Response by the Center for Open Science to Request for Information: NASA Public Access Plan for Increasing Access to the Results of NASA-Supported Research (Document Citation: 88 FR 31827, Notice: 23-051, Document Number: 2023-10643)

By David Mellor* and Nici Pfeiffer**

*Director of Policy, Center for Open Science, david@cos.io
**Chief Product Officer, Center for Open Science, nici@cos.io

NASA recently requested for information[1] regarding it's response[2] to the White House Office of Science and Technology Policy memo to increase access to data and papers resulting from federally supported research. NASA noted three points of its plan to: 1) enable compliance with the policy through open access paper repositories, 2) improve equity in access to publications, and 3) monitor costs associated with OA publication; and requested input on two additional topics to 4) increase findability and transparency of research, and 5) archive software. The Center for Science (COS) is pleased to offer the following perspective and recommendations concerning the items specified in the RFI.

Firstly, we agree with and support the rationale and pragmatic approach behind this "Green OA" policy, in which results of NASA-supported research must be made publicly available through any one of a number of options, including the databases that are made available. Increased use of preprint services, an encouragement by NASA to use them regardless of final publication, and preferentially linking to any OA paper (preprint or otherwise) would show additional support and normalization of these platforms and make progress toward supporting the other benefits that these methods provide. These benefits include a lack of publication bias, increased speed of disseminating findings, and decreasing the prominence of journal-based metrics or heuristics.

Recommendation: NASA should preferentially link to any free-to-read research output (i.e. the preprint of the revised and accepted article version or the OA published version) in any system that points to research outputs and ask authors or grantees to do the same.

In regards to monitoring the evolving cost of open access publishing, we realize that the trend of increasing article processing charges (APCs) represents a significant burden to the scientific research community that has, until recently, been borne by subscription costs paid for by university library systems. Shifting costs from one public entity to another will likely perpetuate the troubles of the current system. Under both systems, normal market forces cannot be applied because third-party payers bear the cost of paying additional money to access work that is often produced and reviewed by public sources. While "Green OA" are the most fair short-term strategy to enable access to outputs while not putting additional burden on research grant budgets, it remains to be seen how widespread adoption of that solution will alter this market.

---

[1] https://www.federalregister.gov/documents/2023/05/18/2023-10643/request-for-information-nasa-public-access-plan-for-increasing-access-to-the-results-of
[2] https://www.nasa.gov/sites/default/files/atoms/files/nasa_ocs_public_access_plan_may_2023.pdf

As this evolves, one critical piece of information is the amount of money spent by funders on APCs. Transparency on how much federal resources are spent on APCs will help stakeholders to determine if those costs are reasonable and, if not, provide an avenue for feedback and response to these increasing costs.

Recommendation: NASA should collect and disseminate annual costs spent on its own work and its grant-funded research for APCs. This transparency will enable researchers, as well as all taxpayers, to understand how that money is being spent and to assess the reasonableness of those expenditures.

In response to NASA's request for input on increasing findability of research outputs, we support the increased use of persistent identifiers such as ORCiDs for individuals, DOIs for digital objects, the use of the Funder Registry and Crossref's DOI grant registration workflow, and Research Resource Identifiers (RRiDs[3]) for physical objects such as geologic samples, model organisms, antibodies, and similar materials. Beginning with DOIs for NASA grants, NASA will enable the proposal process to include key metadata and PIDs at the earliest possible stage of the research lifecycle. Grant DOIs collect metadata on the grant proposal title, investigator names and institutional affiliation, along with investigator ORCiDs, project descriptions, funding amounts, project start and end dates, and funder name and identifiers. By NASA taking this first step of assigning the award with a persistent identifier and contributing this stage to the scholarly record, it signals the importance of these actions and models the practices expected by researchers. It will further advance the ability to track outputs of the award project.

In addition to the use of PIDs, good metadata is essential for findability and reusability, but is notoriously difficult to implement. In an effort to increase the applicability and ease of creating relevant metadata for research outputs, COS is working with the Center for Expanded Data Annotation and Retrieval (CEDAR) to create metadata fields that are modifiable and more easy to use. This will help to obtain the benefits of good metadata in a more convenient workflow for researchers. Basic metadata provided for DOI registration with both the Datacite and Crossref metadata schemas for the appropriate object is sufficient for tracking related outputs and relationships between objects. However, there is a different set of metadata, domain and methodology specific metadata, that serves the needs of researchers in having sufficient and useful information to allow for reusing, reproducing, remixing, and having trust with shared data and outputs. CEDAR supports domain and methodology communities in developing these specialized metadata schemas and applying them within repositories managing sharing and access to research outputs. The Research Organization Registry (ROR) is a database of funders with unique IDs that enables grant identification through PIDs.

Recommendation: NASA should register with ROR and the Funder Registry and consistently register all grant metadata, relationships and related PIDs for each award granted. In addition, the research investigators for all NASA awards should be given the grant DOI and be required to link all related outputs, including data, code, materials, and papers to the grant award DOI.

---

[3] https://www.rrids.org/

NASA should support its research communities by fostering workshops for metadata schema development to enable the rich metadata necessary for sharing, reuse, reproducibility, and overall impact of its funded research outputs.

Software has become a critical component to run research experiments, conduct analyses, and provide a turnkey instance of the research environment used in investigations. Therefore, it is essential that the software code, containers, parameters, versions, etc used for the experimental setup and analysis are provided alongside the data and results to enable replication, reproduction, and reuse of findings. The archiving and preservation of these environments, code, containers, etc. there were used must be a necessary step in the research process. Several tools exist to support this, such as Docker, and other git style repos, however, archival is not always done. OSF enables software archiving by linking the GIT repository and other components to an OSF project. By registering that project, all software developed in a GIT repository can be archived, which creates an immutable, time-stamped copy while allowing the GIT repository to continue to ch.

Recommendation: NASA should identify ways that research software can be archived in a persistent repository so that it remains immutable while also being FAIR-compliant. OSF is one tool that enables such features, but by pointing to additional platforms that enable this feature, NASA will help to preserve scholarly outputs.