RFI Response: Public Access

Subject: **Request for Information: Public Access to Peer-Reviewed Scholarly Publications, Data and Code Resulting from Federally Funded Research**

Response from: American Meteorological Society
45 Beacon Street
Boston, MA 02108

Submitted by: Dr. Keith L. Seitter
Executive Director
E-mail: kseitter@ametsoc.org

The American Meteorological Society (AMS) welcomes the opportunity to comment through this OSTP request for information on the public access to peer-reviewed scientific information.

The issues associated with providing public access to peer-reviewed publications, data, and code are quite different and AMS feels each should be addressed independently.

Peer-reviewed scholarly publications

Most scientific publishers, and certainly scientific society publishers such as AMS, having been moving toward more open access of the scholarly journals for many years. Publishers have been converting existing journals to open access, as well as launching new journals that are open access. The most highly ranked publication published by AMS, the *Bulletin of the American Meteorological Society*, has made all its peer-reviewed content open since 1997. Further, many scientific journals, including all AMS journals, have adopted a hybrid publishing model that allows articles to be open at the time of publication after payment of a OA fee, rather than 12 months after publication as mandated by U.S. funding agencies. Further, many publishers, including AMS, have been generous in allowing content to be placed on institutional open access repositories earlier than 12 months after publication. AMS has gone a step further to encourage — and provide the permissions to allow — those repositories to provide the final published form of the article rather than the accepted manuscript to avoid having two versions of the research results available.

Nearly all scientific societies have expressed an explicit goal of increasing the open availability of the research they publish and have been actively working toward sustainable business models that would allow that. At the same time, those society publishers have concerns of losing the important value to the scientific community that is provided by having subscription revenue in addition to author charge revenue:

- Subscription revenue allows author charges to be kept low while still providing the resources needed for high quality publication, with that quality referring to both the rigor of the peer-review process and the editorial quality of the final published form.

- Subscription revenue provides a source of funding that allows a generous author charge waiver policy for those authors who do not have institutional support to cover publication fees. Our experience is that this situation is not limited to scientists from developing countries — though it certainly includes those scientists — but also U.S. researchers who have recently completed degrees or for some other reason have ended up at a new institution that will not provide funding in support of publishing research completed somewhere else.
- With a dual funding source of author charges and subscription revenue, societies can look to publications generating excess revenue that can be used in support of the scientific community. For many societies, this is the only source of funding for some critically important programs that serve the scientific enterprise.

Despite these extremely important considerations, society publishers like AMS have been moving as aggressively as they can toward new business models that provide for more open access publication of research results in ways that are sustainable.

We note here that we have some concern that business models that depend solely on author charges have unfortunate incentives compared to those that include subscription income. When a publisher must depend on subscription income, there are strong incentives to be as selective and rigorous in peer review as possible so that the quality of the published research is beyond question and worthy of a subscriber's funds. When author charges are the sole source of income, there is an implicit incentive for publishers to accept more papers for publication since each provides additional revenue, and even publishers with the best of intentions may relax standards somewhat to meet budget realities. (And, of course, the so-called "predatory publishers" will knowingly forego the quality of the science in order to maximize profits.)

Given the above discussion, the primary recommendation of the AMS is two-fold:

1. **Embrace as fully as possible the hybrid journal model that allows individual articles to be provided as open access on publication with the payment of an additional open access author charge, and commit to having federally funded research cover that additional charge (with reasonable limits set for it).**
2. **Allow publishers in the physical sciences to continue evolving their business models over the next few years (ideally five) toward sustainable fully open access journals, and provide for a longer timeframe for publishers covering the social sciences where existing cultures and best practices do not include the author charge practices that have become ubiquitous in the physical sciences.**

The above recommendations do not resolve all the barriers toward sustainable open access of peer-reviewed research results. It should be recognized that specific publishers and specific journals may have unique circumstances that impede the transition toward a fully open access model. Some of these special circumstances were revealed in the publisher meetings held by OSTP in early 2020. AMS would hope that an overarching value of ensuring benefit to the scientific community would always override seeking a "one size fits all" solution.

