

February 16, 2024

Re:Docket Number:2024-01673. Response to “Notice of Second Interested Party Feedback Period Regarding Increasing Public Access to the Results of USDA-Funded Research”

Thank you for the opportunity to provide input on USDA's public access plan, "Increasing Public Access to the Results of USDA-Funded Research." We are librarians at the University of Minnesota who support researchers in making their work open access and curating, managing, and sharing research data. We are submitting these comments as individuals, not representatives of the University. Our comments are organized according to the questions posed in the Federal Register.

### **(1) How can USDA best implement its plans to improve public access to USDA-funded research results?**

In addition to the points we raise in response to (2), we make the following recommendations to USDA:

- PubAg should be upgraded to have a PMH-endpoint (<https://www.openarchives.org/pmh/>) so that it is able to be connected to Unpaywall (<https://unpaywall.org/>), making materials in the repository more findable. For example, Unpaywall data is used in Web of Science and Scopus, two major literature databases. AgEcon Search (<https://ageconsearch.umn.edu/?ln=en>) is an example of a long-running agricultural research-focused repository that is included in Unpaywall's sources.
- USDA should provide much clearer guidance to researchers regarding data sharing, including a summary of major points to know and consider, where to share data, who the policy applies to, and what is meant by the term "data." The statement attempting to define data (“The USDA public access policy will apply to any digital scientific research data asset connected to a scholarly publication covered by the USDA public access policy. The public access policy will also extend to other digital scientific research data assets as currently outlined in DR 1020-006, Public Access to Scholarly Publications and Digital Scientific Research Data”) leaves questions as to whether researchers will be required to share only the data underlying publications, all data, or some other subset of data; in what format to share data; costs associated with sharing “big data”; timelines; and exemptions. While this information may be scattered throughout the long, formalized policy, it is not easy to locate, understand, or interpret. We recommend USDA use the National Institutes of Health (NIH) data sharing web pages as an example. Implementation and compliance can be greatly improved with this step.
- If the USDA is not going to host all of the publications and data, USDA should provide a list of “agency-designated repositories” where researchers can share their data, including “big data.” This list may not need to be exhaustive, but should provide examples and explain why they are considered “agency-designated repositories” (i.e., have been vetted by CoreTrustSeal, follow FAIR data practices). The White House’s Desirable Characteristics of Data Repositories (<https://www.whitehouse.gov/wp-content/uploads/2022/05/05-2022-Desirable-Characteri>

[stics-of-Data-Repositories.pdf](#)) is a useful starting point, but examples specifically for USDA researchers, for the types of data they work with, would be helpful.

## **(2) How can USDA update or refine its policies to improve public access to USDA-funded research results?**

In addition to the points made in (1), we recommend USDA:

- Clarify when publications must be made publicly accessible. E-versions of articles are frequently published on journal websites ahead of print, sometimes several months or even more than one year afterwards. The policy should clearly state whether the requirement applies based on the date of e-publication or the date of publication in a numbered volume/issue.
- Standardize language in the plan with other federal agencies. The use of the term "digital persistent identifiers (DPIs)" is not standard and may confuse researchers trying to comply with USDA policy. These are more commonly called Persistent Identifiers (PIDs). Also, "data asset" is not a commonly used term, which will lead to unnecessary confusion about what USDA requires; we recommend using the term "dataset." We support USDA's plan to recommend the use of ORCID iDs and DOIs and suggest also requiring the use of Research Organization Registry (ROR) identifiers, which are the only open (CC0 licensed) PID for research organizations. ROR is already integrated into metadata sources like Crossref and DataCite and supported in ORCID.
- Work with other federal funding bodies to clarify similarities and differences in expectations. Researchers may use funding from multiple bodies for a single research publication or dataset (e.g., USDA and National Science Foundation, USDA and NIH). Use of non-standard language and differing expectations will introduce unnecessary complications, making it difficult for researchers to keep track of requirements and ensure that they are in compliance with all of their grants.
- Follow through on the steps listed in section 2.4 "Public Input in New Policy Development", including rechartering the USDA Public Access and Open Science Forum, following a formal Tribal Consultation Process, and consulting with all of the external stakeholders listed, using these processes to identify additional stakeholders whose input should be considered.
- Provide more guidance and advocacy for Indigenous data sovereignty. We appreciate that Tribal Consultation Processes are listed once in the Public Input section; however, work should be done *with* the community, not *around* the community. We recommend referring to NIH's Supplemental Information to the NIH Policy for Data Management and Sharing: Responsible Management and Sharing of American Indian/Alaska Native Participant Data (<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-214.html>) as a guide.
- Define "long-term preservation" and address issues of accessibility beyond "machine readability." The policy states "USDA will require digital scientific research data assets covered by public access requirements be published in a machine-readable format by a data repository that is recognized by reputable registries and provides: • Public access for search, retrieval, and analysis. • A digital persistent identifier, such as a DOI. • Long-term preservation of the data asset." We applaud the requirement for a DOI, but

