

open access

Thu 1/12/2012 3:57 PM

Amber Langston

I support open access. As a taxpayer, I'm interested in the advancement of medical science. I oppose any efforts to curtail open access as it would definitely slow down progress of medical cures, which our world sorely needs.

Please don't ban this access between researchers, general public, and others.

--

Amber



Response to Office of Science and Technology Policy
**Request for Information: Public Access to Peer-Reviewed Scholarly Publications
Resulting From Federally Funded Research**

Submitted January 12, 2012 by:
Julia Zimmerman, Dean of Florida State University Libraries
jazimmerman@fsu.edu
Tallahassee, FL

Background

Florida State University Libraries are in the early stages of exploring the role a research library plays in public access to peer-reviewed scholarly publications resulting from federally funded research. In the past 18 months the libraries have hired an E-Science librarian, an Associate Dean of Digital Scholarship and Technology Services and a Scholarly Communications Librarian who will all be working to address the concerns in the academic community surrounding adaptations in scholarly publishing. Additionally, in October of 2011, the Faculty Senate of Florida State University unanimously passed a resolution expressing support for open access in principle and faculty who chose to adapt their publishing habits. As part of this initiative and in response to the Request for Information issued by the Office of Science and Technology Policy, Florida State University Libraries offers the following statements as our recommendations.

Statement

[How can policies for archiving publications and making them publicly accessible be used to grow the economy and improve the productivity of the scientific enterprise? What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?]

Public access facilitates productivity and effective use of library resources.

Access to information is essential in contemporary society, as echoed in the mission of FSU Libraries, "...to support and enhance the learning, teaching, research, and service activities of the Florida State University by **providing organized access to quality information in all formats**, promoting information literacy, preserving information and **engaging in collaborative partnerships to disseminate ideas** to advance intellectual discovery." Supporting our faculty and researchers by offering greater, more immediate access to peer research will improve the caliber of scholarship that our institution produces. As a stakeholder on the consumption side of the publishing cycle, the research library will benefit from freely accessible and fully reusable research by transferring funding from increasing journal subscription costs to offer greater support and services to researchers. Training and resources could be put into assisting faculty with intellectual property issues and concerns, providing technical assistance in digitization and archival processes and more, all generally increasing the productivity of research. Therefore in order to fulfill our mission and support the continued excellence of our specific research community, our primary recommendation is that publications resulting from publicly funded research should be (1) made freely accessible and (2) fully reusable (3) without commercial



restriction (4) within six months of publication at most, and that a uniform policy be adopted by governmental agencies and funding bodies in support of the listed criteria.

[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?]

Public access policies simplify intellectual property concerns by providing standards.

As a member organization of the Association of Research Libraries, Florida State University Libraries align with its goal to “establish alliances and develop relationships that promote open collaboration among stakeholders in the scholarly communication system.” Building on our recently adopted open access resolution at Florida State University, the Libraries are interested in pursuing dialog with publishers, Federal agencies and researchers on the topic of intellectual property and ensuring the rights of our faculty in regards to their publications and data.

Specifically, we believe it is in the best interests of faculty to be informed that copyright to their works is fully theirs, and that publishers need only require the rights necessary to distribute their works upon permission from faculty. We recommend that universities and libraries (1) adopt open-access policies that begin with the retention of copyrights for the faculty author or scholar and outline how the work can be used, with allowances for archiving in institutional or field-specific repositories. Further, we recommend (2) the adoption of a government-wide public-access policy that also begins with author’s rights and exists as a standard from which publishers and other stakeholders can build their own copyright transfer agreements for publications.

[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? 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?]

University libraries are prepared to preserve and provide access to digital content.

Institution-specific guidelines and models for stewardship of scholarly content are necessary. As Florida State University Libraries enter the space of open archives through our newly established institutional repository, ensuring interoperability and standardized metadata is a foundational concern. In addition to universities and libraries, a federally sponsored repository based on the same standards of openness and interoperability would be a welcome entity. A federal repository would provide a central location and would be openly searchable to the public and researchers alike, with content and metadata harvested from institutional repositories. The establishment of a repository at the federal level has the potential to encourage access, produce collaborations



between publishers, funding agencies, libraries and researchers, and build on the infrastructural standardization that the web is constantly undergoing.

[Please identify any other items the Task Force might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.]

Public access to research supports the mission of libraries to provide access to information.

Finally, providing public access to peer-reviewed scholarly publications supports the ideal of an unintended audience. As a research library, one of our most valuable assets is the serendipitous discovery a student or faculty member may have when searching for materials. In our technological moment, those discoveries have greater potential and impact online, where international, underprivileged and transdisciplinary readers have the opportunity to stumble across world-class research. It is our final recommendation that the Office of Science and Technology Policy recognize access to information, including and especially the peer-reviewed research of federally funded scholarship, as a human right, and work diligently to propose and enact policies that provide that opportunity.

Remaining on the cutting edge of science and technology, as well as the arts and humanities, depends on researchers' ability to innovate based on their peers and colleagues. It is essential that research be made freely accessible and fully reusable to the public that funds it, producing two distinct outcomes: the profile of the scholar, its supporting institution or university, the funding body, and the publisher of the work are all increased; and, more importantly, public knowledge is built, allowing global citizens to increase the quality of life and mind that makes our historical position so profound.

For more information regarding this submission, contact:

Julia Zimmerman
Dean, Florida State University Libraries
Tallahassee, Florida
850-644-5211
jzimmerman@fsu.edu

or

Micah Vandegrift
Scholarly Communications Librarian
Florida State University
Tallahassee, Florida
850-645-9756
mvandegrift@fsu.edu

Introduction and Perspective

The comments and recommendations contained herein are in response to an RFI issued by the Office of Science and Technology Policy (OSTP) on November 3rd 2011 and published November 4th in the Federal Register 68518-68520, "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research."

I am responding as a relatively new small business owner with a current business plan of providing Aerospace engineering services to government agencies, commercial industry, academic institutions, and non-profit institutions. Supported by over 30 years experience working for a large company in the Aerospace industry, I have been competing, in particular, to perform studies on innovative advanced concepts and technology. This business focus which is particularly sensitive to access to research results has made me personally aware of the topic of this RFI. Prior to the issue of this RFI, I was not aware of previous OSTP public consultations (2009 and 2010) or a report issued by the Scholarly Publishing Roundtable (Jan 2010) but had wondered on numerous occasions if public access was perceived as a concern and if a path to express my experiences existed. So, I welcome the opportunity to respond to this RFI.

Comments on Current Public Access to Scholarly Publications Resulting from Federally Funded Research

There are two aspects of current public access to scholarly publications that I wish to comment on, affordability and timeliness, but before I address those I would first offer a general comment. The scope and quality of research results available through today's internet access is truly remarkable and has almost certainly not yet reached its full potential. This access is undoubtedly contributing to the advance of both science and engineering. I would like to make special note of NASA's Technical Reports Server (NTRS) as representing an honest attempt to make available NASA funded work. Also of note is Wikipedia whose numerous technical and scientific entries frequently have pdf versions of peer-reviewed articles contained in the list of references.

Having said that I would like to comment first on the cost of obtaining electronic copies of peer-reviewed scientific and technical papers from the sources that are of primary interest to me in my aerospace endeavors. Below is a table showing my primary sources, a typical cost per paper, and a brief description of the sources' breadth of publications and/or primary publication of interest.

Publishing Source	Typical Cost per Paper	Brief Description of Offerings
Elsevier ScienceDirect	\$31.50	full-text scientific database offering journal articles and book chapters from more than 2,500 peer-reviewed journals including Acta Astronautica, Planetary and Space Science, and Icarus
American Institute of Aeronautics and Astronautics (AIAA)	\$25	9 journals plus proceedings
Springer	\$34.95	33 journals in astronomy, 272 journals in physics including Celestial Mechanics and Dynamical Astronomy
American Institute of Physics	\$28	15 journals

While the cost of an individual paper or two would not seem an undue financial burden to bear for either a casual or professionally interested reader. However, it has been my experience that 10-15 papers may be needed when researching a topic. This quantity is needed to understand where the current published state-of-the-art is and what research areas have already been explored. Several irrelevant papers may be acquired since abstracts must be used to judge the relevance and significance of the contents. In a typical year I may research 5-6 distinct topics. Using an average paper cost of \$30, my yearly expense to obtain research results can easily reach \$3000. This is a non-reimbursable expense for any contracts awarded as a result of the information obtained but is often required to fulfill proposal requirements to

discuss the impact and relevance of the proposed work. As you are probably aware, individuals employed at corporations, academic institutions, and research institutions can usually obtain the same papers at discounted or no direct cost if their employer or associated library has obtained an institutional or corporate subscription. This is a distinct advantage for those individuals over the general public.

The second aspect of current public access to scholarly publications that I wish to comment on regards the timeliness of the availability of the information. Most government funded studies I am familiar with require a final report of the results to be delivered at the conclusion of the study and usually require the results to be presented at conferences or published in professional journals. There can be a significant time lag between the conclusion of a study and the publication of the results in a peer-reviewed forum. This time lag can easily one year or more. The time lag is understandable, the peer-review process does require some time and the chosen conference's date or journal publication date will generally be asynchronous with the conclusion of the research. While understandable the time lag is no less consequential to those without access to the final report or not in close association with the principal investigator. Typically the next phase of funding will have already been competed before the prior phase's results have been published.

Recommendations

I generally endorse the Scholarly Publishing Roundtable's core recommendation:

"Each federal research funding agency should expeditiously but carefully develop and implement an explicit public access policy that brings about free public access to the results of the research that it funds as soon as possible after those results have been published in a peer-reviewed journal."

And their additional recommendation that:

"Agencies should establish specific embargo periods between publication and public access. An embargo period of between zero (for open access journals) and twelve months currently reflects such a balance for many science disciplines. For other fields a longer embargo period may be necessary."

However, per my timeliness discussion above, I would modify the Scholarly Publishing Roundtable's recommendations as follows:

Each federal research funding agency should expeditiously but carefully develop and implement an explicit public access policy that brings about free public access to the results of the research that it funds as soon as possible after the research study has been completed and the results delivered to the funding agency in a final report.

Agencies should establish specific embargo periods between research conclusion/final report delivery and public access. An embargo period of between zero and four months prior to the next related solicitation should be established.

Recognizing that the final report has not been as thoroughly peer-reviewed and cannot benefit from possible post-study analyses, funding agencies could establish separate longer embargo periods between publication and public access for conferences and professional journals.

The Scholarly Publishing Roundtable's Report pointed out the size of revenues from scholarly journal publishing, approximately \$3 billion in 2008 in the U.S. market. While the publishing firms and professional societies in this market do provide a valuable service and add value, their product relies on federally funded research paid for by taxpayers. The federal government should examine the current pricing structure of the existing players in the market and determine if they are realizing a reasonable (on the order of 15% for many federal contracts) but not excessive profit. As noted in the Roundtable Report, Open Access (OA) publishers do not depend on subscription revenue because their costs are recovered up front through other revenue streams, such as publication fees, advertising revenue, sponsorships,

institutional subsidies, grants, or some combination of these. With revenue secured beforehand, these business models permit free access to and liberal reuse of published content. Such publishers may choose not to obtain rights or to retain only those rights necessary to assure the integrity and preservation of content.

In conclusion, I believe a combination of a policy of research Final Report accessibility that is rigorously applied and a restructuring of pricing practices of the publishing industry of federally funded research will result in much improved accessibility for the general public. This potentially opens up access to those who can use the new knowledge transformatively, to readers from education and the general public who seek knowledge for many purposes, and to officials and other analysts who might wish to track and assess the effectiveness of the investment of public resources. These benefits will serve to grow our economy and maximize the benefits of the monetary investment made on behalf of the U.S. taxpayer.