Data

AMS has a formal Policy Statement on "Full, Open, and Timely Access to Data" (see: https://www.ametsoc.org/index.cfm/ams/about-ams/ams-statements/statements-of-the-ams-in-force/full-open-and-timely-access-to-data/), and in that statement AMS reaffirms its "commitment to a policy of full, open, and timely access to data that are critical to the advancement of atmospheric and related sciences, the provision of products and services for the benefit of society, and the promotion of commerce and private-sector activities. Adopting such policies could accelerate scientific discoveries, broaden and enhance participation in scientific enterprise, promote entrepreneurship, and benefit society."

As noted in the statement:

> AMS encourages its stakeholder communities to provide full, open, and timely access to environmental data and derived data products, as well as all associated information necessary to fully understand and properly use the data (metadata). In this context, *full* means that all data and metadata should be available, *open* means that it should be available to anyone who requests it, and *timely* means that it should be available as soon as possible, particularly in the case of data critical to human health and safety. These data are at the foundation of efforts to ensure public safety and national security, as well as efficient management and use of weather- and climate-sensitive sectors and systems such as water resources, transportation, and agriculture. Environmental data are used to protect critical infrastructure and support scientific publications, and they are essential for routine and high-impact weather forecasting and warning and climate monitoring. "Data," in this statement, refers to entities or outputs used as evidence of phenomena for the purposes of research or scholarship. The spectrum of data is diverse and includes in situ and remotely sensed observations, environmental predictions generated by numerical models, and data products derived from integrations of observational and model-generated sources.

All publications should follow a comprehensive data archiving and access policy (see, for example, the AMS policies https://www.ametsoc.org/index.cfm/ams/publications/ethical-guidelines-and-ams-policies/data-policy-and-guidelines/)

Specific disciplines may face additional challenges associated with making data broadly available. For example, the datasets associated with environmental observations can be so large as to be difficult to manage in traditional ways. In addition, observational datasets may undergo multiple levels of quality control and adjustment processes, with the final research results having dependencies on the processes used in each step, making it difficult to know which versions of the datasets need to be archived (from raw unprocessed to fully quality-controlled) in order to fully meet the needs of other researchers. Finally, it may not be possible to archive all data in perpetuity due to the sheer magnitude of these files. It will be important for OSTP to work with agencies to develop guidelines as to which datasets should be permanently stored and which ones should be archived for an agreed fixed period of time.

There are often significant costs associated with the collection, creation, storage, value-adding, and distribution of data. AMS has noted that it is therefore reasonable for commercial data providers to charge for access to their data at rates commensurate with their value and costs.

Government agencies, academia, and private-sector entities face additional challenges in meeting the goals of open data. Some of these challenges include policies that inhibit the widest possible use of data, such as resource constraints, intellectual property rights, ethical and legal constraints related to human privacy or personal data ownership, and inadequate infrastructure to properly maintain and administer datasets. As one example, the European Union's General Data Protection Regulation (GDPR) regulates the use of data that may be directly or indirectly related to individuals from the European Union.

The AMS recommends:

A set of principles and recommendations outlined in the AMS Policy Statement "Full, Open, and Timely Access to Data" (see: https://www.ametsoc.org/index.cfm/ams/about-ams/ams-statements/statements-of-the-ams-in-force/full-open-and-timely-access-to-data/**), which seeks to balance the many challenges in ways that align with the growing international movement to enable scientific data to be FAIR: Findable, Accessible, Interoperable, and Reusable** (Wilkinson, M. D., and Coauthors, 2016: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. https://doi.org/10.1038/sdata.2016.18).

Code

Making the computer code used in a research study freely available offers its own challenges. These include:

- The code may be specific to a particular hardware or operating system configuration that is not easily replicated.
- Given the rate at which hardware and system software is updated and replaced, the code used for a particular study may not be executable on currently available systems even a short time after the research project has been completed.
- Sections of the program may include proprietary code that the researcher is not licensed to share.
- Even if best practice software standards are followed in the documentation of the code, its use may require levels of expertise that are not widely available in other research groups.

These challenges are longstanding and not easily overcome, and any policy on making code openly available to other researchers must recognize and accommodate the limitations in achieving that goal. It is unreasonable to place as burdens on researchers the requirement that shared code must be maintained into the future as executable code, or even that the researchers are responsible for providing anything more than industry standard documentation on the code.

AMS recommends:

Policies covering the open availability of code used in research studies should require the code to be documented following industry standards, including documentation on the hardware and system software used to execute the code. The policies should not require researchers to provide any additional assistance in using the code or modifying it to be executable on another system.