

## **Comments of the Medical Library Association and Association of Academic Health Sciences Libraries**

### **Re: OSTP Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research**

**A Notice Distributed by the Science and Technology Policy Office on 11/04/2011**

**Submitted December 23, 2011 via email: [publicaccess@ostp.gov](mailto:publicaccess@ostp.gov)**

These comments are submitted on behalf of the Medical Library Association (MLA) and Association of Academic Health Sciences Libraries (AAHSL), and respond to questions 1, 2, 3, 4, 7, and 8.

The Medical Library Association (MLA) (<http://www.mlanet.org>) is a nonprofit educational organization with 4,000 health sciences information professional individual and institutional members worldwide. Founded in 1898, MLA provides lifelong educational opportunities, supports a knowledgebase of health information research, and works with a global network of partners to promote the importance of quality information for improved health to the health care community and the public.

The Association of Academic Health Sciences Libraries (AAHSL) (<http://www.aahsl.org>) is composed of the libraries of 124 accredited U.S. and Canadian schools as well as 26 associate members. AAHSL supports academic health sciences libraries and directors in advancing the patient care, research, education and community service missions of academic health centers through visionary executive leadership and expertise in health information, scholarly communication, and knowledge management.

**Question #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 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 U.S. economic growth and improve the productivity of the American scientific enterprise?

The Medical Library Association (MLA) and Association of Academic Health Sciences Libraries (AAHSL) have been able to observe firsthand the significant benefit of public access to publications arising from National Institutes of Health (NIH) funded research including its cost-effectiveness. The NIH public access policy model has proven to be cost effective, representing about 1/100<sup>th</sup> of one percent of the overall NIH budget (\$3.5-4.6/million annually based on a \$30 billion budget). The PubMed Central database (<http://www.ncbi.nlm.nih.gov/pmc/>) contains over 2,000,000 articles and current usage is now up to 1,000,000 users per day (and growing!). The spread of usage across all the articles in the database is not what might generally be expected (e.g. a small subset of articles getting the most usage), but rather consists of 99% of the articles being accessed on an annual basis. It is important to note that the Medical Library Association's premier journal, *Journal of the Medical Library Association (JMLA)* has

been an open access publication since 2001 and provides immediate access upon publication through PubMed Central (back to vol. 1, 1898!) (<http://www.ncbi.nlm.nih.gov/pmc/journals/93/>).

The 2009 Joint Information Systems Committee (JISC) report **Economic implications of alternative scholarly publishing models: Exploring the costs and benefits** (<http://www.jisc.ac.uk/publications/reports/2009/economicpublishingmodelsfinalreport.aspx>) notes that the number of authors publishing scholarly works is in many fields a small subset of the potential readers of those works. Further, (readership) of scientific research publications is much greater and more widespread than the production (authorship) of those publications. The dramatic level of usage of the openly accessible NIH information in terms of both quantity and breadth clearly demonstrates an extremely heavy demand that would bear out these points.

The JISC report further also identifies significant benefits impacting the economic and commercial aspects of public access to scholarly research, including:

- Positive impact on innovation, quality of service, and productivity particularly in healthcare, high technology and knowledge-intensive industries
- Potential for emergence of new businesses and industries (weather service information derivatives)
- Reduced search and discovery time and cost
- Reduction of the chances for duplicative research

Google Scholar (<http://scholar.google.com/>) and GoPubMed (<http://www.gopubmed.org/>) provide further examples of private investment building on federally-funded information.

In a follow-up study, authors of the JISC report examined and reported to the Scholarly Publishing and Academic Resources Coalition (SPARC) on the economic impacts of an open access mandate in their July 2010 publication, **Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs** (<http://www.cfses.com/FRPAA/>). This study concluded that the net present value gains of expanding an NIH-style policy to all other U.S. science agencies would be around \$1.5 billion, with a conservative estimate of a five times cost ROI benefit to the United States. The study was based on a six-month embargo and if there was no embargo, the return would increase to closer to \$1.75 billion. While the parameters for the models were based on a Federal Research Public Access Act (FRPAA) model and a six-month embargo, it is clear there would still be a significant cost-benefit of a variation to that model or a longer embargo period.