Submitted by:

Walter Mirczak
Sole Proprietor
Within the Heliosphere

Mirczak@Verizon.net

419 Via El Chico
Redondo Beach, Ca 90277



January 12, 2012

Ted Wackler
Deputy Chief of Staff
Office of Science and Technology Policy
Attn: Open Government
725 17th Street, NW.
Washington, DC 20502

Submitted via e-mail to publicaccess@ostp.gov

Dear Mr. Wackler,

The Society for Conservation Biology (SCB), a global community of conservation professionals which publishes *Conservation Biology*, among other journals, submits these comments in response to the request by the Office of Science and Technology Policy (OSTP) for input on the Administration's interest in enhancing public access to scholarly publications resulting from federally funded research. The following comments echo similar concerns expressed by our sister societies in the Ornithological Council, a consortium of twelve scientific ornithological societies in the Western Hemisphere. However, as an international society with many members in developing countries who would greatly benefit from increased access to scientific publications, our comments differ in that we emphasize both the risks and potential benefits of open access.

Much of the literature in SCB's journals reports research funded in whole or in part with federal funding.

We share the Administration's view that increased access to scientific information benefits society. Scientists want to increase the dissemination and impact of the information they generate. We support broad access to the scientific and medical literature and have in fact established a task force on these issues that may be able to work with your office as you consider these questions in the future. We are concerned, however, about the impact of free access on scientific societies, and in particular, the idea that one model is appropriate to all scientific publishers, regardless of size, revenue, or current publishing model.

Ensuring Fair Public Access: The issues addressed here are part of the overall debate about ensuring that we learn as a society and apply our knowledge well, with the help, more than the hindrance, of our governments. In the debate between intellectual property rights and the benefit of public access, we would expect OSTP to help resolve the matter through recommendations to the Administration and to Congress for solutions that provide fair return on scholarly investment as well as fair access to that knowledge, data or analysis in publications or symposia, for the public good. This is also the essence of copyright and patent law – weighing limits on public access just enough to encourage



investment in creativity and research while providing for public beneficial use. As science and technology evolve it makes sense to review that balance from time to time.

We are grateful to OSTP and the House Committee on Science and Technology for convening the Scholarly Publishing Roundtable. The Scholarly Publishing Roundtable report acknowledges the differences among scientific societies, but we would like to emphasize that, scientific societies serve society in many other ways — such as nurturing the development of new scientists and offering impartial expertise to guide government policy — and it is critical that enhanced access to scholarly publications not be achieved by sacrificing these other important benefits to society. We suggest options to prevent those negative outcomes.

In the public notice, OSTP asked:

(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?

We would like to begin by reminding OSTP of the costs of mandated, one-size-fits-all open access publishing. Many not-for-profit societies rely heavily on the revenue generated by the publication of journals. That revenue derives from membership dues, often with subscriptions included, and library subscriptions. We have already seen declining membership resulting from the fact that university students and faculty members have virtual open access because they can obtain online, full-content papers from hundreds of journals through their university libraries. The convenience of having one's one copy so as to avoid a trip to the library once had value; without that value, some forego membership. Library budgets at most universities and colleges – particularly the state-funded universities – have declined significantly over the past three years and that has caused a reduction in subscription revenue.

This is beginning to lead to fewer publishers and fewer papers published in the peer-reviewed literature. Papers may still be self-published, but self-publication is no substitute for peer-reviewed publications that have passed the scrutiny of expert review and editorial review. While peer-review is not perfect, its failures are relatively few and the vast majority of published papers are improved by this valuable process. Publication in established journals also increases accessibility because these papers are simply easier to find and are more likely to persist than those self-published on websites that may or may not be maintained over long periods of time.



Many not-for-profit societies typically do not have sufficient revenue to hire staff and undertake alternate activities that might generate revenue to replace the loss of publication revenue.

Recommendations to OSTP or the government to “grow existing and new markets related to the access and analysis of peer-reviewed publications that result from federally funded scientific research?”

- 1) **The Federal Government, in cooperation with other governments and other research and scientific analysis entities, such as the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES) and the Intergovernmental Panel on Climate Change (IPCC) could underwrite the creation and maintenance of the online public access websites** so that societies such as ours can make more content more freely available. SCB and many others are now engaged in a global effort to make the best possible array of publications, data, and analysis of them available to the world’s leading scientists working together to advise international bodies, governments and others concerning decisions on and about climate change (IPCC) and biodiversity and ecosystem services (the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services - IPBES). OSTP should consult with these and other international bodies and the Library of Congress concerning how US agencies and the Library of Congress might best work together to expedite the creation of on-line access to publications and related decision-support tools via the most complete and effective access to publications and related analyses.
- 2) **OSTP could work with the Smithsonian Institution, the Library of Congress and others to create an online directory of public access websites** and a mechanism to maintain orphaned open access websites.
- 3) **OSTP could work with the Library of Congress, the Smithsonian and others to create an open-access citation system**, which is essential to realizing the full value of open access scientific literature. Readers of a paper can of course determine which papers were cited by that particular paper, but the ability to find subsequent papers citing that particular paper is still limited to those able to afford access to an online citation system such as the Thomson-Reuters Science Citation Index. And more than a mere citation system, a set of discipline-specific annotated bibliographic databases would be invaluable to anyone delving into the enormous body of literature.
- 4) **Ensuring Federal Decisions Are Informed by the Best Possible Scientific Publications and the Data underpinning them:** Finally, to ensure that the Federal Government has access to the best possible evidence for Federal decision-making, OSTP should address the question of bringing that data and analysis, such as are found primarily in peer reviewed journals, even if it was not originally Federally funded, into Federal processes and data bases when it is most needed, as in the Federal Executive and Legislative decision-making processes. SCB’s final recommendation in our December 2008 Recommendations to the Obama Administration and Congress was to restore, to the extent possible, the practice of paying potential interveners, such as scientific societies, academics, public



interest groups, and others in federal rulemaking proceedings for information that would likely not otherwise be as fully available to the agency. The Carter Administration had begun to do this by 1978 (e.g., in FERC and DOE proceedings) and began to expand it via Executive Order Number 12044. This was done in part in order to avoid unnecessary litigation and to arrive at better decisions sooner and more efficiently by building better administrative records that included a wider array of expert evidence early on. A rider approved by Congress stopped some forms of that practice. Peer review is now used by Federal agencies in some situations but that does not reach as many decisions as it could. Given evidence that better decisions depend on better records of decision, and that both depend on an objective understanding of what science knows, the Administration and Congress can change that and remove other impediments and proceed wherever possible to bring the best science to government and then more fully to the non-governmental and private sectors. Rather than be a net expense, if those firms or groups of firms standing to profit from the use of a Federal resource or permit could be required to build the small cost of such data acquisition and analysis into the price of their products or services that depend on Federal lands, resources, or permits, then the costs could be internalized appropriately, rather than borne entirely by the tax-payer, and still managed by the government under public scrutiny and in the public interest. This would both increase the support and use of publications and apply them in the public interest at a minimum of net public expense.

A greater appreciation of all of the issues raised in this process could be inspired by OSTP offering to brief the Congressional Research Service and Committee staffs and the Scientific Integrity Officers of each agency on these issues and alternative actions that OSTP is considering.

The Scholarly Publishing Roundtable, an ad hoc working group convened by OSTP and the House Committee on Science and Technology (January 2010), recognized that a twelve-month embargo might not be adequate for some scientific disciplines. Protecting the revenue associated with access to what is considered current or recent content might require delaying public access for several years unless ways are found to reward the producers more fully and more quickly, including some of the steps noted above. The cited half-life of the journals published by societies ranges from a few to as many as 10 years. If societies determine that revenue loss associated with access to papers not yet available in their own fee-free archives would be minimal, they may choose to decrease the duration of the embargo.

We also wish to remind OSTP of the cost associated with publication charges. The journals published by our many societies charge very low publication fees; such as \$100 per page and many will waive some or all of the publication cost if the author is unable to pay for publication. Unlike other societies that are able to maintain relatively low page charges because membership fees are sufficient to subsidize the cost of publication, our society charges on the order of \$80 per year (we are instituting an increase to that level



now) and we offer substantially reduced rates to students and young professionals and members in the developing world. Increased page charges would erode research grants and increased membership dues would likely result in fewer members, and, in turn, reduced membership revenue. As membership revenue is a substantial part of overall revenue, this decrease could jeopardize the existence of the society.

(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?

Any public access policy must include a provision that the original copyright holder retains all intellectual property rights conveyed by law. To the extent that a publications database is maintained by a federal agency, the agency should require that those accessing its holdings read and acknowledge the intellectual property rights of the holder. These acknowledgments should be maintained by the agency providing public access and made available to the copyright holder upon request.

(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?

Had Congress wanted to mandate a central repository, it could have done so when reauthorizing the America COMPETES Act. Instead, the legislation directed the working group to look for standards to maximize interoperability and to take into account existing standards. We also note that the assumption that an agency repository will suffice in perpetuity is a faulty one. At this moment, the U.S. Geological Survey is terminating the National Biological Information Infrastructure. Some of the databases will be incorporated into other programs (though not necessarily made available to the public) and some will be lost.

That being said, there should be a registration system whereby every repository that holds federally funded papers is reachable through a central directory and a provision that if a



repository becomes orphaned, the central agency repository may take it over. Even then, the society should be permitted to first try to find another organization to maintain its holdings.

In our field, the development of metadata standards for data repositories is quite mature. From the development of the Darwin Core, first issued in 1998, to the 2009 release of the metadata standard, this body of standards now supports numerous extensions for use across organismal biology. It is recognized internationally and in wide use. Requiring this large body of literature to be deposited into a centralized database would impose an undue burden if that database uses different standards. It would also make it more difficult to retrieve data associated with the literature and vice versa.

(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?

(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?

This is an issue that our task force noted above may be able to assist OSTP in addressing but we have no specific recommendations at this point.

(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?

By making sure that agencies search, use and cite the applicable peer-reviewed journals and presentations at scientific symposia when the journal articles are not yet available, in each proposal for a Federal rule or guidance, and final Federal Register notice or available archive in support of such notice, and encouraging Congressional Committees and international bodies to do the same. And by covering, reimbursing or paying for the costs incurred and the value contributed by the authors and societies for such publications and presentations.

(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?



In our field, conference proceedings (where they exist) rarely consist of the full text of a talk along with the associated slides or other media. Proceedings are more commonly a listing of talks and perhaps abstracts. Full-text or not, they are rarely peer-reviewed even though they may refer to peer reviewed literature. Because these talks present new information in many cases not yet in the published literature and new analysis of it, it may be useful to ensure that these materials be made open access or that any society or agency maintain a public access repository for these materials **by making sure that full costs of managing that process are met by a combination of governments and other entities with little delay or on a regular, contracted basis. In fact, in recent years many of SCB's symposium organizers have been required by SCB to answer the question "Which agencies would use the information to be conveyed in your proposed symposium and how?" We have done this to help choose the most symposia that are most useful for the policy and practice of conservation. For Federal and other agencies to help support an archiving and open access system for such presentations could be a very big contribution to both scientific societies and the agencies and personnel dealing with the issues addressed in such symposia.**

(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?

As previously stated, the Scholarly Publishing Roundtable recognized that a twelve-month embargo might not be adequate for some scientific disciplines. Protecting the revenue associated with access to what is considered current or recent content might require delaying public access to some journals of some societies for several years, although SCB has many members and leader who very much want to work toward a much faster sharing of our publications with the public. The cited half-life of the journals published by some societies ranges from 4 to 10 years. They may eventually determine that a shorter embargo period will not reduce the level of paid access,. Establishing an upper limit or a sliding scale that takes into account the extent to which the society relies on journal revenue could be reasonable and fair, if these metrics are established in consultation with scientific societies.

