**MEMORANDUM**

**TO:**  **Dr. Lisa Nichols, Assistant Director of Academic Engagement, Office of Science and Technology Policy**

**FROM:**  **Association of American Universities;** Contact: Katie Steen, katie.steen@aau.edu
**Association of Public and Land-grant Universities;** Contact: Kacy Redd, kredd@aplu.org
**Council on Governmental Relations;** Contact: Jackie Bendall, JBendall@COGR.edu

**DATE:**  **May 6, 2020**

**Re: OSTP Request for Information- Public Access to Peer-Reviewed Scholarly Publications, Data and Code Resulting from Federally Funded Research- Document No. 2020–03189**

On behalf of our organizations representing the higher education and research university community, we greatly appreciate the Office of Science and Technology Policy's (OSTP) efforts to seek comments on ensuring public access to scholarly outputs resulting from federally funded research. Stakeholder input and buy-in across the research enterprise will be imperative as we seek to develop and advance federal policies to appropriately manage and sustain the public sharing of federally funded research. Our organizations have appreciated the ongoing dialogue with OSTP and federal agencies on this matter and look forward to continuing our engagement on this topic.

Enabling public access to federally funded research results accelerates scientific inquiry and ensures research integrity; both goals are critical to enabling the scientific discoveries which advance our nation's health, drive U.S. global competitiveness, and ensure our overall well-being. The creation and dissemination of new knowledge, as well as ensuring public access to such knowledge, are core to our missions as research universities and a responsibility we take seriously. With support from NSF and NIH, AAU and APLU are convening a series of workshops, meetings, and national summits aimed at accelerating public access to research data at our member institutions. In our comments, we outline opportunities for further collaboration with the federal government and other stakeholders as well as existing barriers to enabling public access to federally funded research results.

**1). What current limitations exist to the effective communication of research outputs (publications, data, and code) and how might communications evolve to accelerate public access while advancing the quality of scientific research? What are the barriers to and opportunities for change?**

Coordination and Management
Ensuring public access to research requires significant coordination across a wide array of public and private entities and communities of practice, many of which have a variety of operating procedures (e.g. funders, researchers, institutions, scholarly disciplines, scientific societies, service providers, etc.). An added challenge

to managing research sharing efforts is the decentralized nature of academic communities. This decentralization is also reflected in the variety of public access policies across the federal agencies. Policy harmonization across agencies is needed to incentivize researchers to engage in the open sharing of research outputs and assist institutions in compliance. A possible solution is the creation of more one-stop-shop access points for researchers that integrates grantee and funder operating procedures and requirements. One illustrative example is the PASS System developed by Johns Hopkins University, which is making great strides in simplifying the reporting, sharing, and compliance components of federally funded research.

## Data Expertise and Standards

Publicly sharing research data presents technical challenges that necessitate federal guidance and coordination to ensure accessibility, quality, and efficiency. There is considerable variation across disciplines in the type and size of data produced from federally funded research. While some disciplines have adopted best practices for data sharing, curation, and dissemination, many disciplines still lack standards and consensus. The variation and lack of standards makes it more difficult for researchers to share data and for institutions to comply with federal data sharing policies. Clear, disciplinary-based data standards are needed to facilitate data sharing, especially in disciplines that lack consensus. This will require active engagement with key disciplinary societies.

Private industry employs a significant number of data experts that are needed in the public sector research enterprise. To bolster data expertise within the government and at institutions, we need investments in discipline-specific and interdisciplinary data fellowships and traineeships. Developing and incentivizing researchers with data skills to stay in the public sector will strengthen our human and technical infrastructure while providing additional opportunities for collaboration with the private sector.

## Costs

Providing public access to research requires financial investments across the scientific enterprise. The community currently lacks the human capital and technical infrastructure to enable full access to taxpayer-funded research. To meet this challenge, stakeholders, including the federal government, have an opportunity to come together to develop, build, and financially sustain the human and technical infrastructure required to fully realize the goals of public access.

Many existing business models in scholarly communication with paywalls effectively lock out potential users, beneficiaries, and universities because of excessive subscription costs. This model has proven difficult to modify because peer-review, the cornerstone of quality science, is typically housed within these existing models. Our world-class scientific enterprise has the leverage and the ability to explore new models (e.g. "open platforms/models") where peer-review is managed more directly by the academy. Rethinking current practices and models is an important step in lifting the financial barriers that currently exist to accessing research results.