Numerous studies have made the case that open access shortens the research progress and efficiency in scientific areas as noted by Alma Swan, in her paper "Open Access and the Progress of Science" in the May-June 2007 issue of **American Scientist** (<http://www.americanscientist.org/issues/pub/open-access-and-the-progress-of-science>). She further posits the emergence of semantic computer technologies to work more effectively on the published research findings can track the evolution of ideas, topics, and fields and facilitate trend analysis. This provides great value in enabling research planners, policymakers, and funders to make better decisions in the interest of scientific progress and applications that derive from it.

In summary, the importance of quicker and open access to new research findings cannot be underestimated in its critical role in spurring the faster application of research, thereby triggering a faster cycle for follow-up research, new findings, and new applications. All of these in combination

provide significant economic benefits including new business opportunities, creation of new jobs, reducing duplication in research, and speeding the research and development process.

**Question #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 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?

As librarians, members of the Medical Library Association (MLA) and Association of Academic Health Sciences Libraries (AAHSL) support protection of intellectual property through copyright, including the exclusive rights that copyright holders enjoy related to the distribution and use of published materials. However, we maintain that federally funded research represents a unique situation where, in exchange for public funds, investigators are expected to make their information publicly available. This public availability is important because it allows other investigators to build on cutting-edge discoveries more quickly, promotes the free flow of ideas amongst researchers without barriers, adds transparency and further accountability to a previously ambiguous area of federal spending, and speeds up the process of translation of federally funded scientific discovery to better clinical care that everyone can benefit from.

The intellectual property interests of such diverse constituents cannot be satisfied by a laundry list of solutions. The complexity of finding a balance between needs is vital in order to: encourage a more open exchange of information from the research community, strengthen and shorten the bench-to-bedside approach of translational research, and enhance the affordability and distribution of scientific and scholarly research. Public access policies can be successfully implemented by respecting and working within the current copyright framework. One way could be by implementing appropriate licenses – such as Creative Commons' (<http://creativecommons.org/>)CC-BY license which is an attribution license that lets others distribute, remix, tweak, and build upon work, even commercially, as long as credit is given for the original creation.

The joint MLA/AAHSL statement, “Public Access to Health Information: Finding a Balance” recognizes that health sciences librarians are essential in facilitating a balanced approach to providing public access to information while maintaining ownership rights. Access to information is critical to advancing science and promoting healthy people. Protection of owner’s rights, including intellectual property, attribution, and compensation is also important in a democratic society. Knowledge is indispensable in making informed decisions in health care, education and research.

All of the comments and recommendations provided here are consistent with copyright law. Librarians have always educated faculty, students, researchers and clinicians with regard to intellectual property rights and fair use, and we believe that these efforts should continue.

In response to OSTP’s question, “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 of rights of publishers, scientists, Federal Agencies and other stakeholders?”, MLA and AAHSL believe that all policies should be reviewed and evaluated for their ability to solve even a part of this many layered issue.

Making federally funded health care information readily available to the public is vital to the nation's health and furthers research, innovation, and development of knowledge. To provide a return on our society's investment in scholarly research, health care information must also be organized, communicated, and preserved for current, historic, and future access by the public.

**Question #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?

The Medical Library Association (MLA) and the Association of Academic Health Sciences Libraries (AAHSL) maintain that making health care information more readily available to the public is vital to the nation's health and furthers research, innovation, and development of knowledge. We firmly believe that having access to timely, relevant, and accurate information is vital to the health of our nation.

