




Journal of eScience Librarianship


putting the pieces together: theory and practice

“What About Your Friends?”: How a Collaborative Transdisciplinary Training Approach Supports FAIR Data Sharing Principles in Federally Funded Research

Shavonne Hedgepeth, University of Maryland, College Park, MD, USA 

Erin Antognoli, USDA National Agricultural Library, Beltsville, MD, USA, erinantognoli@gmail.com 

Sara Duke, USDA Agricultural Research Service, College Station, TX, USA 

Alison Rehfus, USDA Agricultural Research Service, Fort Collins, CO, USA 

Abstract

Objective: This paper details a pilot project to establish a baseline for current data management planning activities and offers more targeted data management planning training to researchers.

Methods: The authors incorporated a collaborative transdisciplinary approach, leading to the development and delivery of a series of surveys to gain accurate feedback about current workflows, policy adherence, and identifying knowledge gaps.

Results: Using formal survey results and informal feedback from researcher interactions to inform targeted training sessions and materials results in a more productive and collaborative experience for researchers and leads to more complete and structured data management plans.

Conclusions: Understanding researchers' current practices and needs is crucial to developing effective training and resources to help improve data management planning and workflows.

Received: July 14, 2023 **Accepted:** November 14, 2023 **Published:** December 20, 2023

Data Availability: Survey questions and responses are available under the article [Supplementary Files](#).

The *Journal of eScience Librarianship* is a peer-reviewed open access journal. © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

See <https://creativecommons.org/licenses/by/4.0>.

♻️ OPEN ACCESS

Introduction

Research Data Management (RDM) enables researchers to fulfill data sharing requirements and expectations established by funding agencies and publishers (Kanza and Knight 2022). Enforcement of new federal guidelines for data sharing practices has left research scientists unsure of the appropriate ways to meet policy requirements. While scientists possess intimate knowledge of the types of data they collect and are skilled at managing their data on their own computers and networks, they need help to identify appropriate long-term storage and access repositories. They also welcome guidance to create metadata that follows the FAIR principles to make data Findable, Accessible, Interoperable, and Reusable when making their data publicly available (GO FAIR, n.d.). Federal agency leadership also needs clarity on how to implement policies and provide clear guidance to the scientific community.

Federal guidelines stipulate that data resulting from federally funded research must be publicly accessible and work toward being FAIR. In response to federal memoranda on public access requirements, the USDA Agricultural Research Service (ARS) instituted an agency-wide policy in 2020 requiring researchers to submit a Data Management Plan (DMP) as part of their detailed project plan development to the ARS Office of Scientific Quality Review (OSQR). However, as guidelines and directives continue to evolve, researchers have conveyed difficulty understanding current requirements and best practices to satisfactorily complete their DMPs. As a result, researchers requested targeted guidance from the National Agricultural Library (NAL) to effectively manage their data throughout the entire data life cycle and to comply with current policies. Researchers specifically requested additional guidance on choosing a suitable repository for data storage and access to facilitate open science and support policy adherence. By using a participatory approach this project aided development of the NAL's targeted training and resource offerings by conducting pre-training surveys of researchers about their data management practices and needs (Cargo and Mercer 2008). The authors then tailored solutions to address knowledge gaps as individual National Programs (NP) approached review. Survey respondents helped identify discrepancies in researchers' understanding of data storage and access requirements, support data librarians' efforts to prepare targeted training, and facilitate communication with scientists to promote FAIR data sharing practices that adhere to both funder and journal requirements.

Background

The topics of data management and public access have become prominent focal points in the ongoing discussions surrounding the modern research ecosystem. However, the substantial push toward public access for data began less than two decades ago and lacked clear guidance on achieving compliance (Kriesberg et al. 2017). This resulted in further guidance and policy designed to make federally funded research data more publicly accessible and facilitate permanence. Librarians and information professionals have advocated for standardization in best practices and released guidelines to aid the advancement of data management and sharing as well as compliance with policy (Kanza and Knight 2022).

On January 4, 2011, President Barack Obama signed the America COMPETES Reauthorization Act of 2010, mandating the enhancement of research capabilities and coordination within the scientific community (U.S. Government Publishing Office 2011). This document reinforced the original America COMPETES Act published in 2007 and reaffirmed the policy to “invest in innovation through research and development, and to improve the competitiveness of the United States” through initiatives to fund and provide public access to STEM research. (Gonzalez 2015).

