



## ASSOCIATION OF AMERICAN UNIVERSITIES

The Association of American Universities (AAU) appreciates the opportunity to submit comments to the Office of Science and Technology Policy (OSTP) in response to its Request for Information, “Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research.” AAU is an association of 59 U.S. and 2 Canadian universities distinguished by strong programs of research and graduate education. AAU universities are major contributors to the international scholarly publishing system as well as primary consumers of the products of that system.

The comments below draw substantially on the public access recommendations of the Scholarly Publishing Roundtable.<sup>1</sup> The Roundtable was created in June, 2009, by the House Science and Technology Committee in cooperation with OSTP to develop consensus recommendations for expanding public access to the journal articles arising from research funded by agencies of the U.S. government. Sec. 103 of the American COMPETES Reauthorization Act (P.L 111-358) reflects a number of the recommendations of the Roundtable report. AAU strongly supports the Roundtable recommendations and the provisions of Sec. 103.

**(1) Are there steps that agencies could take to grow the existing and new markets related to the access and analysis of peer-reviewed publications that result from federally funded scientific research? How can policies for archiving publications and making them publically accessible be used to grow the economy and improve the productivity of the scientific enterprise? What are the relative costs and benefits of such policies? What type of access to these publications is required to maximize US economic growth and improve the productivity of the American scientific enterprise?**

The government-university partnership forged during World War II has generated extraordinary benefits to the nation. The Federal government has invested substantially and effectively in university research and graduate education through competitive, merit review processes; universities in turn have built facilities, recruited faculty, conducted high-quality research, and educated successive generations of students who carry forward the U.S. research enterprise in academia, industry, and government. More than 50% of economic growth since World War II has been due to technological advances, many of which have stemmed from scientific, medical, and engineering research at the nation’s research universities.

The benefits of university research are conveyed primarily through the broad dissemination of high-quality scholarly publications drawn from that research, and dissemination is embedded in the mission of the university. The current system of scholarly publishing has been extremely successful in producing and distributing high-quality, peer-reviewed publications to those who can benefit from and build upon that information. However, the globalization of science

---

<sup>1</sup> A copy of the Scholarly Publishing Roundtable report may be found at [http://www.aau.edu/policy/scholarly\\_publishing\\_roundtable.aspx?id=6894](http://www.aau.edu/policy/scholarly_publishing_roundtable.aspx?id=6894).

research and the dramatic increase in the volume of research and research publications has put strains on this system. From the university perspective, these strains have been felt acutely in the increasing pressure on the serials budgets of research libraries as those libraries struggle to maintain access to this rapidly expanding body of journal literature. Smaller institutions, start-up companies, and independent scholars face even greater challenges to maintaining access to research publications.

The extraordinary advances in digital technologies and communications capacities provide important new opportunities to both increase access to and reduce the cost of scholarly publications. Federal agency public access policies can exploit these advances to provide free and open access to the results of research that they fund. However, such policies must be constructed in ways that sustain the capacity of publishers to maintain publishing quality and integrity as they incorporate digital technologies into their operations and evolve their business models accordingly. One clear example of such a policy is the inclusion of embargo periods between the publication of journal articles in peer-reviewed journals and the availability of those articles or their final accepted manuscripts in freely accessible public access repositories. The shorter the embargo periods, the greater the benefit to the public; but such embargo periods need to be of sufficient duration for subscription journal publishers to recover their costs of publishing.