For the public to obtain the most benefit from federally funded peer-reviewed publications, a centralized approach provides the most logical and efficient means to store scholarly publications. Currently, these publications are mounted on hundreds of different platforms many of which do not allow public access. Each platform requires a user to learn a new method to search and retrieve publications. A centralized approach such as that found with PubMed Central (<http://www.ncbi.nlm.nih.gov/pmc/>) allows users to search and locate articles in one centralized location. For entrepreneurs looking for ways to data mine the literature to develop scientific and technological innovations, one centralized database will allow for easy, efficient and reliable retrieval.

An example provided by a technological start-up investor highlights the frustration often experienced when entrepreneurs seek to use scholarly publications. The most recent scientific literature is locked behind pay walls controlled by publishers, and the entrepreneur investor cannot access the wealth of information funded by federal research unless he or she subscribes to thousands of journals. For most medical school libraries, the cost of subscribing to the most important journals can cost between two and three million dollars each year. This cost is beyond the reach of most entrepreneurs and start-up companies. Medical librarians, on a daily basis, must turn away scientists in their communities who want to use scholarly publications in their labs and offices due to the restriction of license agreements that do not allow access by unaffiliated individuals. Centralized management of federally funded research output would allow all who need access to capitalize on the investment that the American public has made in scientific research.

The MEDLINE database (<http://www.nlm.nih.gov/pubs/factsheets/pubmed.html>) is an example of the important centralized role the federal government has in the organization and standardization of access to the scholarly scientific literature. Begun as Index Medicus in 1879, MEDLINE is indispensable to researchers, clinicians, teachers, and students looking for both the most recent and historical health sciences research results. MEDLINE is an example of the important stewardship role that a federal agency has maintained over a long period of time to organize and make research results accessible.

Long-term stewardship of federally funded research should be the responsibility of the federal government. It is recognized that publishers add value to scholarly publications through the editorial and production process. However, the commercial publishing industry over the last several decades has been extremely volatile with numerous takeovers, mergers and acquisitions. Many publishing companies have been acquired by investment companies whose main focus is to provide a profit margin for investors. In this climate, it is easy to imagine publishers abandoning scholarly publications and their long-term stewardship if a publication is no longer profitable. There is little incentive beyond the profit margin to maintain long-term stewardship of scholarly publications.

In the world of the past, when scholarly publications were produced and stored in print form at hundreds of libraries around the world, long-term stewardship was the responsibility of research libraries. Today, libraries subscribe and license scholarly publications in electronic form and generally do not own the content to which they subscribe. In this setting where ownership of printed volumes is no longer the norm, it is crucial to have a central repository for scholarly publications maintained by the federal government acting as a steward of our nation's research output.

MLA and AAHSL believe that the safest method for long-term stewardship of published research results is to archive the results in more than one repository. However, the federal government is the most appropriate entity to provide permanent stewardship of research which it has funded through tax payer dollars. It is in a unique position to ensure that publicly-funded research articles are permanently preserved, made accessible, and useable into the future.

**Question #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?

The Medical Library Association (MLA) and Association of Academic Health Sciences Libraries (AAHSL) maintain that collaborative models built upon public-private partnerships that ensure long-term preservation and access to the results of federally funded research are a win-win strategy for everyone—publishers, research professionals, and the public. Publishers should be encouraged to participate in public/private partnership by providing approved repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles. Currently, none of the 50+ research funders with public access policies, including the Howard Hughes Institute (<http://www.hhmi.org/>), and the National Institutes of Health (NIH) (<http://www.nih.gov>), use proprietary sites as the final archive. In the case of the NIH, publishers often deposit the final published article in PubMed Central (<http://www.ncbi.nlm.nih.gov/pmc/>) on behalf of the researcher which is made available to the public at the end of the embargo period (no later than one year following official date of publication) and in a manner consistent with U.S. copyright law. PubMed Central also serves as a digital archive for the final publication. This benefits the publishers by removing the burden of maintaining the print and digital archive copies while ensuring compliance with U.S. copyright law. As publishing companies merge and shrink, the public and research community benefit from knowing that permanent access is ensured.