The NAL is one of five national libraries in the United States and houses one of the world’s largest collections devoted to agriculture and its related sciences (USDA National Agricultural Library, n.d.). In 2012, the NAL launched support services for digital scientific research data. The NAL’s services are founded upon data curation as a service for enhancing the value of research data for public access and re-use. They offer data management plan training and review services for scientists.

In 2013, the federal government sought to further facilitate an increase in research innovation, open science, and collaboration. The Director of the Office of Science and Technology Policy (OSTP), John P. Holdren, released the “Memorandum on Increasing Access to the Results of Federally Funded Research.” This memorandum directed federal agencies with research and development budgets above \$100 million to develop plans to share research data. (Holdren 2013). To fulfill the requirements set forth in the 2013 OSTP Public Access Memo, scientists require additional, specifically targeted guidance (Kriesberg et al. 2017).

To facilitate increased data sharing and public access for USDA-supported research data, the NAL created the Ag Data Commons to serve as a public, federally compliant, generalist data repository and catalog in 2015 (FAIRsharing.org 2018).

The FAIR Guiding Principles, a set of integral guidelines for proper data stewardship, were published on March 15, 2016 (GO FAIR, n.d.). The FAIR Principles emphasize the importance of sharing machine-readable data to allow for retrieval and reuse by machines or humans. The FAIR data principles have emerged as a standard for improving data management practices and facilitating data-driven scientific discovery. These principles provide a comprehensive framework for enhancing data management practices in federally funded research projects. Making research data FAIR is crucial for maximizing the impact of scientific research. However, the ability to achieve the FAIR principles heavily relies on robust data preservation strategies and RDM ecosystem (Wilkinson et al. 2016).

On September 14, 2020, to provide contextual guidance and expectations surrounding data sharing, the USDA ARS Office of National Programs released an internal document titled “Policy & Procedure 630: Data Management & Public Access Requirements for ARS,” otherwise known as “P&P 630.0.” This policy serves as an outline for the minimum requirements for managing ARS data and marks the first mandate in data sharing and public access for the agency (ARS Office of National Programs 2020). Further efforts to report on data management and public access are evidenced by the Office of Scientific Quality Review (OSQR),

a congressionally mandated entity, independent and objective within the ARS, tasked with ensuring the highest scientific quality for the Agency's People, Projects, and Programs. (U.S. Department of Agriculture, n.d.).

The diversity of research subject areas in the USDA and beyond, combined with lack of standardization in data management expectations, makes identifying and assessing suitable digital repositories challenging for many researchers seeking to meet the requirements laid out in these new policies. In response, on May 30, 2022, the National Science and Technology Council's (NSTC) Subcommittee on Open Science (SOS) released the "Desirable Characteristics of Data Repositories for Federally Funded Research." This document provides guidance on evaluating data repositories for compliance with federally funded research mandates. According to this document, well-documented and aggregated data stored in suitable data repositories serve as fundamental components in increasing reproducibility, shared knowledge, and innovation. Attributes researchers and data managers should consider when deciding on a repository's suitability to store their data fall under three main categories: Organizational Infrastructure, Digital Object Management, and Technology (The National Science and Technology Council 2022).

On July 20, 2022, the Office of the Chief Scientist within the U.S. Department of Agriculture released the "Departmental Regulation (DR) 1020-006 - Public Access to Scholarly Publications and Digital Scientific Research Data." This policy mandates the USDA make all peer-reviewed, scholarly publications and digital data assets associated with unclassified scientific research accessible to the public within 12 months of paper publication (Office of the Chief Scientist 2022).

Most recently, on August 25, 2022, Interim OSTP Director, Dr. Alondra Nelson released an updated memo requiring agencies and departments conducting federal research to update their public access policies to make federally funded research data publicly accessible immediately upon publication and without embargo no later than December 31, 2025 (Nelson 2022). This document also stressed the use of persistent identifiers (PIDs) and referred to the NSTC NSPM-33 definition of a PID as a globally unique, persistent, machine resolvable and processable digital identifier with an associated metadata schema (The National Science and Technology Council Subcommittee on Research Security 2022).