"long-term preservation" should be defined. Repositories have varying preservation and deaccession policies. Additionally, machine-readable formatting is not necessarily accessible or equitable and these aspects of the requirement should be considered.

### (3) How can USDA ensure equity in publication opportunities?

To meet the requirements of the 2022 White House Office of Science and Technology Policy Memorandum to make federally funded publications available "in an equitable and secure manner", we recommend several steps USDA could take to increase equity and prevent the introduction of new inequities in publishing opportunities.

- We suggest USDA:
  - Continue to use PubAg (improving it as noted in (1)) and accept the author's accepted manuscript (AAM) version as the primary mechanism for compliance.
  - Clarify, in the plan and in communications to grantees, that fees publishers request are **not** required for compliance with the funding requirements.
  - Consider relying on federal purpose licenses under 2C.F.R. § 200.315(b). This will ensure that if an author signs a publisher agreement that conflicts with the public access policy requirements, USDA will retain the permissions necessary to make the work publicly accessible. This route to compliance should be clearly explained in all documentation provided to research grantees.
- We recommend emphasizing in USDA's policy and documentation provided to grantees that the primary path to complying with the public access requirement for publications does **not** require payment of article processing charges (APCs) to publishers.
- Although USDA-funded researchers may write article processing charges (APCs, fees charged by publishers to make articles open access) into their grants, not all researchers who conduct research in agricultural disciplines have USDA funding. USDA should take care to avoid contributing to a system in which all authors are required to pay to publish their research. The 2023 United Nations Open Science Conference (<https://www.un.org/en/library/OS23>) emphasized that to address the challenges facing our planet (e.g., achieving the United Nations' Sustainable Development Goals), all researchers, globally, must be able to share their research. USDA's policies should promote equitable publishing models and proactively avoid potential adverse outcomes.
- In addition to promoting non-APC based routes as the preferred mechanism for compliance with the public access policy, USDA could highlight open access (OA) journals that do not charge fees, such as journals using the Subscribe-to-Open (<https://subscribetooopencommunity.org/>) or other "Diamond" OA (<https://www.coalition-s.org/diamond-open-access/>) models. We suggest that USDA should also consider supporting publishers that use these models, to contribute to a scholarly publishing system that is open to all, regardless of grant funding status. These are equitable publishing models in which no one has to pay to read and no authors pay to publish. This model is a focus for funding bodies in such as Science Europe, cOAlition S, OPERAS, and the French National Research Agency, which have together created an "Action Plan for Diamond Open Access" that USDA could consider supporting (<https://zenodo.org/record/6282403#.ZDhEvXbMI2w>).