That said, however, the best of both worlds may be achieved if OSTP can work with interested parties to find a way to help the producers recover the expected fair return on their investment more quickly. If this is the information age, then a 21st century version of the Interstate Commerce Commission to ensure fair public access to commerce in information may be warranted, and all those affected should be partners in the exercise.

Currently, there are numerous journals in organismal biology, wildlife biology, and ecology that have no public access, even for material that is decades old. This may be as much a function of the cost to convert older formats and maintain a website as it is about



the loss of revenue. Societies that do not have the financial resources to provide public access to older volumes should be given assistance to make access available.

Sincerely,

John M. Fitzgerald
Policy Director





Public access to federally funded scholarly publications

Comments submitted to the Office of Science and Technology Request for Information
January 12, 2012

To Whom It May Concern:

I'm writing today on behalf of the more than 100 organizations that comprise the Alliance for Taxpayer Access (ATA), in response to the Office of Science and Technology Policy's Request for Information (RFI) dated November 3, 2011, seeking input on the issue of Public Access to Scholarly Publications Resulting from Federally Funded Research.

About the Alliance

The Alliance for Taxpayer Access (ATA) is a coalition of advocacy, academic, research, and publishing organizations that supports open public access to the results of federally funded research. The Alliance was formed in 2004 to urge that peer-reviewed articles stemming from taxpayer-funded research become fully accessible and available online at no additional cost to the American public.

The diverse members of the coalition are committed to the general principle that American taxpayers are entitled to open access on the Internet to the articles that result from research funded by the U.S. government, and that facilitating broad access to these articles is an essential, inseparable part of our nation's investment in science. ATA members firmly believe that this (and other scientific information) should be shared in cost-effective ways that take advantage of the Internet, stimulate further discovery and innovation, and advance the translation of this knowledge into public benefits. Enhanced access and expanded sharing of information will lead to increased use of this information, and will deliver an accelerated return on the taxpayers' investment.

As 41 Nobel Prize-winning scientists wrote in an open letter to the U.S. Congress:

“For America to obtain an optimal return on our investment in science, publicly funded research must be shared as broadly as possible... As the pursuit of science is increasingly conducted in a digital world, we need policies that ensure that the opportunities the Internet presents for new research tools and techniques to be employed can be fully exploited. The removal of access barriers and the enabling of expanded use of research findings has the potential to dramatically transform how we approach issues of vital importance to the public, such as biomedicine, climate change, and energy research.”

http://www.taxpayeraccess.org/supporters/scientists/nobelists_2009.shtml

We thank the Office of Science and Technology Policy for organizing this important discussion. The Alliance shares the Administration's view that enhancing access to this information will promote advances in science and technology, encourage innovative use and application of government-supported research, and fuel commercial development and economic growth.

The Alliance supports the implementation of government-wide public access policies to facilitate the sharing of scientific results, and make this level of access a reality. To maximize the taxpayer's return on our nation's investment, such a policy should ensure that all members of the public are able to immediately access and fully reuse digital articles reporting on the results.

The Alliance supports building on the successful framework of the NIH Public Access Policy, and recommends that an expanded policy include the following components:

- ***Public access to the published results of federally funded research should be a mandatory requirement across all agencies.*** As the experience of the NIH has shown, a voluntary policy is not enough. The NIH saw less than 5% of eligible authors deposit their manuscripts under a voluntary policy. However, after the policy was made mandatory in April 2008, the percentage manuscripts deposited quickly rose to over 60%, and has continued to rise steadily since then.
- ***Articles that result from federal funding should be made freely accessible to the public immediately upon publication.*** The Alliance feels strongly that immediate access optimizes benefits from this research to the public. However, we recognize that this may not be practical in every discipline. We support the inclusion of an author-determined embargo period of up to six months as an acceptable compromise.
- ***Articles should be housed in permanent, interoperable digital archives.*** The results of federally funded research should be archived permanently, in interoperable repositories (maintained or approved by the agency) that allow this critical layer of information to be freely linked to the wealth of other publicly accessible databases.
- ***Access may be either to the author's final manuscript or to the final published version.*** The requirement for deposit of the author's raw final manuscript, rather than the final published article, makes it possible to also monetize value-added enhancements beyond what is available in the public repositories. Where the publisher allows, access to the final published version is also desirable. Where the publisher allows, access to the final published version is also desirable.
- ***The reuse rights associated with articles should be clearly articulated, and should ensure that articles can be both read and fully used.*** The Alliance supports ensuring this through the use of a license that works within the current copyright system and at most requires attribution to the author, such as the Creative Commons Attribution (CC BY) license.
- ***Articles should be presented to the public in a standard digital format that allows them to be fully read and used.*** The Alliance supports XML as the preferred standard. While we support the additional inclusion of PDF files, PDF alone is not an acceptable format, as it does not support robust enough linking and searching.

• ***Implementation should be closely coordinated across all agencies to ensure seamless compliance.*** The Alliance strongly believes that public access policies must be as closely coordinated across agencies as possible, and that multiple policies with multiple implementation requirements would result in unnecessary overhead and costs.

Our responses to the specific questions in the RFI follow below.

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 publicly 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?

Encouraging Commercialization

Articles reporting on the results of taxpayer-funded research are an important component of our nation's research output. To create the optimal environment to encourage such commercialization, the complete collection of full-text articles reporting on publicly funded research should be made immediately, freely available to the public¹. Members of the public must also be ensured the rights to fully use these articles without commercial restriction.

Enabling Open Access² to these articles will accelerate the ability of individuals and companies to construct new services and products, and ensure that the value of the public's investment in this research is fully realized. It will create a business climate where all stakeholders can apply ideas generated from this research more quickly, speeding the launch of new products, services and new markets.

The publishing industry is well-positioned to be among the primary beneficiaries of opening access to articles reporting on publicly funded research results. Creating a body of articles openly available to the public creates the opportunity for individuals and companies to add value to the research through the creation of new tools and services for searching, text mining, data mining, indexing, translation, and other services. These new services create the opportunity for new jobs (and new tax revenue) to be generated – directly stimulating both innovation and economic growth.

This kind of business development is already happening in areas where a significant number of open-access articles exist. Companies such as Mendeley, which offers integrated academic search and peer recommendations built on a collection of open-access papers, are rapidly achieving success, with Mendeley boasting a user base of over one million individuals in just two years of operation³.

¹ It is important to note that the full text of peer-reviewed articles must be made accessible -- not merely abstracts, summaries, or un-peer reviewed grant reports -- for the full value of these articles to be leveraged by the public.

² <http://www.soros.org/openaccess/read>

³ <http://www.mendeley.com/our-users/>

Additionally, a robust new market of open-access journals -- journals that make their content freely available to all users, with no restrictions for reuse other than appropriate attribution to the author -- is also flourishing. More than 7,300 open-access journals are currently being published in a broad spectrum of disciplines.⁴ Innovative companies, such as the U.S.-based Public Library of Science (a member of the Alliance for Taxpayer Access), have lead the way in demonstrating the financial viability -- and desirability -- of this new publishing model. The growth of this new market segment has been so dramatic that a new trade association, the Open Access Scholarly Publishers Association (OASPA) has now been established to help promote its further development⁵.

While creating new markets and business models in the publishing industry is one important outcome of a public-access policy, it is important to remember that it is part of a larger goal: to encourage the use of publicly funded research to spur increased commercialization in other business sectors. Creating a government-wide policy that results in an openly accessible database (or set of databases) of publicly funded articles will provide opportunities for companies of all kinds to build on this information. This is particularly true for industries such as biotechnology and the pharmaceutical industry, where the ability to interact with leading-edge research results is part of the lifeblood of the company. They (and their investors) count on these resources to be able to deploy a research and development strategy that keeps them on the cutting edge of new ideas and knowledge, so that they can translate these ideas quickly into marketable products and services.

Improving Scientific Productivity

Besides providing an environment in which commercialization can be optimized, ensuring full open access to articles reporting on the results of taxpayer funded research can also play an important role in improving scientific productivity. The research community has long recognized the opportunity that providing immediate, barrier-free, online access presents to researchers to work **faster**, by enabling them to get to research articles and incorporate new findings into their research more rapidly. In biomedical disciplines, for example, the need to rapidly collect, evaluate and understand the work of colleagues is readily apparent. Taxpayers fund basic biomedical research with the expectations that it will lead to new discoveries, and ultimately, to new treatments and cures. Expediting this process is directly in the best interest of the taxpayer.

Ensuring open access to scientific articles can also help scientists to incorporate **more** information into their work more efficiently. With the continued increase in papers generated from scientific research, enabling an open-access environment where computers can serve as a new category of reader of publicly funded research papers is essential.

In biomedicine alone there are currently more than 19 million citations and abstracts covered by the National Library of Medicine's search engine, PubMed. These include ~830,000 articles published in 2009, up from 814,000 in 2008 and 772,000 in 2007.⁶ The growth rate gives no indications of slowing, particularly as emerging economies like India, China and Brazil continue

⁴ <http://www.doaj.org>

⁵ <http://www.oaspa.org/>

⁶ <http://www.nature.com/news/2010/100127/full/463416a.html>

to accelerate their research outputs. Researchers need to be able to employ new semantic and computational tools to contextualize ideas contained in papers, identify new relationships, and significantly expand the breadth of research threads that they can effectively pursue.

A government-wide policy that facilitates the creation of an open-access environment will also allow – and encourage -- more people to participate in the scientific research process at many levels. Researchers in a variety of disciplines are already using open-access environments (for both data and publications) to help them expand their pool of collaborators in specific research areas, as well as to help create new pathways to solutions.

In Alzheimer's research, experts (led by Neil Buckholtz, chief of the Dementias of Aging Branch of the Division of Neuroscience at the U.S. National Institute on Aging, and Dr. William Potter, a neuroscientist at Eli Lilly) established the Alzheimer's Disease Neuroimaging Initiative (ADNI), a novel, public-private collaboration that posts all of its data on Alzheimer's on an open public Web site. ADNI has made thousands of brain scan images and clinical and neuropsychological data available to researchers around the world, and has generated a wealth of new research papers, as well as more than 100 new studies testing drugs that may slow or stop the disease.⁷ The ADNI model is already being replicated other areas, most notably in Parkinson's disease research.⁸

Along with increasing the sheer number of participants, an open-access research environment also increases the diversity of participants in the research process. It helps to promote access and reuse of information by researchers in loosely related (or even unrelated) fields that might not otherwise have access to the full corpus of research articles. This increases the value of our scientific research investment, by increasing the efficacy of scientific discovery. In the Autism community, Sophia Colamarino (Stanford University Medical School, and former Vice President for Research at Autism Speaks, also an Alliance member organization) has spoken eloquently from the patient advocacy and researcher funding perspectives. She notes that, because there is no routine treatment for Autism, families are routinely responsible for learning about therapies and treatments that may be appropriate for them.

Her experience has shown that while families are inundated with information from a variety of sources, what is most easily available may not always be credible. Because of the barriers that subscription and pay-per-view pay walls present, families have easy access to all but the most scientifically valid information.⁹ Providing immediate, barrier-free access to articles that report on the results of taxpayer-funded research access empowers family members and caregivers to be better, more informed advocates, and gives them a positive outlet by allowing them to participate in progress first hand. Barriers to accessing published research literature cause families to struggle – unnecessarily – to find the most rigorous data necessary to make informed decisions.

⁷ <http://www.nytimes.com/2010/08/13/health/research/13alzheimer.html>

⁸ http://www.michaeljfox.org/living_PPMI.cfm

⁹ <http://www.berlin9.org/bm~doc/berlin9-colamarino.pdf>

Costs and Benefits

Costs

The potential costs and benefits of taxpayer access to publicly funded research articles are of deep interest to the Alliance. Many helpful sources of data are available to draw on -- in particular, the data provided by the National Institutes of Health (NIH), whose successful public-access policy already ensures full accessibility to articles reporting on the results of the ~\$30 billion of basic and applied research that it funds annually.

