

Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally funded Research

From: Patricia A. Steele
Dean of Libraries
University of Maryland, College Park

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

First, we must accept the premise that open is the future. Scholars will accept no less. The greater the access, the better chance that markets will emerge. It is impossible to project who will take a spark from information available freely – information that came from the public and should thus return to us. Information is the basis of growth thus; this access is determinant factor for economic growth and success. It encourages new thinking, interdisciplinary connections and innovation. One aspect to consider is the concept of the citizen scientist that is beginning to express itself more clearly.

Using studies such as the Houghton Report one can project a 5-8 time ROI on the relatively modest investment (\$3-5 million) in opening access to federally funded research. A \$1.5 billion impact on the economy from this investment is not out of scope in light of the studies and experience. If the investment were measured against the number of users – over 500,000 daily, the cost per search is very efficient. When one considers the potential use of the information so gleaned, value increases. Of course, the information should be fully free and able to be re-used. When one has a public commodity of such demonstrated value, any impediments to its access and use are non-sensical and destructive to the public good.

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 that result 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 or other stakeholders?

Faster availability, use of Creative Commons CC-BT licenses, embargo periods, all need to figure into the mix that permits the government to provide good access and computational use across all aspects of the record, including data sets. The better the use enabled, the greater ROI.

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?

In relation to this public corpus, the Federal government has responsibility. Those of us in academic research libraries understand that we have responsibility now and into the decades for the preservation and access to materials under our care. This means mirror sites, state-of-the-art digital curation and creative partnerships to leverage our investment. As a community, we know that we have not reached even a fraction of the preservation we must accomplish. (Witness the Columbia-Cornell study that revealed only 15% of their digital assets was being preserved in the existing infrastructure.) Nothing less can be asked of the government for these materials under their care. Costs are negligible and need for full service is critical.

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?

As I noted above, academic libraries are fully committed to our role in long-term access and preservation. We have formed the HathiTrust as a community effort that is managed and controlled by academic libraries. It is but an example of the emerging, interdependent network under development. Part of that network involves multiple sites, a variety of architectures and strong governance that focuses on expansive use and preservation over the decades. It would seem reasonable that working with academic libraries in these efforts would leverage investments on both sides.

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 was discussed regarding preservation, metadata access has received much attention and is becoming more and more sophisticated. One assumes machine readability/interoperability, controlled vocabulary, the coupling of data with the publication, etc. NISO and others continue to develop in this area and standards such as Dublin Core, OAI-PMH, and Europeana Semantic Elements exist and can serve as basis for the work still to be done. Leveraging with the suitable actors within the context of the target of full open access must undergird this work.

6. How can Federal agencies that fund science maximize the benefit of public access policies to U.S. 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?

Since libraries have such a stake in this endeavor, working in concert establishing policies that promote better use can be effective – research profiles, tool creation, use measurements, and the like. Many of us have technology institutes that are working on the same issues and reflect the emerging policy needs. Our common goal is the increased, effective use of the repositories. The better integrated with our campus research offices and the grant process, the more successful. One outcome that I took from the recent Berlin Open Access Conference was how important it is to demonstrate to the public the specific values of open access. In this case, how access has made a difference to specific researchers. Public libraries call these “stories” but they are important. The government is on the side of the angels in this initiative and needs to get the proper credit.

7. Besides scholarly journal articles, should other types of peer-reviewed literature, while minimizing burden and costs for stakeholder, including awardee institutions, scientists, publishers, Federal agencies, and libraries?

Conference proceedings, chapters and other educational materials that result from federal funding should be available. Clearly the established model must be different since, in some of these cases, authors do receive some payment as opposed to journal articles. Goal should be access for the public to its research and resources.

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?

The standard embargo period seems to be up to 12 months. This assumes that an embargo period is necessary. Ideally, there would be none, but moving from that ideal world, there is no evidence that there need be any differential on time period to reflect various disciplines. There has been evidence in the business world that giving away information, such as Google, attracts users and thus an evolving business model. Certainly the experience of the Royal Society that discovered that they derived less than ½ of 1% of their revenue from their extensive backfiles, is evidence that one cannot make assumptions about business models. There are many factors that need to be taken into consideration in any discussion of embargo periods. Libraries long have struggled with the cost of the journal literature and monitor the production of articles in a discipline, the cost and cost history, our own budgets, bundled practice, and similar factors. Also to be figured in this analysis is the percentage of journals supported by federal research.