Challenges

Researchers and data managers at the USDA ARS face many challenges in their ability to comply with the numerous and evolving public access mandates and guidelines that support open transdisciplinary science. A series of surveys conducted prior to data management training sessions, along with feedback gathered during data management office hours, highlighted specific areas where researchers expressed a strong need for assistance. The identified areas where researchers sought help included:

1. Researchers had difficulty finding and navigating the numerous documents with mandates and guidelines and keeping track of policy updates.
2. Scientists and librarians refer to the same themes using different terms. Researchers expressed a need to bridge the language gap to support more effective research.
3. Researchers desire targeted guidance on agency approved repositories. Additionally, previous workflows may not be compliant in the present day.
4. A common theme among USDA researchers' feedback includes resource constraints: a lack of time, people, and funding to achieve data management objectives.

Methodology

Pilot project initiation

The pilot project to address data sharing needs began in October 2022 when the authors, consisting of the NAL Data Curation Team Lead, a USDA Plains Area Statistician, a USDA Information Management Specialist, and a Digital Curation Fellow from the University of Maryland, College Park's College of Information Studies, formed a small team to address the USDA ARS researchers' needs for data management guidance. This project, through collaboration with researchers and ARS National Program Leaders (NPL), aimed to synthesize numerous sources of data management policy into clear guidelines to assist with completing DMP documents and selecting appropriate repositories according to policy.

Iterative evaluation and participatory research

On October 20, 2022, the NAL data management team delivered a general data management training webinar to researchers in [ARS National Program 216 - Sustainable Agricultural Systems Research](#). During the question-and-answer period, researchers elucidated the challenges they encountered with data management planning and choosing an appropriate data repository. The authors, using this session as a baseline for understanding the pain points experienced by researchers when attempting to meet data sharing requirements, began evaluating available repositories to guide researchers toward desirable data repositories for long-term storage.

The team conducted an extensive evaluation of over 200 digital databases and repositories previously used as subject specific collaborative data sharing and storage solutions by USDA researchers (Antognoli, Sears, and Parr 2017). This work established a clear divide between previously used solutions that constitute a repository as opposed to a collaborative online database resource. Databases are designed to store structured data and provide features like transaction management, data indexing, and complex query support (Elmasri and Navathe 2021). "A repository holds data, makes data available to use, and organizes data in a logical manner. A data repository may also be defined as an appropriate, subject-specific location where researchers can submit their data" (National Library of Medicine, n.d.). Due to the considerable overlap in features, the NSTC

Desirable Characteristics of Data Repositories for Federally Funded Research document provided a detailed and comprehensive framework for evaluating the qualities of repositories to further clarify requirements for appropriate data repositories. Each storage solution in the list was evaluated using the principles outlined by the NSTC document (The National Science and Technology Council 2022), and each received a subject category and site navigation evaluation. These databases covered a wide range of research areas, totaling over eight distinct subject domains such as Genomics and Genetics (plant and animal), Plants and Crops, Agroecosystems and Environment, and Livestock. When employing this set of evaluation criteria, fewer than half of the previously used storage solutions strictly adhere to current federal compliance regulations and guidelines for data repositories. The remaining database resources were classified as non-compliant because they lack preservation standards, submission guidelines, or standardized metadata practices. Non-compliant databases also failed to provide a PID for deposited data.

Following the review of the 200 previously used databases and data repositories, a short list was synthesized to include roughly 10 compliant generalist repositories and 10 compliant subject-specific repositories based on knowledge of the types of research being conducted in the Plains Area.

Survey Administration

In February 2023, the authors released a survey for researchers within the Natural Resources and Sustainable Agricultural Systems [ARS National Programs](#) to gather information about their preferred data repositories and their experience uploading their data to repositories (Appendix A). The survey was conducted using the Survey Monkey platform. National Program Leaders, primarily in natural resources and environmental fields with upcoming review cycles, forwarded the survey link to their researchers. Over the course of three weeks, the initiative received an encouraging response, with a total of 123 participants providing their feedback out of 312 survey recipients for a roughly 40% response rate. The survey results yielded valuable insights that can now be used to guide the NAL in addressing the existing gaps in data stewardship and policy compliance. This feedback will play a crucial role in shaping the NAL's training strategies by providing a deeper understanding of researcher knowledge concerning data repositories, enabling the development of better support and guidance in navigating data management obligations within the research community.

As a result, the authors used findings from the first survey to narrow the scope to National Programs in the process of submitting their five-year project plans for their OSQR review. The next target group to receive a survey ahead of their training webinar was [National Program 107 - Human Nutrition](#) (Appendix B).