This policy, which covers approximately one half of the total U.S. annual investment in scientific research, has proven to be extremely cost-effective. NIH reports that it costs \$3.5- \$4.6 million annually (on a total \$30-billion budget) to administer its public-access policy. This represents an investment of only about 1/100th of one percent of the NIH's overall \$30 billion operating budget to ensure that the 90,000-95,000 articles generated annually to report on NIH-funded research are readily accessible to all potential users.¹⁰

The NIH also reports a deep demand for these articles, with more than 500,000 unique users from all sectors of the public accessing the PubMed Central database each day to view and retrieve articles.¹¹ Many of these users are members of the 100+ organizations represented by the Alliance. An effective, government-wide public-access policy can likewise be implemented in a cost-effective manner, by leveraging this existing infrastructure to minimize unneeded duplication of efforts, and utilizing the investments already made by the NIH.

Benefits

Significant economic research has been done, in the U.S. as well as internationally, on cost-benefit analyses of various policy approaches to ensuring greater access to articles reporting on the results of publicly funded research. Detailed economic analyses have been conducted on proposed national policies in Australia, the U.K., the Netherlands and elsewhere, providing sound methodologies for policy makers to use in considering the potential impact of such policies. These studies have consistently demonstrated that the adoption of policies to encourage the open sharing of research results – including scientific articles – has a significant economic upside for national economies.¹²

Perhaps most germane for the purposes of this RFI is the 2010 study conducted by Houghton et al., examining the potential impacts of opening up access to articles reporting on the results of all U.S. federally funded scientific research, under a policy similar to that of the current NIH Public Access Policy. Houghton and his colleagues examined both the costs and potential returns to the public investment in R&D, and provide a working model to be used for further testing and refining estimates as additional data becomes available.¹³

¹⁰ <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>

¹¹ <http://olpa.od.nih.gov/hearings/111/session2/Testimonies/PublicAccess.pdf>

¹² <http://www.cfses.com/projects/Easi-OA.htm>

¹³ Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs, Houghton et al. (2010)

The initial Houghton et al. modeling suggests that providing open access to all articles reporting on U.S. scientific research under a model similar to the current NIH policy would (very conservatively) result in at least a five-fold increase in ROI, with the benefits of the policy estimated to be approximately 8 times larger than the costs. They further estimate that the net present value gains of expanding an NIH-style policy to all other U.S. science agencies over time would be on the order of \$1.5 billion (net the costs of running the archive). Of that number, approximately 60% is estimated to accrue directly to the U.S. economy.¹⁴ The Alliance for Taxpayer Access is strongly supportive of a policy that pursues such an approach.

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?

To best support the goals of accelerating scientific discovery, innovation and the creation of new markets, any public-access policy should ensure not only full *accessibility* of scientific articles, but also full *utility* of the articles in the digital environment.

The Alliance supports the creation of a government-wide public-access policy that works within the current copyright framework by requiring full open access to articles reporting on the results of federally funded research under a mechanism such as the Creative Commons Attribution (CC BY)¹⁵ license. This kind of an approach is consistent with protecting the copyrights of both authors and publishers.

While the NIH Public Access Policy provides an excellent benchmark for most aspects of government-wide policy, it can be substantially improved upon in the area of rights retention. The Department of Labor's Trade Adjustment Assistance Community College and Career Training (TAACCCT)¹⁶ grant program provides a more appropriate exemplar. The TAACCCT program requires that grant recipients license content created from grant funds under a Creative Commons Attribution (CC-BY) license. This framework ensures broad access and reuse for anyone wishing to utilize this federally funded research output, while also ensuring that proper credit is given to the author.¹⁷

Additionally, taxpayers also need access to these articles sooner than the current term of copyright allows. Ideally, articles reporting on the results of publicly funded research should be made accessible to the public immediately upon appearance in a journal. However, an initial interim, phased approach might prove a practical way forward. This type of approach might be constructed to include:

¹⁴ Op. cit.

¹⁵ <http://creativecommons.org/licenses/by/3.0/>

¹⁶ <http://www.doleta.gov/grants/pdf/SGA-DFA-PY-10-03.pdf>

¹⁷ <http://epsiplatform.eu/content/topic-report-no-23-creative-commons-and-public-sector-information-flexible-tools-support-psi>

- First, providing an appropriate period of embargoed access (no longer than 12 months) where current rights appropriate under copyright apply;
- Second, after the expiration of the embargo period, full reuse rights under an appropriate license such as CC-BY apply.

It should be the explicit goal of any government-wide public-access policy to make the results of federally funded research as useful to taxpayers as possible. Broad reuse allows both researchers and businesses to unlock additional value from our public research investment – now, and for decades to come. Restrictions that limit how users can work with these digital articles will result in only a fraction of their value being delivered, and unnecessarily reduce the subsequent return to the taxpayer.

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?

The federal government is the appropriate entity to provide permanent stewardship of these articles, and is in a unique position to ensure that publicly funded articles are made permanently accessible, and useable. To ensure this, any public-access policy that is developed must give the federal government adequate rights to archive and distribute articles reporting on publicly funded research. Currently, the National Library of Medicine (NLM) appropriately fulfills this crucial role for articles generated by NIH-funded research by housing articles in their PubMed Central (PMC) digital repository.

NLM has indicated¹⁸ that they are willing to expand their role and accept articles from any other federal science agency, providing an immediate, cost-effective potential solution to taxpayers. Alternatively, NLM has also indicated that the software supporting PubMed Central is freely available in the public domain, and was explicitly designed in a modular form to be easily shared with other entities that might wish to use it. This option provides another cost-effective mechanism that ensures the interoperability of multiple federal agency archives.

This type of approach does not preclude other, non-governmental entities from participating as partners in a decentralized approach. An effective federal public-access policy could involve multiple repositories maintained by third parties, as long as those repositories support access and use conditions that allow all interested parties to build on the content contained in them. Repositories that meet conditions for public accessibility, unrestricted use rights, interoperability and long-term preservation of articles can play an important role, encouraging innovative public-private partnerships.

Having the federal government retain custody of a master copy of these articles is critical in minimizing the possibility of exclusive arrangement that inhibit the ability the widest possible

¹⁸ <http://olpa.od.nih.gov/hearings/111/session2/Testimonies/PublicAccess.pdf>

community of stakeholders and businesses to use these articles, and ensure that new services and products can be readily built from them, enhancing the taxpayer's return on their investment.

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?

Public-private partnerships can play an important role in leveraging the unique capabilities of a broad range of potential service providers, and create opportunities for the development of new products and services to built on publicly funded information. A key aim of the America COMPETES Act (whose goals this RFI has been issued to facilitate achieving) is to improve the competitiveness of the United States through investment in research and development. As such, it is critical that any public-private partnerships be constructed to ensure that all potential service providers have an equal opportunity to participate. The Alliance firmly believes that under no condition should any one site, organization or company be the single point of access for taxpayer-funded articles.

This is particularly important as it relates to small businesses that may experience difficulty with entering markets given access conditions or restrictive copyright/reuse provisions. Constructing a partnership that unfairly advantages a limited number of participants will result in a less competitive environment, rather than facilitating the kind of environment that encourages robust participation by all stakeholders.

The Alliance notes that the publishing community is only one stakeholder group whose interests must be considered in this context. The federal government should also carefully consider the other potential partners, particularly libraries, archives, and higher education institutions. These organizations have experience in access to (and preservation of) information, and also have a wealth of experience and existing infrastructure that can be leveraged. Developing a public-access policy that includes roles for these kinds of organizations would greatly increase prospects for the viability and long-term sustainability of such partnerships.

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?

Metadata enables the interoperability, search, discovery and analysis of articles reporting on federally funded research, and should also be used enable specific actions that can be taken on digital articles, as well. To be as useful as possible, metadata associated with federally funded

articles must be both machine-readable and machine-interoperable, and should facilitate the robust use, reuse and analysis of digital articles.

The Alliance recommends working closely with experts in the university library and digital repository communities, as well as other expert organizations to build out a minimum core set of metadata. While we understand that Dublin Core is the current standard in this regard, we also understand that this will only enable the minimum amount of discovery and download to take place. Broader metadata specifications are needed to make full use of the information contained in federally funded articles and to active the aims of the federal government of improving scientific productivity and accelerating commercialization.

To maximize the value of this information, additional metadata is needed to also facilitate archiving and preservation, and to encourage the development of new services (such as text mining, visualizations, etc.). It is important to ensure that any metadata standard or framework not only meets current needs, but is also flexible and extensible enough to support potential future uses. This is particularly critical to ensure that connections between articles and digital data can be supported.

Close consultation with established entities that are working on standards and best practices in this area, such as NISO and the Library of Congress, will also be helpful and should be actively pursued.

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?

The benefit of public-access policies to U.S. taxpayers can be maximized by making the complete collection of full-text articles reporting on federally funded scientific research immediately, freely accessible to the public. Taxpayers also must be guaranteed rights to fully use these articles without commercial restriction. The federal government should provide long-term stewardship over the repositories that house these articles, in partnership with organizations such as libraries and archives. Access conditions and reuse rights that at most require author attribution (Creative Commons Attribution CC-BY license or similar) must also be clearly articulated, to enhance scientific productivity and encourage the full range of potential stakeholders to build secondary services and generate new products and markets from this content.

For any public-access policy to be successful, there must be consistency of requirements and of implementation across all federal agencies. Creating multiple, disparate access policies – or even compliance requirements – for different federal science agencies would introduce needless confusion and expense into the system, and greatly increase the compliance burden on the grantee and their home institution.

Uniform requirements and procedures regarding deposit of peer-reviewed articles should be established across all federal agencies covered by a public-access policy to reduce the cost and complexity of compliance.

Effective implementation strategies that minimize the burden on the researcher can also play an important role in maximizing the returns to the taxpayer, by raising compliance rates and ensuring that the complete corpus of articles reporting on federally funded research is widely available in a timely manner. This will have the added benefit of supporting informed, transparent, science-based federal budget and policy decision making by increasing federal agency accountability and providing agencies with an improved accounting on the outcomes of their funded research. It will also give Congressional budget drafters, appropriators, and authorizers better information to accurately assess the value of existing expenditures, and to target funding on the most promising research areas.

Any federal public-access policy should also be constructed in a way that encourages the development of additional tools and services to facilitate both the work of the researcher, and the federal agency. Encouraging the integration of articles with agency (and home institution) grant management systems is an important potential way to improve agency accountability, as well as to provide increased information to the public on the results of the research that their tax dollars support.

An effective public-access policy centered on creating accessible databases of research articles can also create opportunities to build productivity management tools – like enhanced bibliographies or Principle Investigator (PI) Profiles – that are of wide use to researchers, institutions, and federal agencies, further leveraging the value of the taxpayers investment in scientific research.

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 Alliance for Taxpayer Access firmly believes in the that principle that taxpayers have the right to the results of publicly funded research – and that this right applies to **all** outputs of research, from data and articles, to educational materials (book chapters, texts, conference proceedings, etc). We believe that these outputs resulting from publicly funded research should also be made readily accessible to the public.

However, we recognize that different conditions and expectations apply to different types of outputs. For example, authors are not paid for journal articles, but may in fact be compensated for the creation of book chapters. Data sets may contain confidential or personal information that may not be appropriate for unrestricted access or reuse. Access policies that reflect these differences while holding true to the basic principle of public access may need to be constructed.

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?