Finally, access to (or creation of) repositories often comes at a significant cost. At a minimum, federally funded research should include such costs as allowable direct expenses. In many cases, the curation of data necessary before putting it in data repositories occurs after grants close, in which case the costs are not allowed as a direct charge. This is a problem that needs to be addressed to support greater sharing of data.

Openness, Security, and Privacy
We understand the need to balance openness and security within the scientific research enterprise and take seriously the responsibility to protect U.S. intellectual property when necessary. With this balance in mind, our institutions are actively working to update and strengthen campus policies aimed at supporting a research environment that broadly disseminates unrestricted research while appropriately protecting classified and export-controlled research. In addition to security concerns, our institutions also take seriously the privacy of human research subjects and their personal information and diligently work to adhere to government-imposed controls and requirements for care and handling of "Controlled Unclassified Information." In some instances, lack of consistency between federal requirements and expected practices for maintaining privacy and security can lead to confusion when trying to publicly share research outputs. For this reason, any federal policy on public access must be developed and implemented in concert with relevant privacy and security practices.

**2). What more can Federal agencies do to make tax-payer funded research results, including peer-reviewed author manuscripts, data, and code funded by the Federal Government, freely and publicly accessible in a way that minimizes delay, maximizes access, and enhances usability? How can the Federal Government engage with other sectors to achieve these goals?**

Streamlining Data Guidance and Services
The federal government has a unique role to play in meeting the challenges posed by the variation in type and function of research data across disciplines. Researchers are experts in their field but not necessarily experts in the best data sharing practices, e.g. those that align with FAIR principles. To minimize the delay in preparing data for public use, detailed guidance and user-friendly infrastructure is needed. We suggest the federal agencies solicit input from data experts, universities, and agency staff to develop and endorse specific data standards and practices. As recommended by the Government Accountability Office (GAO) in their report on public access to research results, we also suggest agencies work to establish single points of access for researchers to deposit data. Additional direction from the agencies, including in one-to-one program officer and researcher communications, is needed to support research compliance and solidify operating procedures on campus.

Developing a Data Workforce
In addition to guidance and single points of access, federal investment in developing data expertise at the agencies and in the broader research enterprise is necessary. Across the research community, a gap exists in that researchers skilled in their discipline may not be skilled in their discipline's data sharing practices. As the research community works to transition to an increasingly digital world, the development of the scientific workforce must support new traineeships and fellowships that seek to fill this gap. While technical data skills are critical, new and innovative approaches to science (e.g., collaborative/team science, open science, interdisciplinary research) require experts skilled in both the discipline and data dissemination. To bolster and develop data expertise within the government and at institutions, we recommend federal investments in discipline-specific and interdisciplinary data fellowships and traineeships.

Infrastructure
Our existing research infrastructure, both technical and human, presents challenges and opportunities in enabling public access. Our comments have outlined the difficulties posed by the variation in types of data across scientific disciplines, lack of single points of access to repositories, gaps in data expertise, and the need

for more platforms with appropriate reuse characteristics. Strategic federal investments in key infrastructure will enable and incentivize the scientific community to share research and further incorporate open science practices into the enterprise. To determine areas for strategic investments, we recommend the federal government undertake an assessment of current platforms, services, and systems that support the public sharing of research and identify areas of need. This assessment could inform sustained federal investment decisions. In parallel, we also suggest the federal government engage, through agency requests for proposals (RFPs), in pilot projects with scholarly societies, institutions, and other partners to develop, build, or sustain discipline-specific data repositories and systems that allow for interdisciplinary data analysis across repositories.

Costs

As universities endeavor to support scholarly communication in the digital age, shortcomings in our current system have highlighted the need to adapt and rethink our current models for disseminating research. The COVID-19 pandemic has further underscored the need to minimize the delay in sharing data and articles in a way that ensures quality, usability, and reproducibility. Specifically, we encourage the agencies and OSTP to support the scientific enterprise in exploring new research dissemination models where peer-review is managed more directly by the academy. This may include peer-review managed by scholarly societies or other self-organized and proven models that ensure the quality of research articles at reasonable costs. Rethinking current practices will lessen the financial barriers to accessing research results.