Plains Area data management office hours

Concurrently, in March 2023 the USDA ARS Plains Area began holding data management office hours. These sessions included a brief presentation that introduced the history and context of public access and policy releases that directly impact USDA researchers so they understand why they must now complete DMPs and share their data in open repositories. Other topics included reviewing the basics of DMPs, the

scope of the data for public access sharing, and how to choose an appropriate data repository. The focus of these sessions involved researchers managing data better for themselves, with a long-term goal of building a more reproducible workflow.

In total, from March 2023 to June 2023, a Plains Area statistician conducted 23 two-hour sessions with up to seven attendees per session. The statistician on the team provided the link to the scientific community being studied. Her experience with research data in the organization and relationships with the community helped guide strategies to target different research specializations for the surveys and interpret results. The idea to use office hours as a mechanism to disseminate guidance for data sharing based on policy and fulfilling requirements in a more personal interaction discussion forum came from her 20 years of experience interacting with the scientists one-on-one. Office hour sessions generally reached capacity and the informal nature of the sessions allowed scientists to provide valuable feedback on their primary pain points.

Results

Survey respondents

Of the 123 respondents from the first survey, over 52% stated that they had never uploaded their data into repositories previously (Figure 1), highlighting a significant gap in the utilization of repositories for data management among the survey respondents.

Only 27% of survey respondents reported having a preferred repository (Figure 2). When asked to name their preferred repositories, respondents provided an even split of compliant (e.g., Ag Data Commons, Ameriflux, Dryad, and Environmental Data Initiative) and non-compliant options. Minimum features that make a repository compliant with policy expectations include persistent identifiers for data assets with the ability to provide ongoing public access for data without privacy, security, proprietary, or confidentiality restrictions, as well as adequate preservation and backup assurances to ensure safety of the assets long-term. Nearly 55% of respondents needed help finding a repository (Figure 3), and 65% requested more training on how to evaluate and select a repository (Figure 4). Responses came predominantly from researchers in National Programs representing a wide range of scientific domains with upcoming OSQR review cycle due dates, reflecting an urgency for support from the NAL. Survey results indicated that finding and choosing an appropriate repository proved a significant barrier to ensuring compliance with federal and USDA policy.

The second survey respondents, although a smaller group (n=11), echoed the same challenges discovered from the results of the first survey. Seventy percent of the survey respondents in group 2 confirmed they uploaded their data into repositories and provided examples of the data repositories used within the last five years. Among those were REDCap, MG-RAST, Ag Data Commons, all of which embody the desirable characteristics identified by the NSTC's SOS. Despite most respondents' experience using a repository, the group was still split on having a preferred repository with a few respondents listing "journal" as a repository. Sixty percent of respondents indicated they didn't need help finding a repository. However, on the question

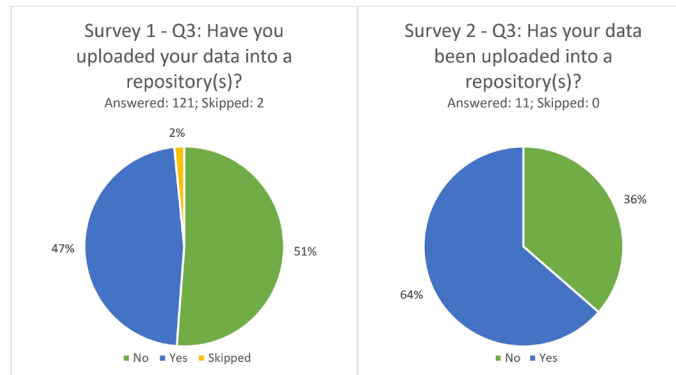


Figure 1: The majority of researchers from the first survey had never uploaded their data into a repository before. The second survey respondents represented a much smaller, more targeted research area with better established data management protocols. Roughly 1/3 of the smaller, more focused group had never uploaded their data into a repository compared to over half from the more general population of survey respondents from the first survey.