To optimize their scientific and commercial utility, articles reporting on the results of federally funded research should be made immediately available to the public in freely accessible digital repositories. The federal government should also consider providing support to cover reasonable publication fees for those authors who opt to publish their articles full open-access journals (those that are immediately freely accessible, and enable full reuse rights such as those supported by the Creative Commons Attribution (CC-BY) license).

Given the rapid growth in the number of open-access journals, and the increasing adoption of the open-access model by publishers across the journal marketplace, we note that the use of embargoes only benefits one subset of publishers that use a very specific, subscription-dependent revenue model. Open-access publishers, whose business models replace subscription fees with article processing fees, institutional subsidies, advertising, and other revenue streams, have very different revenue models, and receive no clear benefit from embargoes of any length. However, to accommodate those journal publishers who choose to continue to rely on subscription income, an author-determined embargo period that is as short as possible – preferably 6 months -- could be considered.

The discussion of the inclusion of embargoes in public-access policies often centers exclusively on their potential to protect publisher revenues. However, since one of the goals of an effective federal public-access policy is to balance the needs of all stakeholders, it is also important to consider the impact of embargoes on other stakeholders. Embargoes of any length come with a cost in terms of decreased public access and a negative impact on the degree to which an article's availability fosters further research and development.¹⁹

Some publishers have argued that public-access policies -- including the NIH Public Access Policy, which includes a lengthy 12-month embargo period – will discourage individuals and institutions to subscribe to journals, and cause them financial harm. The Alliance believes that any data provided by publishers documenting such a negative impact should be carefully considered; however, we know of no studies that directly examine this hypothesis nor any documented examples of journals whose financial viability has been significantly damaged by public-access policies.

In examining the length of embargo periods currently in use, a maximum embargo period of six months has emerged as the norm among biomedical research funders, with the NIH an outlier allowing 12 months. In other disciplines, embargoes of maximum 12 months are most prevalent

¹⁹ Houghton, Rasmussen and Sheehan, page 8. 2010

in research funder policies around the globe.²⁰ This is also consistent with the current voluntary practices of many publishers. Highwire Press, one of the premier online hosting services for scholarly journals, currently lists hundreds of journals in a variety of disciplines that make their articles freely accessible after a 12-month (or shorter) embargo period.²¹

Conclusion

The Alliance for Taxpayer Access, with its diverse membership of consumer groups, patient groups, universities, students, and library organizations, strongly supports the establishment of policies that ensure fast, free, public access to the results of research funded by our tax dollars. We believe that the NIH played an important leadership role in establishing a clear, successful blueprint for public accessibility to the results of its publicly funded research. We note that many other research funders around the world – both public and private – have established policies that share many of the characteristics of the NIH Public Access Policy, and encourage the U.S. federal government to implement an expanded version of the NIH policy to all other federal science agencies in an expeditious manner.

On behalf of the Alliance, we look forward to working with you to help ensure that the public's investment in research is maximized to the fullest extent. If you have any questions or comments, please don't hesitate to contact us.

Sincerely,



Heather Joseph
Spokesperson
Alliance for Taxpayer Access

(Attachment)

²⁰ <http://www.roarmap.org>

²¹ <http://highwire.stanford.edu/lists/freeart.dtl>



Members

Action to Cure Kidney Cancer - NY
AIDS Action Baltimore - MD
AIDS Vaccine Advocacy Coalition - NY
American Association of Law Libraries - DC
American Library Association - DC
American Medical Student Association - VA
Amherst College Library - MA
Anaheim Public Library System, Central Branch - CA
Appalachian State University Library - NC
Arthritis Foundation - DC
Asian & Pacific Islander American Health Forum - DC
Association of Academic Health Sciences Libraries - WA
Association of Cancer Online Resources - NY
Association of College & Research Libraries - IL
Association of Maternal and Child Health Programs - DC
Association of Research Libraries - DC
Association of Southeastern Research Libraries - GA
Autism Speaks (formerly National Alliance for Autism Research) - DC
Autosomal Recessive Polycystic Kidney Disease and Congenital Hepatic Fibrosis Alliance(ARPKD/CHF Alliance) - PA
Barth Syndrome Foundation - FL
Boston College Libraries - MA
Boston Library Consortium - MA
Bowdoin College Library - ME
Bowling Green State University Libraries - OH
Breast Cancer Network of Strength - IL
Carnegie Mellon University Libraries - PA
CFIDS Association of America - NC
Chemists Without Borders - CA
Chordoma Foundation - NC
Christopher Reeve Foundation - DC
Cold Spring Harbor Laboratory Library and Archives - NY
Colorectal Cancer Coalition - DC
Committee for Economic Development - DC
CUNY – City College Libraries - NY
CURE: Citizens United for Research in Epilepsy - IL
Cutaneous Lymphoma Foundation (formerly Mycosis Fungoides Foundation) - MI
Coalition for Heritable Disorders of Connective Tissue - DC
Colorado State University - CO
Conquer Fragile X Syndrome - FL
The Creutzfeldt-Jakob Disease (CJD) Foundation - OH
Cystinosis Research Network - IL
Denison University – William H. Doane Library - OH
Down Syndrome Treatment and Research Foundation - CA
Eastern Kentucky University Libraries - KY
Electronic Frontier Foundation - CA
Emory University Libraries - GA
Essential Action - DC
Facing Our Risk of Cancer Empowered (FORCE) - FL
Francis Countway Library of Medicine (Harvard Medical School) - MA
Fred Hutchinson Cancer Research Center - WA
FreePatentsOnline
Genetic Alliance - DC
Global Neuroscience Initiative Foundation - WA
GNU EPrints - UK
Grand Valley State University Libraries - MI
Greater Western Library Alliance - MO
International Mosaic Down Syndrome Association (IMDSA) - TX
International Journal of Medical Sciences - MD
IP Justice - CA
IsoDivalent 15 Exchange, Advocacy and Support (IDEAS) - OR
Kent State University Libraries - OH
Knowledge Ecology International (formerly Consumer Project on Technology - CPTech) - DC
Linda Hall Library of Science, Engineering & Technology - MO
Loyola University Chicago Libraries - IL
Medical Education Online - MI
Memorial Library, Berry College - GA
National Coalition for PKU & Allied Disorders - MA
National Fragile X Foundation - MI
National Tay-Sachs & Allied Diseases Association - MA
New England Biolabs - MA
Oberlin College - OH
Ohio Library and Information Network - OH
Parent Project Muscular Dystrophy - OH
Planetree - CT
Polio Survivors Association - CA
Prader-Willi Syndrome Association - FL
Pseudoxanthoma Elasticum (PXE) International - DC
Public Knowledge - DC
Public Library of Science - CA
Purdue University Libraries - IN
Right to Research Coalition - DC
Scholarly Publishing and Academic Resources Coalition (SPARC) - DC
South Dakota State University, Hilton M. Briggs Library - SD
Special Libraries Association - VA
Spina Bifida Association of America - DC
Students for Free Culture - FL
Swarthmore College - PA
Tourette Syndrome Association - DC
Trinity University Coates Library - TX
Tufts University Libraries - MA
Turner Syndrome Foundation, Inc. - NJ
Universities Allied for Essential Medicines - NJ
University of Colorado at Boulder Libraries - CO
University of Connecticut Libraries - CT
University of Kansas - KS
University of New Hampshire - NH
University of North Carolina - Chapel Hill, School of Information and Library Science - NC
University of Wisconsin – Madison Libraries - WI
University of Wisconsin Oshkosh – Forrest R. Polk Library - WI
Utah Academic Library Consortium - UT
Veterinary Information Network - CA
Washington State Board for Community & Technical Colleges - WA
Wayne State University College of Nursing - MI
Williams College Libraries - MA

January 12, 2012

Ted Wackler,
Deputy Chief of Staff.
Office of Science and Technology Policy
Submitted via email: publicaccess@ostp.gov

**RE: SIIA Comments to FR Doc No: 2011-28623, Public Access to Peer-
Reviewed Scholarly Publications Resulting From Federally Funded
Research**

Dear Mr. Wackler,

On behalf of the Software & Information Industry Association (SIIA), thank you for the opportunity to comment on the Office of Science and Technology Policy (OSTP) Request for Information (RFI) issued November 3, 2011 regarding "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research."¹

SIIA is the principal trade association of the software and digital information industries, representing more than 500 leading companies that develop and market software and electronic content. Our members include leading technology companies that provide the backbone of the Internet, as well as electronic publishers whose investments provide the public with a wide variety of information products and services covering nearly every subject matter imaginable, including publishers of peer-reviewed scientific literature, books that incorporate findings from government research, as well as databases and graphics that assist researchers in better analyzing, understanding and using research information. These industries have long produced significant knowledge-based, value-added jobs to our economy and our Nation's innovation base and are committed to continue doing so without interference with their rights as publishers.

SIIA has a long history of supporting open e-government, dating back to the turn of the century when we worked closely with Congress and the Administration in support of the E-Government Act of 2001. From the affirmation that the Government's information is a national asset, to the objective to harness new technologies to rapidly disclose information and engage citizens, SIIA strongly supports the President's commitment to openness, transparency and collaboration established in his memo to executive agencies on his very

¹ [76 FR 214, November 4, 2011.](#)

first day in office,² and the ensuing Open Government Directive that has defined this Administration.³ In response to the questions posed in the RFI, I submit the following recommendations on behalf of SIIA.

1. Federal public access policies should be limited to the direct results of publicly-funded research, not expanded to include value-added works provided by the private sector.

SIIA strongly supports government policies and initiatives aimed at ensuring broad public access to the results of publicly-funded research. However, it is essential that these policies and initiatives be limited to the *direct* results of publicly-funded research and not extend to value-added information products that result from private sector investments and publishing. The peer reviewed journals and other value added products and services that private-sector publishers, including commercial publishers, professional societies, and university presses publish are not merely the result of the publicly-funded research. Rather, these works add further value by incorporating comments, interpretations and additional expert insights to enhance their customers' understanding of the research activities.

This is a very clear, significant distinction that cannot be over-emphasized. As OSTP seeks to maximize return on Federal investments made in R&D, and to leverage those investments to stimulate scientific and technological innovation and competitiveness, we hope that you will recognize this considerable distinction between Government—public—information and the value-added works that result from the substantial investment and contribution made by the private sector.

Unfortunately, the recent RFI proposes to extend “broad public access to the peer-reviewed scholarly publications that result from federally funded scientific research.”⁴ The principle of public access should apply to the direct results of government funding, such as government reports, not to value-added products such as copyrighted, peer-reviewed publications.

2. Federal public access policies should recognize and seek to preserve the valuable role that scientific publishers play in the peer review process and their contribution to the economy.

Additionally, as OSTP conducts its review on public access to the results from federally funded scientific research, it should pay close attention to the impact of its

² [Memorandum](#) for the Heads of Executive Departments and Agencies on Transparency and Open Government (January 21, 2009).

³ [Memorandum](#) for the Heads of Executive Departments and Agencies on the Open Government Directive (December 8, 2009).

⁴ 76 FR 214, November 4, 2011

recommendation on scientific publishers. Section 103(9) of the America COMPETES Act enacted in early 2011 requires OSTP to “take into consideration the role that scientific publishers play in the peer review process in ensuring the integrity of the record of scientific research, including the investments and added value that they make.”⁵

The private sector publishing industry—including both for-profit and not-for-profit publishers—has set the high-quality standard for scientific, technical and medical (STM) information that exists today.

STM publishers and their employees contribute positively to our nation’s economy—a fact that should also be weighed against the purported public benefit of forcing journal publishers to share their works freely without compensation or further control on how their copyrighted works are distributed and used. Non-profit and commercial publishers invest hundreds of millions of dollars every year in the peer review, editing, publishing, disseminating and archiving of scholarly journal articles. There are over 1,000 STM publishers that employ some 30,000 people and indirectly support an additional 20,000 workers in the United States. These U.S.-based employers publish approximately 45 percent of all peer-reviewed research papers worldwide. For many U.S. publishers, over 50 percent of their revenue comes from foreign subscriptions—billions of dollars per year—making this a very strong U.S. export industry.