As detailed above, building and sustaining the appropriate research infrastructure is critical to the long-term sustainability of public access to federally funded research. Maintaining and increasing America's return on investments in science requires investment in research infrastructure. To ensure alignment with FAIR data standards, it will be important that government agencies help to play a role in establishing and maintaining such infrastructure for the core scientific disciplines for which they provide significant levels of federal support.

Openness, Privacy, and Security

To improve understanding within disciplines and across universities, agencies should provide specific guidance on the balance between public access, security, and privacy. The federal government is better positioned to indicate the appropriate balance and how it expects the research community to steward federal funds. To facilitate compliance with federal policies and enhance research quality, we recommend agencies provide clear and consistent rules and policies which appropriately balance the need for scientific openness with security and privacy, in consultation with stakeholders. We also recommend that agencies request that Principle Investigators incorporate both information about how research results will be shared (with attention to privacy concerns, etc.) or, secured, if necessary, in the data management plans that they submit as a part of their grant proposals.

**3). How would American science leadership and American competitiveness benefit from immediate access to these resources? What are potential challenges and effective approaches for overcoming them? Analyses that weigh the trade-offs of different approaches and models, especially those that provide data, will be particularly helpful.**

Scientific Innovation, Economic Impact, and Integrity

Enabling public access to research outputs allows researchers to address new questions that cross disciplinary boundaries. Tackling national and global challenges requires an interdisciplinary approach that leverages expertise and studies from across many different scientific communities. The National Academies of Sciences,

Engineering, and Medicine's (NASEM) 2018 report, Open Science by Design: Realizing a Vision for 21st Century Research, outlines how sharing research data in a machine-readable format for computational analysis speeds the discovery of new patterns and relationships that can cross disciplinary boundaries. With appropriate quality checks in place, the open sharing of knowledge will help us uncover new patterns and insights across fields. For example, Paradigm4, a data management system started at MIT before growing into its own company, allows users to draw upon open data across a range of fields. The computational platform enables researchers to analyze multiple data sets quicker than ever before at scale. The culture of data sharing underlying Paradigm4 and other initiatives undergirds the next frontier of science and medicine as we look to better understand the interconnectedness of our world.

With increased innovation and discovery comes increased economic output. A 2014 study commissioned by the Omidyar Network suggests open data has the potential to unlock $3.2 trillion in economic value annually. For the U.S. to fully realize this potential, we must invest in the tools, services, and infrastructure that enables scientists to easily share research outputs and collaborate. To effectively leverage resources, the research community and the federal government should work together in partnership to evaluate needs and costs within the enterprise.

The open sharing of research data not only promotes collaboration and innovation within the scientific community, it also helps build the public's trust in science by ensuring accountability and transparency. Sharing data allows other researchers to re-analyze and reproduce studies to test reliability and maintain research integrity. American competitiveness hinges on the public and policymaker's trust and belief in science and the research community's ability to demonstrate its value. While communicating science can be difficult, sharing data publicly helps build trust in our researchers. A Pew Research study found the majority of Americans are more apt to trust research when the data is openly available.

COVID-19 Pandemic, Public Health
As made clear in OSTP's call to publishers to make COVID-19-related research available to everyone, the current pandemic further highlights the need to share research results so medical professionals and scientists can collaborate on the development of life-saving treatments and ultimately, a COVID-19 vaccine. We are grateful for the research community's rapid response to the virus and are committed to continuing our work with the federal government and industry partners. Investments in open science infrastructure that enable public sharing of quality research outputs better prepares American science leadership to fight the coronavirus and other public health threats.

**Any additional information that might be considered for Federal policies related to public access to peer-reviewed author manuscripts, data, and code resulting from federally supported research.**

The implementation period of any new public access policy must appropriately consider the down-stream effect on universities, scientific and disciplinary societies, and the scientific enterprise more broadly. Federal agencies should consider the significant changes university practices and academic culture will have to undergo to adapt to new policies in a way that ensures integrity and quality. Achieving such cultural change will not happen overnight. Appropriate time will be needed to enable universities and scientific societies to develop and implement new models and costing mechanisms to ensure broad based and more immediate public access to research results. Moving too quickly to implement new government-wide public access policies could have a damaging, as opposed to positive, effect on universities' ability to conduct and effectively disseminate new scientific knowledge generated by their faculty and students. Continual engagement with the university community during the implementation of any policy will be critical.