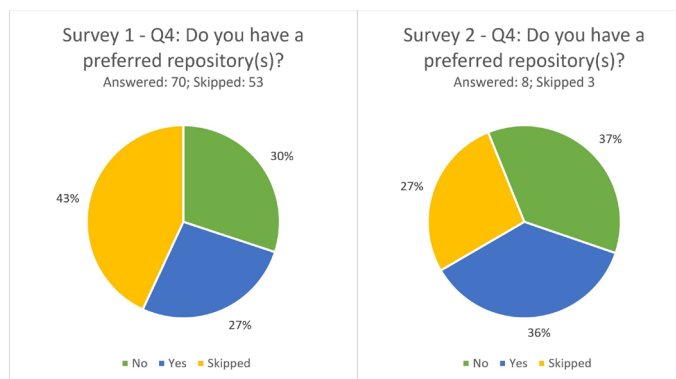


Figure 2: The majority of researchers surveyed did not have a preferred repository. Respondents who had never uploaded their data to a repository before were instructed in the previous question to skip this question. Those that named previously used repositories provided a nearly even mix of compliant and non-compliant storage solutions in both surveys.

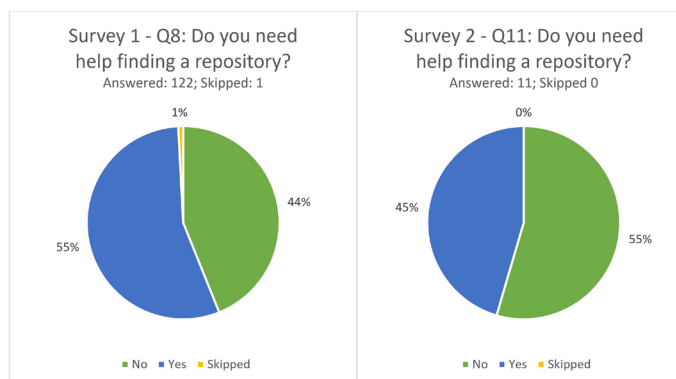


Figure 3: The majority of researchers surveyed needed help finding a repository

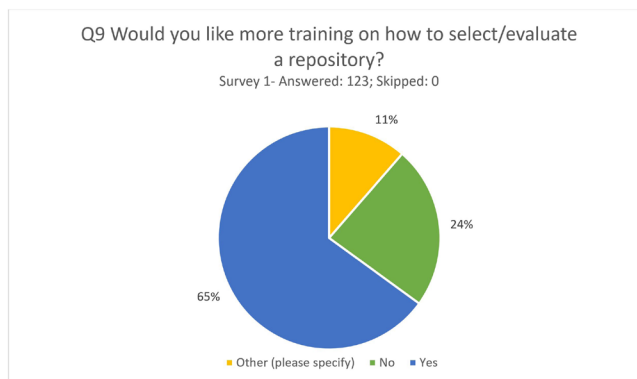


Figure 4: The majority of researchers surveyed would like more training on how to select/evaluate a repository. This question was not asked in survey #2 because the research domain of that group was more focused with a goal of informing researchers about existing compliant repository workflows.

inquiring what they would like to see in a training session, some of the responses indicated they need more targeted guidance on “best practices for different repositories,” “an overview of what the ARS can provide,” “demonstrations for uploading data sets in general purpose repositories,” and “best methods for data sharing.”

Like their peers in survey group 1, researchers suffered from unclear guidance on repository selection and emphasized the importance of establishing best practices around data management and storage preservation.

NAL DMP review service and template

The difference between working project storage and long-term preservation storage needed clarification as evidenced by both the survey answers where researchers listed repositories they have used (e.g., “I have uploaded data as supporting information to journal articles. This is not my preferred approach.”), as well as from information gleaned during DMP reviews conducted by the NAL in the several years leading up to the most recent public access policies. The NAL has offered a DMP review service to ARS researchers since 2018. While the criteria for review were less structured early on, the library has since more closely aligned review criteria and guidance to suit the expectations outlined in P&P 630. The NAL continues to update guidance with the release of new policies and guidelines and actively collects valuable feedback from researchers to target specific data management topics that require attention and improvement. By continuously updating its guidance and incorporating researcher input, the NAL strives to provide comprehensive and relevant support in the ever-evolving landscape of data management. This ongoing information stream, combined with the survey results, led the NAL librarians to develop the first draft of a DMP template to help guide researchers in accounting for all requirements for data management and sharing. The initial general DMP template provided fill-in-the-blank style guidance to form a starting point for completing the DMP. Among those features included distinct statements about working project storage and preservation storage, prompts

for declaring file formats and standards, and public access features required by USDA and federal policy. Additionally, repositories often dictate acceptable data types and file formats, data and metadata standards, preservation policies, and public access mechanisms. Because DMP content largely relies so closely on a researcher's repository choice, the authors identified this as an area to expand guidance and services (Figure 5).