Subscriptions to STM journals continue to evolve from a basic subscription to a hardcopy journal, to electronic access to a database of current and archived articles published by not-for-profit and for-profit publishers. So while in many cases subscription fees have transitioned to fees for online access to peer reviewed works, these services are still critical to pay the cost of the peer review, editing, publishing, distribution, archiving, and quality control process. Moreover, many publishers have already instituted additional services that allow their readers and users to better analyze, evaluate and incorporate information to enhance their own knowledge and further research activities. Public access policies that require this information to be freely available around the world within a certain period of time would undermine the critical business model that promises to sustain the high-quality standard for STM scholarly published works.

A broad policy mandating free public access to final published, copyrighted journal articles arising from research funded by agencies of the U.S. Government would severely compromise the ability of STM publishers—particularly the smaller not-for-profit publishers—to retain subscribers or charge access fees to recoup their peer-review and quality control costs for producing first-rate STM scholarly works. Such a policy would also

⁵ America COMPETES Act (PL 111-358)

threaten the ability and willingness of these publishers to continue providing innovative products and services going forward.

A policy that eliminates journal publishers' ability to recoup their investment would likely force publishers to begin levying substantial author fees to recover the cost of publication, or to simply stop publishing entirely. Either of these alternatives threatens to undermine the critical STM peer review and publishing model that is so effective today, and the industry as a whole. Under the former scenario, a shift to a predominantly author-fee based system, the objectivity of journal publishers would be compromised by a significant reliance on author fees. Under the latter scenario, a decline of small publishers, authors and researches could lead to an overall deterioration in the high-quality publication process provided by a the competitive publishing industry that exists today.

3. Federal public access policies should support public-private collaboration to improve interoperability and achieve the widest possible dissemination and discoverability of publications that analyze and interpret research.

SIIA has long been a proponent of the use of open standards and open formats developed with input from a broad range of stakeholders to maximize interoperability. Additionally, we agree that improved scholarship and access can best be achieved by promoting interoperability among various research databases and publication platforms. This approach should also support OSTP's goal of enhancing the effectiveness of search and discoverability across journals and articles. To this end, SIIA endorses the continued efforts of OSTP and relevant agencies to work cooperatively with the research community and private sector publishers in the promotion of open-standard formats that can facilitate greater interoperability, broad access and long term preservation of both data and peer reviewed scholarly publications. However, federally mandated use of particular platforms or formats does not foster interoperability; rather it would stifle the important consensus process under which technical criteria, methods, processes and practices are developed to suit the needs of the broadest number of uses and users to maximize the availability of information, including products and services provided by private sector publishers.

Conclusion

In summary, SIIA fully agrees that taxpayers should have access to the output of taxpayer funded research, and that the Government should ensure access to its direct outputs. However, the output of the federal funding is the research, its conclusions and data resulting from this research, not the peer-reviewed scholarly publications that are produced by publishers as the result of significant private sector investment.

Further, SIIA strongly supports the continued efforts of OSTP to work collaboratively with the research community and private sector publishers in the promotion of open-standard formats developed by a consensus of all stakeholders that can greater facilitate greater interoperability, broad access and long term preservation of both data and peer reviewed scholarly publications—however, interoperability also should not be sought via federal mandates of particular platforms or formats.

Again, thank you for the opportunity to participate in the public consultation on Public Access. We look forward to continuing to work with you and the agencies throughout the process of developing and implementing public access policies that are effective and appropriate. If you have additional questions based on these comments or would like to discuss further, please do not hesitate to contact David LeDuc, SIIA Senior Director for Public Policy, at dleduc@sia.net or 202-789-4443.

Sincerely yours,

A handwritten signature in black ink that reads "Ken Wasch". The signature is written in a cursive, slightly slanted style.

Ken Wasch
President



IADR/AADR Response to the Office of Science and Technology Policy Request for Information on Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research FR Doc. 2011-28623

To:
Office of Science and Technology Policy
Attn: Open Government Recommendations
725 17th Street
Washington, DC 20502
via e-mail to: publicaccess@ostp.gov

From:
Dr. Christopher H. Fox, DMD, DMSc.
Executive Director
International and American Associations for Dental Research
1619 Duke Street
Alexandria, VA 22314
cfox@iadr.org

Dr. William Giannobile, DDS, MS, DMSc.
Editor-in-Chief
Journal of Dental Research
1619 Duke Street
Alexandria, VA 22314

January 2012

January 12, 2012

Dr. John Holdren
Director
Office of Science and Technology Policy
Executive Office of the President
725 17th Street, NW
Room 5228
Washington, DC 20502

Dear Dr. Holdren:

We write on behalf of the International Association for Dental Research (IADR) and its American Division, the American Association for Dental Research (AADR). The IADR, with over 11,000 members worldwide, including 3,700 members in the AADR, is dedicated to advancing research to improve oral health and to facilitating the communication and application of research findings. The IADR and AADR are owners of the *Journal of Dental Research (JDR)*, a specialized scientific journal that uniquely serves the oral health and dental research community. The *JDR* has one of the top Scientific Impact Factors of any peer-reviewed dental journal. The *Journal* has the top Eigenfactor Score, which measures the number and quality of citations. The *Journal* also has the top Article Influence Score, which is a measure of the influence of articles over the first five years after publication. Given the importance of the *JDR* to oral health, we are pleased to provide a response to the November 3, 2011 Office of Science and Technology Policy (OSTP) **“Request for Input (RFI) on Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research.”**

The IADR and AADR share the belief – of the broader scientific community – that the results of federally funded research should be widely disseminated. We also reiterate our longstanding support for the process whereby publishers are transferred control of copyright and distribution rights in exchange for funding the post-grant peer review and publication process. This relationship, which results in the best science being disseminated to the scientific community as efficiently as possible, continues to be threatened by policy proposals that fail to recognize the very real costs associated with the production of scholarly publications. Each year, scientific publishers invest hundreds of millions of dollars in staff, technology, capital projects, an editorial selection process, and operational funding of independent peer review on all research articles by experts in specialized fields prior to publication. This dynamic can't continue if public access mandates are expanded among federal research agencies along the lines of the National Institutes of Health (NIH) public access model or if existing embargo periods are shortened.

NIH specifically requires submission of the final manuscript only after the manuscript has passed through the publisher's quality assurance peer review processes and determination of acceptability for publication, even though the journal publisher is not a party to the funding agreement for the research. The NIH public access mandate should not be viewed as a success for science or as a model to be replicated, as the long-term viability of scientific journals has been unnecessarily threatened. As an example, for nearly 90 years, the *JDR* had been edited, proofed, peer reviewed, typeset, designed and distributed by employees at the IADR headquarters. However, due to a confluence of factors, not least of which being a government mandated public access policy and the uncertainty of that public access policy expanding or embargo periods shortening, our Board of Directors decided that working with a private sector publisher was the only option to sustain the publication. If this trend continues, the ultimate result will be the consolidation of scholarly journals in the hands of just a few publishers and

publishing decisions based partly on the source of research support as opposed to solely on the quality of research.

The main source of revenue to cover the expenses of our peer review infrastructure, print publication and online version comes from individual and institutional subscriptions. In a typical year, the *Journal of Dental Research* will have about 30% of its accepted research manuscripts with some NIH funding, although it has been as high as 57%. It follows that if the NIH mandate is expanded to additional federal agencies and/or if the existing NIH embargo period is shortened below 12 months, the impact on the *JDR* and other scientific journals would be catastrophic. For a small professional association, we invested significant resources to launch our *Journal* online in 2002, and digitize all of our volumes back to 1919. The only way for the Associations to recoup this investment – not make a profit – is to retain the copyrighted material and to offer individual and institutional subscriptions. A drop in subscriptions in recent years, subsequent to the inception of the NIH public access mandate, was a major contributing factor to our Associations having to cease in-house copyediting and production of the *JDR*.

We recommend the use of post-grant reporting infrastructure as a means to provide the public access to more easily digested information. The scientific community, for whom most scholarly articles are written, has rarely cited a lack of access to federally funded research findings as a problem. The post-grant reporting mechanism continues to be underutilized, as federally supported scientists could easily produce summary results in laymen's terms for public consumption. The lack of a properly utilized post-grant reporting infrastructure should not lead to the taking of publisher's long held copyright as a surrogate for end-of-grant reports. If publishers' copyrights in journal articles continue to be undermined, simple economics will render public access policies obsolete, as a number of journals will no longer be able to fund the cost of post-grant peer review. This would have a devastating impact on scientific integrity, and would leave U.S. scientists at a competitive disadvantage to their peers in other countries.

We look forward to working with the Administration and the entire scientific community to build a better oral health research reporting system for the public. We believe the private sector has made significant strides in IT infrastructure and making central repositories fully compatible and user-friendly. Government collaboration with industry to leverage existing resources would meet the requirements of the America COMPETES Reauthorization Act of 2010 and President Obama's goals of creating a more open and transparent government, while acknowledging existing and well established copyright protections.

Sincerely,



Christopher H. Fox, DMD, DMSc.
Executive Director



William Giannobile, DDS, MS, DMSc.
Editor-in-Chief

IADR/AADR Response to Questions Posed in FR Doc. 2011-28623

(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?

- Agencies should identify specific needs of particular user groups that are not already being met and collaborate with publishers and other stakeholders to meet those needs most effectively. As owners of the *Journal for Dental Research*, we have not, to date, had a request for an article or volume from a patient or other interested party who simply could not afford it. However, if that were to happen, we would be happy to provide the requested article free to that patient. As a result, we fully consider that access is already “open” to our Journal and question the need for additional government intervention. The U.S. economy and scientific enterprise would be best served by government exercising restraint when pushed to issue new mandates that would lead to the collapse of scientific journals. Instead, the federal government could work with researchers to make final grant reports a more useful and accessible tool for the general public.
- Open access government mandates have significant costs to the U.S. economy and the scientific enterprise.
 - NIH’s PubMed Central data indicates 2/3 of users are from overseas, undermining critical export opportunities for an \$8 billion publishing industry that employs 50,000 Americans.

(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?

- The federal government should avoid issuing mandates that take intellectual property without providing funding to support the process that leads to the product.
- The general public derives limited direct information from technical scientific journal articles, and would gain a better understanding of the science being conducted at federal research agencies by the production of more user-friendly end-of-grant reports. These reports are already required, but are not being looked at as a satisfactory means of disseminating scientific knowledge across public populations. We fully support working with the government to make these reports user friendly, freely accessible and interoperable with our articles.

(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 stewardship of scholarly articles carries a cost that is already being paid by publishers. The federal government would be better served by utilizing such funding to support research grants. The *Journal of Dental Research*, as an example, has already made a significant investment in infrastructure to create a user-friendly and innovative online platform. Additionally, Internet search engines, abstracting services, and other tools do an excellent job of ensuring the discoverability of research, and the technology continues to improve. Given current federal budget constraints, it makes little sense for the federal government to duplicate these efforts.

(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?

- If the government improves final grant reports, publishers could link to them.

(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?

- To our knowledge, searching for scholarly publications has not been a barrier to open access. Search engines like Google and Bing are performing well. Our direct experiences with clinicians and clinical researchers working at the patient level are unaware of problems patients cite with regard to access of scientific information. If it exists it likely appears to be a very small minority of individuals. These individuals often times reach the investigators directly who provide a complimentary copy of an article should they seek the in-depth technical information found in a research publication.

(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?