It is important to insert a basic policy statement at this point: federal public access policy must distinguish between the *cost* and *price* of publishing. The real costs of publishing must be met to maintain the essential quality and integrity of scholarly publishing. However, there is ample evidence that some publishers — both non-profit and commercial publishers — have employed pricing policies designed to generate revenue for other purposes — to provide funding for the operation of their societies in the case of some academic or professional society publishers, or to generate exorbitant profits for their stockholders in the case of some commercial publishers. The *costs* of publishing are real, and they must be met to sustain the essential and substantial value-added properties of scholarly publishing. But publishing *prices* that greatly exceed costs in a largely publicly funded enterprise intended to benefit the society that provided those funds are not justifiable and should not be accommodated in federal public access policies. In addition to consideration of the equities of a largely publicly funded enterprise, it is a stark financial reality that universities and their libraries cannot continue to subsidize activities or objectives external to scholarly publishing. The distinction between cost and price is, of course, not clear-cut, but a good-faith effort by all parties involved in the scholarly publishing enterprise to develop public access policies based on real publishing costs will advance the shared goal of advancing scholarship and, thereby, benefiting society. Doing so will require publishers not to inflate the calculation of publishing costs. Doing so also will require consumers and other sectors of the scholarly publishing community not to understate those real and necessary costs.

As noted above, the funding of research and its broad dissemination has been a potent spur to innovation and economic competitiveness. Perhaps the most promising means of dissemination is open access publishing, where the costs of publishing are met at the front end of the publishing process, so that the final, peer-reviewed research report is freely available to all. One of the most challenging aspects of moving from subscription journal publishing to open access publishing is creating a sustainable source of front-end revenue such as publication fees paid through federal grants or institutional funds, or revenue from other funding sources. Where open access publishing is not feasible, at least in the near term, expanding public access through procedures that protect necessary revenue streams to meet publishing costs, such as appropriate embargo

periods for subscription journal articles, can expand access to peer-reviewed articles and thereby expand the benefits of new knowledge.

Given the complexity and diversity of the scholarly publishing system and the number of participants in that system, some of the most promising government policies to expand access to scholarly publications and simultaneously reduce the cost of that access may involve public/private collaborations that engage multiple stakeholders in the highly interdependent system of the production, dissemination, and preservation of scholarly publications. Creative public/private collaborations can extend the benefits of federal agency public access policies by connecting repositories constructed to support rich content interoperability and reuse both within and across such government and non-government repositories. The benefits of these collaborations to U.S. economic growth and the productivity of the American scientific enterprise will be extraordinary.

**(2) What specific steps can be taken to protect the intellectual property interests of publishers, scientists, federal agencies, and other stakeholders involved with the publication and dissemination of peer-reviewed scholarly publications resulting from federally funded scientific research? Conversely, are there policies that should not be adopted with respect to public access to peer-reviewed scholarly publications so as not to undermine any intellectual property rights of publishers, scientists, federal agencies, and other stakeholders?**

Copyright law is the principle mechanism providing intellectual property protection for scholarly publications. Copyright belongs initially to the authors of journal articles, but authors of research articles rarely receive revenue from their articles, and their primary interest is in the dissemination of their research results. Accordingly, they typically assign their copyright to publishers, who are able with that copyright to carry out the peer review, copyediting, production, and dissemination of journal publications and to recover the costs of those activities. Increasingly, authors are negotiating agreements with publishers to retain certain rights for teaching and research purposes while assigning to publishers the rights necessary for them to carry out their publishing activities. The use of Creative Commons licensing provides the flexibility for authors and publishers to specify varying degrees of protection and access to publications. Frequently, authors are interested in maintaining appropriate attribution and article integrity but want to encourage the broadest dissemination and freest use of their articles beyond those conditions, and Creative Commons licenses provide an effective means of doing so.

What should be avoided with respect to public access scholarly publications is federally mandated access under circumstances that impair the ability of publishers to recover their publishing costs or that contravene the legitimate interests of authors. Such circumstances would include embargo periods that are too short for publication cost recovery.

**(3) What are the pros and cons of centralized and decentralized approaches to managing public access to peer-reviewed scholarly publications that result from federally funded research in terms of interoperability, search, development of analytic tools, and other scientific and commercial opportunities? Are there reasons why a federal agency (or agencies) should maintain custody of all published content, and are there ways that the government can ensure long-term stewardship if content is distributed across multiple private sources?**

Federal agency public access policies and the standards associated with them should be centralized to the extent necessary to support interoperability across agency repositories while retaining sufficient decentralization to allow individual agencies to work with their external constituencies to develop policies and procedures designed for their particular missions and the needs and interests of their constituents.