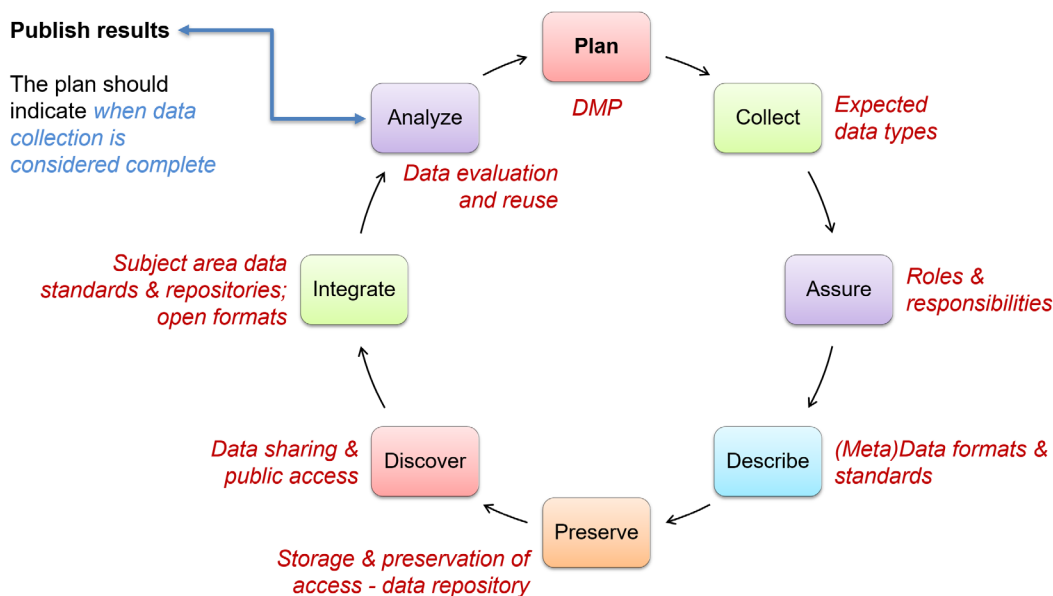


Figure 5: Repository choice influences all other aspects of the data life cycle. DataONE life cycle model (DataONE, n.d.) overlaid with sections and considerations of the USDA ARS DMP.

Data management planning training

Following the initial surveys, the NAL data management planning team began scheduling targeted training for more National Programs and ARS research units, and Plains Area statisticians continued holding data management office hours.

The NAL incorporated survey feedback when planning a webinar and informational panel for NP 107 researchers early in April 2023. Library staff promoted the new DMP templates along with NAL DMP review services and other available resources for researchers to use during this process. In addition to the general template, the NAL created DMP template examples based on several of the repositories and subject areas indicated by this National Program in the survey results. Templates helped to address the language gap in showing researchers what types of information the administration expected in their documentation. Customizing outreach and webinar delivery provided a more relevant training experience for the group.

This allowed for the inclusion of suggested subject-specific repositories and examples that resonated with the participants, ensuring the training content directly addressed their needs and challenges. As a result of the targeted training and outreach, the third quarter of fiscal year 2023 reflected the busiest period for DMP reviews requested from NAL staff to date. Immediately following this training webinar, the NAL data management planning team reviewed eight DMP submissions from NP215 and NP 107, both of which were undergoing their OSQR review cycles that quarter. Several of those researchers decided to use the template, and the completeness and quality of those DMPs increased dramatically while average review times for DMPs decreased. This early evidence indicates that targeted training and resources help guide researchers in their DMP creation and ultimately their data management practices.

Next Steps

The benefits of data preservation in federally funded research extend beyond immediate research needs. Properly preserved data enables scientific progress and contributes to the cumulative knowledge base through citation and reuse (Piwowar, Day, and Fridsma 2007). Moreover, data preservation can help ensure compliance with funding agency requirements, promote transparency and accountability, and support the replication and verification of research findings (Allen and Mehler 2019). By prioritizing data preservation, federally funded research projects can uphold responsible data stewardship and foster a culture of sharing and collaboration (Piwowar, Day, and Fridsma 2007).

While librarians and information professionals typically become involved in data management toward the publication and disposition end of the data life cycle, certain actions can affect researcher behavior further upstream. This includes offering training and information to combat pain points expressed by researchers. This effort can promote data management best practices in a broader capacity during the planning phase of the project (Antognoli et al. 2020).