- The government and the private sector should work together to better disseminate the results federally funded research to the public. Taxpayers could be provided digestible final reports of the research findings, which could also drive public traffic to research agencies in order to increase public interest and support for the science being conducted. We believe that the final progress reports that are required by federal agencies could be made more robust, being written for a public audience and housed on an interoperable and user-friendly IT infrastructure. This model would be similar to the one established under the America COMPETES Act, which required researchers funded through the National Science Foundation to provide a final report that described their research findings, which is then deposited in a central and public repository.

This model can be adopted in a consistent manner with the President's Open Government initiatives, and will respect the long standing copyright protections that have financed the post-grant peer review process. Perhaps most of all, it is a model that makes a clear distinction that the articles contained in peer reviewed scholarly journals are not drafted for a public audience. They are written by researchers seeking validation amongst their colleagues. This model validates and filters the best science into one repository for scientists, helping scientists to more efficiently review breakthroughs and innovations in their own field.

Simply "taking" publishers' accepted manuscripts as a surrogate for the lack of robust public research reports will likely lead to a reduction in the number of scholarly journals, and leave editors with the undesirable economic incentive to maintain a sustainable low level of federally funded open access articles in their journals. Such a policy creates an environment that ultimately harms the U.S. researcher's ability to compete on the global stage, as our researchers are published and cited less than counterparts in other countries.

(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?

- We do not believe that additional types of peer-reviewed publications resulting from federally funded research should be covered under public access policies. New regulations of this type would further impede the ability scientific associations and publishers to generate revenue sufficient to cover the costs of production.

(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?

- We do not believe that one can identify an "appropriate" embargo period, as the useful life of research varies significantly among the various disciplines. As an example, the Association of American Publishers has expressed that across their 37 journals there is a long half-life and lifetime usage of about 4.5 and 19.5 years, respectively. In mathematics, journal articles published in 2009 were as likely to cite articles published before 1998 as after them, and only 10% of the citations were from the previous three years –according to a February 2011 report of the Mathematical Sciences Research Institute. Any embargo period is a dramatic shortening of the period of copyright protection afforded all publishers, and likely to significantly impact publishers' ability to add value and innovate.
- With respect to the NIH public access mandate, we ask that OSTP reject efforts to shorten the embargo period below 12 months. Implementation of the existing policy came at a significant cost to publishers, and a move to a 6 month embargo period – as suggested by some – would likely bring an end to many biomedical research publications.

Additional Comments for Consideration

Importance and Uncompensated Costs of Peer Review Process

NIH has acknowledged the value that is created through the post-grant peer review process by encouraging researchers to seek publication in a scholarly journal.¹ NIH could have chosen to manage this process on their own at any point in time by providing the additional costs and infrastructure for post-grant peer review. However, the publishing process has been a well-functioning and long-standing partnership between research agencies and publishers; agencies fund the application peer review that decides which grants are funded, as well as the research itself. Then, the scientific community relies on publishers to manage the post-grant peer review process to evaluate the merit and authenticity of the conclusions of the research. However, unlike the federal funding provided during the pre-grant peer review process, post-grant peer review is not funded by the agencies at all. There is no federal funding that goes into the publication process. As such, we oppose new government mandates requiring that scholarly publications be made available online without compensation for the work that goes into the product.

Although the days of mailing unedited manuscripts around the world for review are gone, there still exists information technology (IT) infrastructure that is necessary to send manuscripts to reviewers in numerous countries, while being able to capture and evaluate all of their comments. This is an exceptionally intensive and collaborative task, one that incurs real costs both in terms of IT, but also in human capital and labor.

Threats to U.S. Scientific Enterprise

If a journal wishes to maintain their in-house journal operations, the inevitable result of a public access policy will be for editors to simply accept fewer federally funded articles. Journals that publish a majority of federally funded articles will likely see a steeper acceleration in the number of members and institutions dropping subscriptions, as compared to those that are predominately made up of articles not subject to strict public access policies – such as those from the international community or those that are privately funded (as noted above, the *JDR* normally has only about 30% wholly or partially-funded NIH articles). These federally funded articles will represent a liability to any journal, and a publisher or editor will have to manage the number of these articles to ensure sustainability of the subscriber base.

In essence, privately funded articles, which are not subject to an open or public access policy, will have to subsidize the decreased readership from federally funded articles. A ratio of privately funded research versus less federally funded research will have to be maintained so that a journal can maintain readership. In short, a public access policy any more stringent than the current design greatly incentivizes publishers and editors to accept far more non-federally funded articles over those subjected to a public access policy of 12 months or less in order to maintain subscriptions.

With an expanded open access policy, there will be a number of small non-profit scholarly journals that have too high a ratio of articles about federally funded research, resulting in decreased subscriptions that will create an operating loss for the journal. As more and more of these journals outsource their negative-return operations, there will be less of an appetite from large publishers to take on these journals and publications, as the non-open access heavy journals are left to subsidize the heavily open

¹ In the NIH Grants Policy Statement, the NIH “encourages grantees to arrange for publication of NIH-supported original research in primary scientific journals.” However, in the Grants Policy Statement, the NIH also informs the grantee that the NIH has irrevocable authority to take the article from the publisher and reproduce the results as it sees fit. This policy is not only delineated without regard to copyright law, but it also encourages grantees to seek peer review of their work so that publishers can assume all of the costs of peer review and publication, while the NIH can wait for the finished product and then claim ownership of it.

access journals. Eventually, there will be no more outsourcing partners for these journals looking to outsource operations, and they will eventually have to cease operation.

This is an inevitable side effect that will result from a public access model that is less than 12 months. Editors, with an intimate knowledge of a journal's financial viability and status, will tacitly favor non-public access articles in order to maintain an economically viable journal. Researchers will no doubt discover the new economic dynamic surrounding peer review of manuscripts, and those that have been denied publication of strong scientific articles will indignantly wonder if their submission was denied because the journal had already met its "quota" of public access articles. This is an unwanted economic dynamic to introduce to an otherwise fully scientific and meritorious peer review process.

Already U.S. scientists are behind other countries in terms of science funding as a percentage of GDP, both from private and public sources. A strict public access or open access policy, in many instances, would force many publishers to further fuel this competitive disadvantage by lowering the citation rate and publication of U.S. scientists, key factors used to raise the profile of scientist and country at the global level.



11200 Rockville Pike, Suite 302
Rockville, Maryland 20852
USA

American Society for Biochemistry and Molecular Biology, Office of Public Affairs

BENJAMIN W. CORB
Director

JULIE M. MCCLURE, PH.D.
Science Fellow

The American Society for Biochemistry and Molecular Biology appreciates the Office of Science and Technology Policy providing this opportunity to comment on public access to scholarly publications that result from federally funded scientific research. ASBMB sees no justifiable reason for a change in currently existing NIH policies as they relate to open access. ASBMB is compliant with open access policies set forth by the National Institutes of Health and deposits required data on the author's behalf to ensure compliance with these policies.

Response to
Request for Information: Public Access to Peer-Reviewed Scholarly Publications
Resulting From Federally Funded Research
Submitted by the Coalition of Open Access Policy Institutions (COAPI)
January 2012

Background:

Access to scholarly research literature is a crucial concern for universities, colleges, and research institutes worldwide. That concern, in addition to other considerations, has led faculty at many institutions to adopt open access policies designed to disseminate the results of their research as widely as possible.

During a July 19, 2011 teleconference, representatives from 22 North American institutions with existing faculty-initiated open access policies agreed to form a coalition in order to collaborate and share implementation strategies for their policies and advocate at national levels on issues related to their policies. This new alliance, the Coalition of Open Access Policy Institutions (COAPI), was announced on August 3, 2011 in a press release issued by the University of Kansas: <http://www.news.ku.edu/2011/august/3/openaccess.shtml> COAPI has since grown to 41 institutions that have open access policies or are working toward such faculty-led initiatives. COAPI members include leading public and private universities and colleges as well as independent research institutes. We represent an important segment of higher education and research communities in North America.

COAPI has a unique perspective because faculty at our institutions have recognized the importance of greater access to scholarship and embraced it as a core value. They view access to research literature as a critical component of both individual researcher and institutional effectiveness. COAPI faculty and researchers have firsthand experience with the problems created by limited access to research and scholarship and they have demonstrated in a concrete way their belief that broader access will benefit both scholarship and society.

Representatives of COAPI member institutions met in Washington, DC on November 8, 2011 prior to the Berlin 9 Open Access Conference. During the meeting COAPI members agreed that one of our first actions would be to respond to the Office of Science and Technology Policy's Request for Information to provide "recommendations on approaches for ensuring long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally funded scientific research." The following response to the RFI, which has been

approved by COAPI members, was developed by a working group and discussed on two separate conference calls of the full COAPI membership.

Summary recommendation:

The current NIH Public Access Policy, implemented in 2008, applies to the results of approximately one-third of all federally funded scientific research. The NIH policy, while it is not without limitations, has been enormously successful in opening the results of NIH research to a broader audience – to the benefit of science and the general public. There is an urgent need for the federal government to adopt a comprehensive public access policy approach applicable to all major research funding agencies, one that would both extend and improve upon the current NIH policy. COAPI recommends a policy framework that 1) is as uniform as possible for all agencies, 2) is mandatory for all researchers funded in whole or in part by those agencies, 3) results in rapid and open access to the results of peer-reviewed, government-funded research, and 4) allows flexible rights of reuse.

The members of COAPI encourage policymakers to consider carefully the ways in which research information can be both accessed and reused for optimal scientific, economic, and social benefit. Faster public access, with minimal delays following publication, coupled with full reuse rights will result in more rapid advancement of scientific discovery, as well as faster product development and commercialization in all research areas. Such an approach will spur economic growth in broad sectors of the economy, including those of strategic importance such as biotechnology, renewable energy, and sustainable agriculture. It will encourage private investment in enterprises that capitalize on information generated from government-funded research. It will also have optimal benefits for the general public.

Comment 1

[1.a. 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?]

Successful development of markets related to access and analysis of government-funded peer-reviewed publications depends in large part on the speed with which research information is made available and the terms under which it can be used. The combination of rapid public access and liberal reuse rights will drive software development that facilitates new types of information discovery and tools for research. It will create the capacity for new information-based business models that draw on the innovations in information technology, such as the semantic web, which fosters sharing and reuse of information across applications and community boundaries. Full open access in this sense will also foster commercialization of products that increase access to and awareness of specialized research information.

All of these potential capacities will be reduced to the extent that access is delayed through embargoes or that reuse rights are limited unnecessarily.

Text mining, data mining, other forms of information computation, and the creation of derivative works are examples of new research and information dissemination capacities that can be enabled through appropriate reuse rights. An example of one such tool that could be exceptionally powerful in a full open access environment is Action Science Explorer, which is designed to speed understanding of scientific literature. See: <http://www.cs.umd.edu/hcil/ase/> In addition to potential commercial applications, such tools could also be valuable to funding agencies by allowing them to monitor research developments in specific fields as part of the process of setting funding priorities.

A broader federal public access policy framework of the kind we envision will also foster the continued development of open access journals (which now number more than 7,000 titles) and the transition of traditional publishing to open access business models – again to the benefit of science, economic development, and public welfare. Commercial firms – both new firms such as Hindawi and existing ones such as Springer – are clearly realizing the economic benefits of open access through the creation of profitable new journals that follow open access business models. Nonprofit publishers are also experimenting with open access publishing and thereby extending the reach of the research they disseminate. The growth of publicly accessible research information will encourage scholarly publishers (both nonprofit and for-profit) to transition to open access in ways that meet both their scholarly missions and their economic interests. A broader federal public access policy framework will thus both add to and encourage the continued growth of openly accessible research information.

[1.b. How can policies for archiving publications and making them publicly accessible be used to grow the economy and improve the productivity of the scientific enterprise?]