One of the most challenging aspects of the digitization of scholarly publications is the critical need to address the long-term preservation of digital content. The Library of Congress plays the primary government role in preservation by maintaining a central repository for printed copyrighted and public domain works, and that repository was extended in 2010 to include “born-digital” journals. The National Library of Medicine has managed preservation of and access to biomedical literature for 175 years. NIH’s recently adopted Public Access Policy and its creation of PubMed Central provide important mechanisms not only for access to, but also for preservation of, digital content.

Federal agencies acting on their own, however, cannot be the custodians of all published content, since much of that content will fall outside their purview due to the limits of their missions, the constraints of U.S. copyright law, the international scope of scholarly content, and the pragmatic realities of the highly diverse system for the creation, publication, distribution, and management of scholarly publications. But agencies can and should pursue policies and procedures to maintain effective, long-term custody of the published content arising from research they have funded — if not directly, then by contractual or collaborative arrangements with non-governmental entities.

A number of publishers and universities are working on solutions to the functional problems of creating mechanisms for the reliable, sustainable long-term preservation of digital content. Portico, LOCKSS, and CLOCKSS are among the available preservation tools for digital content, but a significant amount of digital scholarly content is not covered by these or other preservation tools. The federal government might be able to provide assistance in addressing the continuing challenges of preservation by implementing a digital preservation program that would provide funding for research on preservation of digital content and provide a forum through which key parties engaged in this effort could coordinate and, where appropriate, integrate their separate initiatives.

**(4) Are there models or new ideas for public-private partnerships that take advantage of existing publisher archives and encourage innovation in accessibility and interoperability, while ensuring long-term stewardship of the results of federally funded research?**

See comments to (5) below.

**(5) What steps can be taken by federal agencies, publishers, and/or scholarly and professional societies to encourage interoperable search, discovery, and analysis capacity across disciplines and archives? What are the minimum core metadata for scholarly publications that must be made available to the public to allow such capabilities? How should federal agencies make certain that such minimum core metadata associated with peer-reviewed publications resulting from federally funded scientific research are publicly available to ensure that these publications can be easily found and linked to federal science funding?**

As noted above, public/private partnerships through which federal agencies, nonprofit and commercial publishers, and universities and libraries develop common standards and procedures supporting full-text interoperability can dramatically expand access to and use of the results of federally funded research. Some of the barriers to fulfilling the potential of such partnerships are technical, but the most difficult — yet solvable — obstacles are procedural: negotiating mutually acceptable terms and procedures across different sectors with common but also differing interests and roles in the conduct of research and the production, dissemination, management, and preservation of research publications.

OSTP should consider using the collective input to this RFI as the basis for creating an ongoing forum that brings together all the key stakeholders in the scholarly publishing system to discuss the terms and conditions for public/private partnerships supporting interoperable search, discovery, and analysis across disciplines and archives. Some potential components of such partnerships now exist or are being developed.

Universities are creating institutional repositories of faculty-produced content — not only research articles, but conference proceedings, teaching materials, and much more. Research libraries have tremendous expertise in the acquisition, organization, dissemination, and preservation of information. Universities and their libraries are collaborating to create richly interoperable repositories of scholarly content. The HathiTrust is a partnership of more than 60 major research institutions and libraries working to ensure preservation of and access to the cultural record. HathiTrust currently houses more than 10 million volumes, including more than 5 million book titles and more than 250,000 journals. Cornell University hosts ArXiv, a freely accessible archive of more than 700,000 digital article preprints in physics, mathematics, statistics, computer science, quantitative biology, and quantitative finance.

CrossRef is a nonprofit, independent organization founded in 2000 by a group of scholarly publishers to create a journal-reference linking service that allows scholars to search and link to over 50 million articles identified by a unique Digital Object Identifier (DOI). With the participation of thousands of publishers and libraries as members and affiliates, CrossRef is providing efficient and reliable citation linking across publishers, with URL pointers to the full text of articles. In 2010, CrossRef implemented CrossMark, which certifies the final published version of scholarly articles. ORCID — the Open Researcher & Contributor ID — is creating a central registry of unique identifiers for individual researchers and a linking mechanism between

ORCID and other current author ID schemes. Publishers have been discussing with federal agencies possible ways to link journal articles to federal funding sources and agency grant reports.

