations	IP Support
Data Collection, Storage, and Management	Creating quality-control mechanisms or procedures	QC
	Developing or reviewing documentation of data (e.g., data dictionary, protocols)	Docs
	Evaluating or recommending data-analysis tools and processes to support sharing and reproducibility	Eval Tools

Phase	Activity	Activity Label on Tableau Visualization
Making Data Broadly Available	Managing active data (e.g., storage, security, backup, lab notebooks)	Active DM
	Consulting on decisions about what data to share or host	Repo Choice
	Consulting, selecting, or applying licenses to data	Licenses
	Creating or recommending persistent identifiers (PIDs; e.g., DOIs)	PIDs
	Creating or reviewing documentation for sharing (e.g., structured metadata, README files)	Review/Curation
	Preparing or consulting on preparing data for sharing (e.g., de-identification, selection, curation, data cleaning, validation, quality control)	Prep
	Checking for compliance with existing Data Use Agreements (DUAs)	Compliance
	Providing or hosting repositories for making data available	Hosting
	Recommending or migrating data file formats be open or more accessible	Transformation
	Submitting data into a data sharing platform (e.g., institutional repository, generalist repository, disciplinary repository)	Submission
Data Retention, Including Preservation, Archive, and Long-Term Access	Consulting on or migrating files to new formats or systems as needed	Migration
	Ensuring data security when appropriate (e.g., PHI/HIPAA, Export Controls, FISMA, student data, and Intellectual Property)	Security
	Making decisions about de-accessioning and removal of research data	Vetting
	Monitoring integrity of preserved data	Integrity
Project Closeout and Compliance	Ensuring funding agency requirements for data sharing have been met	Met Required
	Providing compliance support around research project reports	Report Support

## Appendix C: Ranked Factors Influencing Researcher Decisions on Where to Share Data

This figure complements Table 6 in the report, “Ranked factors with the greatest influence on researcher decisions on where to share their research data.” This figure shows the entirety of the ranked responses for the Researcher Survey question: “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.”