The digital preservation strategies being undertaken by HathiTrust (<http://www.hathitrust.org/>) provide a framework for ensuring long-term integrity of deposited materials. For example, their use of standard

and open content formats that meet community-accepted digital preservation standards, support interoperability among current and future systems, and their regular, automated system checks that verify the integrity of digital content and objects ensure quality control and long-term user access and preservation.

**Question #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?

The Medical Library Association (MLA) and Association of Academic Health Sciences Libraries (AAHSL) believe that public access policies should include book chapters and conference proceedings resulting from federally funded research. Researchers often struggle to find these materials to support their research, and these publications are not easily located even by librarians. To facilitate the research process, an index of book chapters would be beneficial as well as an index of conference proceedings freely available online for publications resulting from federally funded research. Such indexes could be used by librarians, researchers, and citizens who are searching for specific book chapters and conference proceedings.

Librarians do have access to the Online Computer Library Center, Inc. (OCLC) "ProceedingsFirst" database (through subscription), which is an index of worldwide conference proceedings dating back to 1993. However, this database is available through subscription only, so the general public does not have access to it. It would be very difficult for a citizen to locate proceedings from federally funded research via the Internet.

The same is true for book chapters. There is the OCLC Worldcat (available through subscription) that includes detailed tables of contents for some, but not all, books. However, the user must search for books, hoping that detailed content notes are available. There is not an option to search for specific book chapters. Again, the average citizen would have a difficult time locating a specific book chapter via the Internet because there is not a comprehensive index of book chapters resulting from federally funded research.

In summary, public access policies that would make book chapters and conference proceedings readily and freely available online would benefit librarians, researchers, and citizens to make it easier for everyone to locate specific book chapters and conference proceedings resulting from federally funded research.

**Question #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 Medical Library Association (MLA) and the Association of Academic Health Sciences Libraries (AAHSL) believe that immediate, free online availability of the results of taxpayer-funded research is optimal in order to maximize scientific productivity and accelerate commercial innovation. The Public Access Policy currently in effect for publications resulting from research funded by the National

Institutes of Health (NIH) allows for an embargo period of one year. Though there was resistance to this policy from the commercial publishing community, there is no evidence that library subscriptions to scholarly journals have declined because articles resulting from taxpayer-funded research are publicly available in PubMed Central. Since journal content includes many other types of articles, such as editorials, commentary, letters, etc., libraries continue to subscribe to the widest selection of journal titles that their budgets permit.

It is important to note, however, that all types of libraries face unprecedented challenges to maintain their journal collections in the face of serious budget cuts, and subscription costs that continue to rise far in excess of the general inflation rate. Most libraries have been forced to cut journal subscriptions because of declining resources. In this grim scenario, access to the results of taxpayer-funded research will be increasingly restricted to a dwindling circle of scholars who have immediate access to a well-funded library. Those who do not enjoy such access must do without.

In the case of biomedical information, the need to translate the results of taxpayer-funded research into bedside practice, preventive care and public health initiatives is particularly acute. Public access to evidence-based information about treatment advances and improvements in quality of care and patient safety helps to maximize our nation's large investment in our health care system.

Therefore, in order to encourage scientific productivity, promote the nation's health, and accelerate commercial innovation, we advocate a reduction in the current one-year embargo period for public access to the results of research funded by NIH to six months or less. It would be optimal to completely eliminate any embargo period for this essential information.

We appreciate having the opportunity to provide comments on this timely and important initiative. If you would like additional information please contact Mary Langman, Coordinator, Information Issues and Policy, Medical Library Association ([langman@mlahq.org](mailto:langman@mlahq.org)) or Gary Freiburger, President, Association of Academic Health Sciences Libraries ([garyf@ahsl.arizona.edu](mailto:garyf@ahsl.arizona.edu))