Increased collaboration between researchers, ARS area offices, and the NAL can provide a blueprint for more relevant resources for ARS researchers. Through surveys prior to training as well as feedback following training sessions, researchers indicated they need more help choosing an appropriate repository for their data. The NAL provides the [Ag Data Commons](#) as a generalist repository and public access point to accommodate USDA-supported data assets and promote discovery (FAIRsharing.org 2018), but targeted file types, metadata, and other community-driven guidance offered by domain-specific repositories can promote more interoperability and reusability (Horsburgh et al. 2020). While the templated DMP takes a good first step at helping researchers organize their thoughts and create checklists for managing data, taking the guesswork out of choosing an appropriate repository for preservation and access with an eye toward future reusability provides a logical next step in aiding researchers with data management planning.

Once researchers create workflows for storing, preserving, and making their data accessible, future work to expand on these benefits includes creating mechanisms to make data interoperable and reusable by

encouraging the creation and use of data dictionaries and README files. The NAL currently offers templates for these documents but could engage in more targeted outreach like current DMP training sessions. One participant highlighted researchers in Florence, SC who developed a protocol for adding data to Dryad, an open-access generalist repository (Billman et al. 2023). Future work to expand on developing workflows for researchers to deposit data may prove fruitful.

The Plains Area plans to continue office hours and other outreach and engagement efforts to maintain communications about existing and new policies and guidelines.

Conducting pre-surveys before training will continue to help identify the pain points for both broad and narrow research groups. Understanding the program review schedule informed the survey strategy for this round of training, but lessons can apply to other user groups or end goals. To adapt to changing user needs, the NAL data management team intends to conduct regular surveys of researchers. These surveys will precede future training sessions and significant outreach initiatives, ensuring a continuous flow of information and effective communication. This approach aims to deliver highly valuable services that cater to the evolving requirements of the user community.

Conclusion

Challenges in easily accessible knowledge around the criteria for proper data storage cause temporal bottlenecks and serve as a barrier to the decades-long goal of achieving open science (Lortie 2021). Proper data preservation and access in federally funded research projects promote FAIR data principles and cross-collaboration. Researchers, institutions, and funding agencies can contribute to the longevity, usability, and impact of research data, thereby integrating knowledge and maximizing the return on public investments. Achieving workflows that embrace data management best practices will happen gradually with targeted work toward these goals. Information professionals can take steps to understand the researchers' current workflows to bridge the knowledge gaps in the data management policies and best practices through regular communication including surveys and training sessions. With more focused, iterative, and intentional efforts, libraries can provide targeted services to help researchers in the quest for better data management and policy compliance.

Data Availability

Survey questions and responses are available under the article [Supplementary Files](#):

Appendix A: Survey 1 Questions

Appendix B: Survey 2 Questions

Survey 1 Responses: [10.7191/jeslib.755.s837](https://doi.org/10.7191/jeslib.755.s837)

Survey 2 Responses: [10.7191/jeslib.755.s838](https://doi.org/10.7191/jeslib.755.s838)

Acknowledgements

This work was supported by the [University of Maryland, College Park](#) and the U.S. Department of Agriculture, [Agricultural Research Service](#).

The content of this article is based on the presentation entitled “Creating and refining data management planning services for a diverse scientific research community” originally presented at [RDAP Summit 2023](#), available from Open Science Framework: <https://osf.io/f4t2k>.

Competing Interests

The authors declare that they have no competing interests.