- APCs take money away from other important research expenses. In listening sessions and a survey conducted by the American Association for the Advancement of Science (AAAS), researchers 58% reported being unable to attend workshops or conferences and 75% reported being unable to purchase materials, tools, or equipment because funding has been used to pay APCs ([https://www.aaas.org/sites/default/files/2022-10/OpenAccessSurveyReport\\_Oct2022\\_FINAL.pdf](https://www.aaas.org/sites/default/files/2022-10/OpenAccessSurveyReport_Oct2022_FINAL.pdf)).
- Emphasis on APC-based models of open access has the potential to have a greater impact on early career researchers and those at less-well-resourced institutions. In the survey mentioned above, AAAS found that most institutions do not have funds available to support payment of APCs (only 32 of 89 institutions reported having funds for such fees). At a large and research-intensive institution such as the University of Minnesota, we would be unable to pay APCs for all of the nearly 10,000 articles published each year. Even institutions as well-resourced as the Ivy League universities note that paying APCs "detracts from our core mission of open access and more specifically our ability to comply with the proposed policy changes that we so overwhelmingly support. Implementing the Nelson memo via an APC model is antithetical to the equity goals so clearly articulated in the guidance memo and the values of our institutions." (<https://libraries.mit.edu/news/libraries-support-3/34036/>)
- When consulting with the stakeholders listed in section 4.1 of the Plan, USDA should take into account the motivations behind feedback. For example, commercial publishers may be beholden to shareholders and have a primary focus on generating profit. The federal public access policies should be driven by public good, with equity, diversity, and inclusion as the highest standards, not profit.
- If APCs remain allowable costs for USDA grantees, USDA should track the amount of APCs paid from USDA grants as well as costs related to data sharing. APC prices increase every year, and have done so at rates much higher than inflation (<https://doi.org/10.18352/lq.10280>). These fees can range up to \$12,290 per article (<https://www.nature.com/nature/for-authors/publishing-options>). Using grant funding to pay such fees reduces the return on investment for American taxpayers. If APC prices continue to rise, we recommend USDA reconsider whether these fees should continue to be allowable expenses.

#### **(4) How can USDA use partnerships to improve public access and accessibility to results of USDA-funded research?**

In addition to the comments in (2), we recommend USDA consider the following points:

- USDA could consider working with known repositories that allow researchers everywhere, no matter the institution, to share their publications. AgEcon Search (<https://ageconsearch.umn.edu/>) is an example of a repository that has been operating for over 25 years. Data is not yet included in AgEcon Search but that will be explored in the near future.
- The plan states that USDA is migrating Ag Data Commons to Figshare. Although Figshare is a well known repository, curation is a critical process for sharing FAIR data. Figshare does not enhance submitted data sets to comply with FAIR Principles.

- Additionally, Figshare is a for-profit product, which raises concerns over what fee or payment schedule they may adopt in the future. For example, Figshare Curation Service (FCS) can be added to a Figshare-for-Publishers or Figshare-for-Institutions platform subscription. However, relying on services from for-profit providers will have negative impacts on equity—only researchers with deep expertise or money to pay for curation services would be able to share high quality data.
- USDA should engage with the U.S. Repository Network (<https://sparcopen.org/our-work/us-repository-network/>) to have PubAg, Ag Data Commons, and other known USDA repositories participate. This would help increase access to materials and enable accessibility to *share* research data for all researchers, not just those with USDA grant funding.

#### **(4) How can USDA monitor impacts on communities impacted by its public access policies?**

We are unsure what exactly is meant by this question. There are multiple communities that will be affected by USDA's public access policy, each of which will experience different effects:

- Researchers
- Communities being researched
- Taxpayers
- Tribal communities
- Partners
- Research institutions
- Repositories

We recommend USDA clarify section 4 (Outreach and Training) of the Plan, particularly section 4.2, by clarifying who the customers and community stakeholders are, what kind of engagement is desired, and what kinds of lessons learned will be developed and shared. In reference to USDA's plan to "identify audiences and develop initial communications and engagement materials appropriate to the audiences, USDA should take care to ensure the accessibility of all materials created. For example, materials should be written using appropriate language for the audience USDA is trying to reach (e.g., not using jargon), should be made available for those who are vision impaired, and should be available for easy download (e.g., with respect to the digital divide, and differential internet access).

Sincerely,

Allison Langham-Putrow, Shanda Hunt, Shannon Farrell, and Wanda Marsolek

# References

Response to Notice of Second Interested Party Feedback Period Regarding Increasing Public Access to the Results of USDA-Funded Research

<https://www.regulations.gov/docket/NAL-2024-0001/document>

USDA Plan: [Implementation Plan to Increase Public Access to USDA-funded Research Results](#)

Authors: Allison Langham-Putrow, Shannon Farrell, Wanda Marsolek, Shanda Hunt (Scholarly communications and research data librarians, University of Minnesota)