ASSOCIATION OF RESEARCH LIBRARIES

January 2, 2024

Martin Halbert, PhD
Science Advisor for Public Access
National Science Foundation
2415 Eisenhower Ave.
Alexandria, VA 22314

Re: RFI Response: NSF Public Access 2.0

On behalf of the members of the Association of Research Libraries (ARL), thank you for the opportunity to provide comments on the National Science Foundations's "Public Access Plan 2.0." ARL is a membership organization representing the largest research libraries in the United States that are committed to the advancement of open scholarship and open access to accelerate scientific advances and expand diverse, public participation in federally funded research. We appreciate NSF's commitment to making the results of federally funded research widely available without embargo, leveraging persistent identifiers to support scientific integrity, and ensuring equitable access. We submit the following comments to NSF's request for information (RFI).

1. **Overall, do you view public access requirements as having more positive or more negative effects on equity and inclusion in science?**

Mostly positive, but to increase equity and inclusion in public access requirements, ARL recommends that NSF:

- Work with research institutions, their libraries, and their professional and scholarly associations on coordinated education to investigators on their options for no-fee manuscript deposit.
- Explore preprint services to accelerate the sharing of research findings and the potential of recognizing peer-reviewed preprints that are substantially similar to author-accepted manuscripts for compliance with the policy.
- Ensure final publisher PDF versions are accessible. This is critical to enable equitable delivery of federally-funded research results. While publishers of scholarly works can add accessible features most effectively and efficiently, PDFs provided by publishers are not always 508 compliant. We recommend that public access plans include a plan to remediate non-compliant works per U.S. copyright law, which explicitly grants a broad exception for remediation and distribution of accessible works to people with print disabilities. This exception is bolstered by the Marrakesh Treaty.
- Provide rights-retention language (for investigators to use upon submission of manuscripts to journals) that encourages authors to retain their copyrights and assign a Creative Commons Attribution (CC BY) or similar license to their work to enable full reuse rights.

- Adopt an evolving ethical framework, such as the U.S. Department of Commerce Data Ethics Framework,[1] and specifically consider research that involves human participants and local communities.

7. If you have any additional comments about NSF's Public Access Plan, please share them here.

Our recent research[2] has found the average researcher expenditure for research data management and sharing activities throughout the grant project period was $29,800 or 5.83% of their award amount. This research found that the average yearly institutional expense (researcher expense + service provider expense) for data management and sharing is $2,500,000 with a range from approximately $800,000 to over $6,000,000. Further, our research indicates that Libraries and University IT bear the greatest financial and service burden for data management and sharing (Table 1). While researcher direct costs may be included in grant budgets, overall institutional expenses are not yet well accounted for through institutional direct or indirect cost reimbursement.

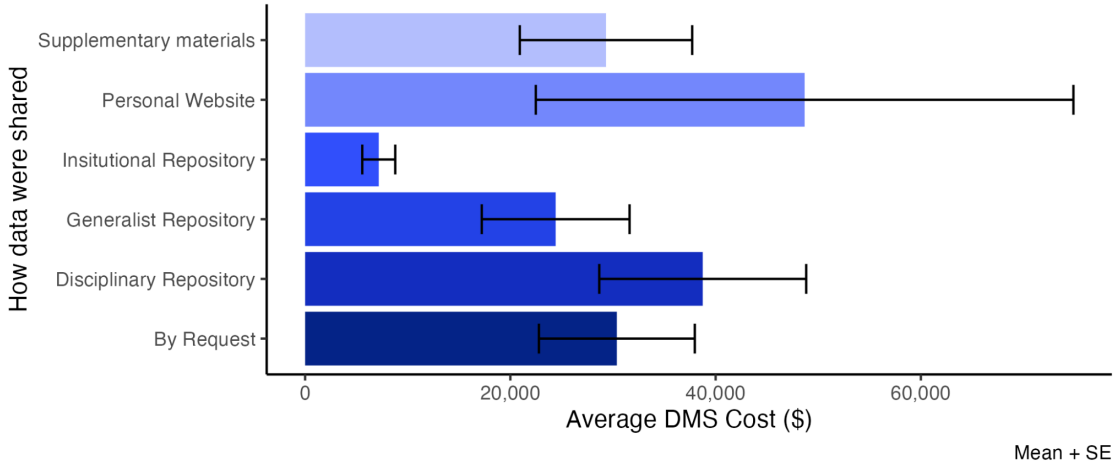
Table 1: Average annual costs to support research DMS services by service provider area (values are rounded to three significant figures)

Service Provider	**Total Average Annual Cost**	**Standard Deviation**
Libraries	**$303,000**	$164,000
IT Offices	**$249,000**	$218,000
Research Offices	**$93,800**	$60,400
Institutes & Centers	**$94,900**	$76,500
Total Cost (Sum)	**$740,000**	$107,000

Additionally, as seen in Figure 1, when leveraging institutional services, such as institutional repositories for data sharing, researcher expenses for data management and sharing had a lower average expense when compared to the average expense when using a different data-sharing location.

[1] United States Department of Commerce."Commerce Data Ethics Framework." (2022). https://www.commerce.gov/sites/default/files/2023-02/DOC-Data-Ethics-Framework.pdf
[2] *Making Research Data Publicly Accessible: Estimates of Institutional & Researcher Expense.* Report. Forthcoming February 2024.

Figure 1: Average DMS costs per funded research project by how research data were shared.



Based upon the information above and to ensure researchers have support to meet requirements, ARL recommends that NSF:

- Minimize the administrative and financial burden on researchers and institutions for compliance by working with institution-based service providers to educate and support the preparation of materials for sharing for public access.
- Specify allowable (and unallowable) costs for data management and sharing activities. This includes clearly stating if data storage and repository expenses post-award are allowed.
- Develop a mechanism to ensure that funds are available post-closeout for publication and research data storage and/or sharing expenses. Post-award publication funding may be particularly important for early-career, postdoctoral, and graduate student researchers whose publication and data-sharing costs may not have been factored into the original grant budget.
- Adopt the Implementing Effective Data Practices report recommendations from higher education associations.[3]
- Review and prioritize the adoption of rights-holder-centered principles for the ethical care of research data and engage stakeholder communities. For example, the CARE Principles for Indigenous Data Governance were created to allow Indigenous People to assert greater control over the use of Indigenous data and knowledge.[4] Tribal stakeholder engagement should be expected for research done on or about Indigenous Peoples.

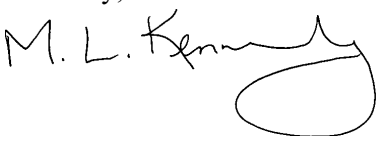
We look forward to continued engagement with NSF during the development and refinement of their Public Access Plan 2.0. We are happy to work with NSF to identify ARL member institutions interested in participating in conversations regarding any of these specific topics. With any questions about these comments,

[3] Chodacki, John, Cynthia Hudson-Vitale, Natalie Meyers, Jennifer Muilenburg, Maria Praetzellis, Kacy Redd, Judy Ruttenberg, Katie Steen, Joel Cutcher-Gershenfeld, and Maria Gould. *Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support*. Washington, DC: Association of Research Libraries, September 2020. https://doi.org/10.29242/report.effectivedatapractices2020.

[4] Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons et al. "The CARE principles for indigenous data governance." Data Science Journal 19 (2020): 43–43. DOI: 10.5334/dsj-2020-043.

please feel free to contact me or my colleague Cynthia Hudson Vitale, Director, Science Policy and Scholarship, ARL, cvitale@arl.org.

Sincerely,



Mary Lee Kennedy, Executive Director
Association of Research Libraries