References

- Allen, Christopher, and David M.A. Mehler. 2019. “Open science challenges, benefits and tips in early career and beyond.” *PLoS Biology* 17(12): e3000587. <https://doi.org/10.1371/journal.pbio.3000246>.
- Antognoli, Erin, Jonathan Sears, and Cynthia Parr. 2017. “Inventory of online public databases and repositories holding agricultural data in 2017.” *Ag Data Commons*. <https://doi.org/10.15482/USDA.ADC/1389839>.
- Antognoli, Erin, Regina Avila, Jonathan Sears, Leighton Christiansen, Jessica Tieman, and Jacquelyn Hart. 2020. “Reproducibility literature analysis - a federal information professional perspective.” *IASSIST Quarterly* 44(1-2): 1-26. <https://doi.org/10.29173/iq967>.
- ARS Office of National Programs. 2020. “ARS P&P 630.0: Data Management & Public Access Requirements for ARS.” September 14.
- Billman, Eric D., Sohoulane, Clement D., and Vanotti, Matias B. 2021. Protocol for Making Data Publicly Available in USDA-ARS. *Protocols.io*. <http://dx.doi.org/10.17504/protocols.io.ewov1o5bolr2/v1>.
- Cargo M, Mercer SL. 2008. “The value and challenges of participatory research: strengthening its practice.” *Annual Review of Public Health* 29(1): 325-350. <https://doi.org/10.1146/annurev.publhealth.29.091307.083824>.
- DataONE. n.d. Data Management Skillbuilding Hub. Accessed 2023. <https://dataoneorg.github.io/Education>.
- Elmasri, Ramez, and Shamkant B. Navathe. 2021. *Fundamentals of Database Systems*. 7th ed. Pearson.
- FAIRsharing.org. 2018. Ag Data Commons. Accessed 2023. <https://fairsharing.org/10.25504/FAIRsharing.83wDfe>.
- Gonzalez, Heather B. 2015. The America COMPETES Acts: An Overview. Washington, DC: Library of Congress. <https://crsreports.congress.gov/product/pdf/R/R43880/13>.
- Holdren, John P. 2013. “Increasing Access to the Results of Federally Funded Scientific Research.” February 22. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

- Horsburgh, Jeffery S., Richard P. Hooper, Jerad Bales, Margaret Hedstrom, Heidi J. Imker, Kerstin A. Lehnert, Lea A. Shanley, and Shelley Stall. 2020. "Assessing the state of research data publication in hydrology: A perspective from the Consortium of Universities for the Advancement of Hydrologic Science, Incorporated." *WIREs Water* 7(3). <https://doi.org/10.1002/wat2.1422>.
- Kanza, S., Knight, N.J. 2022. "Behind every great research project is great data management." *BMC Research Notes* 15(20). <https://doi.org/10.1186/s13104-022-05908-5>.
- Kriesberg, Adam, Kerry Huller, Ricardo Punzalan, and Cynthia Parr. 2017. "An Analysis of Federal Policy on Public Access to Scientific Research Data." *Data Science Journal* 16(0): 27. <https://doi.org/10.5334/dsj-2017-027>.
- Lortie, Christopher J. 2021. "The early bird gets the return: The benefits of publishing your data sooner." *Ecology and evolution* 11(16):10736-10740. <https://doi.org/10.1002/ece3.7853>.
- National Library of Medicine. n.d. Data Glossary. Accessed 2023. <https://www.nlm.gov/guides/data-glossary/repository>.
- Nelson, Alondra. 2022. "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research." February 22. <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>.
- Office of the Chief Scientist. 2022. "USDA Departmental Regulation 1020-006: Public Access to Scholarly Publications and Digital Scientific Research Data." July 20. <https://www.usda.gov/directives/dr-1020-006>.
- Piwovar, Heather A., Roger S. Day, and Douglas B. Fridsma. 2007. "Sharing Detailed Research Data Is Associated with Increased Citation Rate." *PLoS ONE* 2(3): e308. <https://doi.org/10.1371/journal.pone.0000308>.
- The National Science and Technology Council. 2022. "Desirable Characteristics of Data Repositories for Federally Funded Research." <https://doi.org/10.5479/10088/113528>.
- The National Science and Technology Council Subcommittee on Research Security. 2022. "Guidance for Implementing National Security Presidential Memorandum 33 (NSPM-33) On National Security Strategy For United States Government-Supported Research And Development." Accessed July 2023. <https://www.whitehouse.gov/wp-content/uploads/2022/01/010422-NSPM-33-Implementation-Guidance.pdf>.
- U.S. Department of Agriculture. n.d. The Office of Scientific Quality Review. Accessed 2023. <https://www.ars.usda.gov/office-of-scientific-quality-review-osqr/the-office-of-scientific-quality-review>.
- USDA National Agricultural Library. n.d. USDA National Agricultural Library. Accessed 2023. <https://www.nal.usda.gov>.
- U.S. Government Publishing Office. 2011. "America COMPETES Reauthorization Act of 2010." January 4. <https://www.congress.gov/111/plaws/publ358/PLAW-111publ358.pdf>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The Fair Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 160018. <https://doi.org/10.1038/sdata.2016.18>.