These are only some of the initiatives that can be interconnected and expanded to bring the growing corpus of scholarly information together in organized, accessible ways that will greatly enhance access to and use of that information.

**(6) How can federal agencies that fund science maximize the benefit of public access policies to US taxpayers, and their investment in the peer-reviewed literature, while minimizing burden and costs for stakeholders, including awardee institutions, scientists, publishers, federal agencies, and libraries?**

Although U.S. taxpayers benefit in many ways from direct public access to peer-reviewed scholarly publications, the greatest benefit to those taxpayers is through the enhancement of the process of research and development that results from providing scientists and scholars with broader and faster access to the peer-reviewed literature. Providing free public access to peer-reviewed journal articles as soon as possible after publication will be particularly beneficial for scientists and scholars who lack initial access to scholarly publications, but public access repositories, particularly interconnected and interoperable repositories, will benefit those researchers with initial access as well.

It will be extremely important to develop common submission procedures across federal agencies to reduce the burden and cost of contributing to agency public access repositories for universities, publishers, and other entities submitting peer-reviewed content to those repositories. Providing incentives or mechanisms to encourage publishers to submit articles on behalf of their authors, particularly the final published versions of those articles, will both enrich the content of the repositories and reduce the burden on individual researchers and their institutions for submission. Such incentives might include providing links back to the publisher's website or including publisher identification information in repository articles. As digital scholarly content becomes more varied and dynamic, it will be important for federal agencies to work with authors and publishers to develop ways to update repository content as research articles and their data are updated or otherwise modified over time.

**(7) Besides scholarly journal articles, should other types of peer-reviewed publications resulting from federally funded research, such as book chapters and conference proceedings, be covered by these public access policies?**

Providing access to all peer-reviewed publications that report on federally funded research would clearly be beneficial in facilitating access to the full corpus of information upon which new knowledge is built. However, different forms of publications will require different policy treatments. As noted earlier, scientists and scholars typically have no financial interest in journal publications; their interest lies in the broad and rapid dissemination of their work, both to advance their disciplines and to receive recognition for their work. The financial interest in journal publications lies with publishers, who must recover the costs of publishing. For books, however, both the author and publisher typically have a financial interest in the product, the financial value of a book often extends for a longer period than does that of a journal publication, and the roles of authors and publishers in book projects may differ significantly from their roles in journal publishing. Conference proceedings are highly variable by discipline. Nonetheless, providing public access, perhaps under different terms for different categories of content, would

benefit science and scholarship by bringing together a broader array of information into one location.

**(8) What is the appropriate embargo period after publication before the public is granted free access to the full content of peer-reviewed scholarly publications resulting from federally funded research? Please describe the empirical basis for the recommended embargo period. Analyses that weigh public and private benefits and account for external market factors, such as competition, price changes, library budgets, and other factors, will be particularly useful. Are there evidence-based arguments that can be made that the delay period should be different for specific disciplines or types of publications?**

In general, the shorter the embargo period, the greater is the benefit to consumers of the content. However, as noted previously, the embargo period must be long enough for publishers to recover their costs of publishing. The Scholarly Publishing Roundtable discussed this issue at length and examined a considerable amount of publisher data on the citation “half-life” of journal articles. Clearly, the length of time that articles are cited varies by discipline. In the end, however, the Roundtable thought that the data were not dispositive and that developing discipline-specific embargo periods would lead to unworkable public access policies, particularly for agencies such as NSF that fund research in a wide range of disciplines. Differentiating embargo periods by broad categories of disciplines might be feasible and useful, but the best approach for each federal research funding agency developing a public access policy likely will be for that agency to negotiate its embargo period or periods with its constituents, seeking the most appropriate balance between the benefits of short durations and the necessity of durations of sufficient length for cost recovery.