Houghton's work clearly demonstrates the economic value of agency policies that ensure public access to the full results of their funded research. His 2010 study estimates that opening access to all U.S. federally funded scientific articles would result in at least a five-fold increase in return on investment. Specifically, the net present value gains of expanding an NIH-style policy to all other U.S. science agencies is estimated to be on the order of \$1.5 billion. Of that figure, approximately 60% is estimated to accrue directly to the U.S. economy.¹

¹ Houghton, J., Rasmussen, B., & Sheehan, P. (2010). *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*. Report to SPARC. Centre for Strategic Economic Studies. Victoria University. Victoria, BC. See: <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>

Minimal restrictions on the commercial use of federally funded research information will encourage economic growth. Current practices limit reuse rights to either what is allowed by fair use under copyright or what is permitted by licenses that are negotiated between journal publishers and libraries. Most restrictions on use needlessly hamper the commercial development of new products and services and their introduction into the marketplace; they stymie rather than encourage economic development. Appropriate commercial use can be achieved through current copyright law and the licensing framework for agency policies, as discussed below under Comment 2.

[1.c. What are the relative costs and benefits of such policies?]

Numerous studies have demonstrated that openly accessible research information reaches wider audiences and produces more citations than research published under access restrictions. Recent studies are also showing that openly accessible research produces more diversity in follow-on research. It encourages contributions by participants who would have had no opportunity to contribute in an environment with access controls. It thus increases the potential for innovation and the interdisciplinary application of research through a larger pool of participants.

As noted, Houghton's studies have demonstrated the clear economic benefits of opening access to government-funded research. Given his findings, the opportunity costs of not making government-funded research openly accessible are equally clear.

We know from the NIH example that making such research openly accessible is extremely cost-effective, especially when considered in the context of overall benefits. The NIH reports that it costs \$3.5 - \$4.6 million annually (or about one hundredth of one percent of the NIH budget) to provide access to results of its funded research. [PubMed Central](#) is currently used by more than 500,000 users per day, with the majority of users coming from outside academe, underscoring strong demand for this information by the public.

A government-wide public access policy or policies can be implemented by leveraging existing infrastructure in ways that minimize duplication of effort. The investments in software and other resources that already support NIH's PubMed Central and similar repositories can be utilized by other agencies either individually or in a federated model.

A comprehensive federal public access policy framework will have the added benefit of increasing the effectiveness of government research funding. One of the primary motivations of the NIH policy was improved documentation of the outcomes of sponsored research. A comprehensive federal policy will bring that benefit to all of the major scientific research funding agencies. It will also provide

congressional appropriators and authorizers better information to assess the value of existing expenditures and better target strategic funding priorities. It will thus increase agency accountability and support informed, transparent, and evidence-based budget and policy decision-making in accordance with the Obama administration's emphasis on open government.

[1.d. What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?]

With the exception of research covered by the NIH policy, the present system of disseminating the results of government-funded research is clearly inadequate. The system does not adequately serve the interests of government, science, or economic growth. It relies heavily on researchers donating copyright to "toll access" journals that limit access by means of licenses and subscriptions. Dissemination of research information is primarily through academic and research libraries. Given constrained budgets and the current cost of scientific journals coupled with the rapid rate of cost increase over time (which has been significantly above the rate of general inflation in the economy), most libraries simply cannot afford to subscribe to most journal titles. As a result, researchers at most academic institutions lack the kind of access to research information that would enable them to build easily upon the results of previous research. Such limited access greatly reduces the efficiency of our nation's scientific productivity.

Access to research literature is also not optimal in the corporate sphere. Only wealthier corporations can provide even reasonably adequate access to the knowledge that their researchers (who drive product innovation) need or could benefit from. Access to current research literature at smaller companies and incubators is especially limited. Ready access to current research literature is essential for commercial product development, which is a primary driver of innovation that produces economic growth.

Inadequate access to research information also has negative effects upon broader public interests. While that is obvious in the case of health and medical information, the principle applies in many other subject areas. For example, it is important for the public to have access to the latest research information on such topics as environmental toxins and residential energy efficiency. Similarly, current research information is essential in a wide variety of public policy arenas at all levels of government, from federal to state to local. Policy decisions made without awareness of the latest scientific knowledge can result in policies that are less than optimal. Suboptimal policies in turn can have negative economic consequences. Improved access to research information would promote more informed policy debates and decisions at all levels. When scientific development, economic growth, and public welfare are considered together, the combined opportunity costs of poor access to research information are enormous.

The limitations of the present system can be overcome by providing open access to the results of research funded by the federal government. Open access in this sense means that the results of publicly funded research information should be made fully and freely accessible as rapidly as possible with few restrictions on subsequent use. Most restrictions on use will serve only to limit the return on the taxpayers' investment in research. Full reuse rights will enable researchers to build on the results of others in ways that fosters entirely new research capabilities. As noted, they will also speed the process of applying research findings to commercial products.

Comment 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?]

Faculty and staff at COAPI member institutions have considerable experience in designing licensing frameworks that facilitate their open access policies. In general, these policies allow faculty authors to retain all of their original rights under copyright while granting non-exclusive licenses to their institutions and also making copyright arrangements with publishers. The institutional licenses vary to some extent in terms of their scope, but they all have the common purpose of providing a legal framework that allows the works of faculty authors to be made openly accessible by their institutions, while granting publishing entities the limited rights they need to disseminate the published copy. Faculty at COAPI institutions are aware that they benefit most by making their works widely available for subsequent use. Their primary interests are in reaching wide audiences, being credited for their work, and being cited in ways that demonstrate the impact of their scholarship.

If the goals of agency policies are to foster the development of science, encourage economic growth, and serve the public's interests in the broadest sense, then it will be important to construct the licensing framework for the policies according to principles that will facilitate those goals. Doing that requires no change in copyright law. It is only necessary to structure the licenses that authors grant to the agencies (as a condition of their funding) and the licenses that the agencies grant to the public in ways that facilitate both access to and maximum reuse of research information. A [Creative Commons attribution license](#) is an example of a license that would fulfill those purposes. Such a license would allow authors to receive full credit for their works while also creating great flexibility in terms of how their works can be used by others. Licenses that allow only for access to research information – but not subsequent reuse or redistribution to colleagues – are unnecessarily restrictive. Unlike the NIH policy, systematic downloading of articles

should be allowed in order to facilitate flexibility in terms of reuse, for example, by programs that compute on the textual corpus.

Since the licensing framework for the agency policies would be non-exclusive, authors would remain in a position to transfer appropriate rights to publishers. Like the NIH policy, agency policies should be mandatory, with authors required to deposit their final (post-peer-review) manuscripts. In view of that, publisher transfer of rights agreements for federally funded research articles could not be structured in ways that conflict with the licenses that researchers grant to the agencies. Publisher economic interests can be protected by brief embargo periods, as discussed below under Comment #8. During the embargo periods, use of the research information would be governed either by fair use under copyright for journals in print form or – in the case of electronic journals – by the provisions of license agreements. Metadata standards, as discussed below under Comment #5, would include a full citation to the publisher copy of record. Such a policy framework would balance the needs and interests of research authors, agencies, publishers, and the general public.

Comment 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?]

Members of COAPI believe that a centralized or federated approach managed by the federal government is the most appropriate and effective strategy for ensuring interoperability as well as effective search mechanisms and analytic tools. Federally managed approaches are also the most feasible way to facilitate new research capabilities related to reuse (such as text and data mining, creation of derivative works, information discovery tools, and commercialization of products that increase access to and awareness of specialized research information). Even with carefully crafted regulatory requirements, it is clearly more difficult to establish and maintain such capabilities under a decentralized framework that includes partners outside the federal government.

The federal government has a long-term interest in making the results of its funded research permanently available. It is the only entity that has the capacity to make the full corpus of federally funded works publicly accessible, to establish and enforce standards of interoperability that ensure search access across repositories, and to establish and maintain an infrastructure that will allow new services and products to be built from publicly funded information. The federal government's capacity in this regard is demonstrated by its success in implementing the NIH Public Access Policy. As noted above, federal stewardship, as shown by the NIH

example, is cost-effective and its infrastructure can be leveraged by other agencies. A federal approach can also ensure transparency, openness, and accountability.

Primary reliance on a federal government role does not preclude private or third parties from participating in a decentralized approach. We would emphasize, however, that any decentralized approach that involves entities outside the federal government, whether public or private, would need to provide all of the capacities described above – public access, interoperability, search functionality across repositories, adherence to standards, long-term archiving and preservation, openness and accountability, and the potential for creative reuse for research and commercial purposes. If the federal government found that a decentralized approach was feasible and decided to rely on it heavily, then government agencies should maintain mirrored and accessible versions of the decentralized repositories in order to protect the public’s investment and ensure accountability. The federal government’s stewardship over this valuable public good is critical.

Comment 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 noted above, a decentralized approach that involves entities outside the federal government faces significant challenges that would not be present in an intragovernmental approach, especially if one goal of the decentralized approach is to allow and encourage a wide variety of reuse activities (such as text and data mining) that foster innovation in science and that lead to economic development. As noted, such approaches require clear standards for access, interoperability, metadata, search functionality, usage rights, and long-term preservation. The DRIVER project, funded by the European Commission, is one of the best examples of a decentralized, federated repository structure involving cooperation from universities and research institutes in several European countries. See: <http://www.driver-repository.eu/>

Academic research libraries, including members of COAPI, have developed extensive experience and expertise in creating and managing digital archives designed for long-term preservation and access. Examples include [arXiv](#) (now managed by the Cornell University Libraries), the digital repositories of several research universities (such as COAPI members Harvard and the University of Kansas), and the [HathiTrust](#), a major partnership of research libraries and research institutions that is designed to preserve digital books and broader cultural heritage. Given their expertise and focus on long-term preservation and access, research libraries could be important consultants in the development and implementation of federal, interagency and public/private partnerships in a public access policy. Some research universities could also partner with federal agencies to develop repositories for specific subject areas. We note that some academic and research

institutions have partnered with research funders to provide their permanent archives.

Publishers could be encouraged to participate in public-private partnerships by voluntarily providing the final published versions of articles after limited embargo periods that ensure their subscriptions and licensing revenues. However, given their focus on immediate income and the fact that they tend not to have long-term time horizons, commercial publishing firms in particular should not be relied upon solely for digital archiving. It should be obvious that long-term archiving and public access will be made much more difficult when corporate acquisitions, mergers, or business failures occur. For that reason, publishers should provide archiving and public access for the results of federally funded research only if the publishers' sites are mirrored by sites maintained by the federal government or by institutions that provide greater certainty of long-term preservation and access. Publishers would also have to be able to comply with detailed rules for user interface, access formats, and interoperability.

Comment 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?]

The development of “interoperable search, discovery, and analysis capacity across disciplines and archives” depends on the creation of carefully crafted metadata standards that are implemented for all archives containing the results of federally funded research. It is critical that metadata be both machine-readable and machine-interoperable if agency policies are to realize their full potential. Metadata standards for archives should be designed to facilitate the functions of use, reuse, and analysis described above.

Federal agencies, through their public access policies, are best positioned to ensure the creation of metadata standards that will meet the functional goals of their policies. The research library community, including the Library of Congress and organizations such as OCLC, has developed a variety of metadata standards that have been endorsed by standards organizations (NISO, ISO, etc.). These can be drawn upon in developing a broad federal metadata specification.

The specification should support multiple metadata standards in order to develop metadata that is as rich as possible. Some of the primary goals of the specification (along with examples of related standards) would be to: 1) provide institutional information for published sources (grant IDs, funding organization, I2 – Institutional Identifier, etc.), 2) provide descriptive information for both the

repository and published versions ([Dublin Core](#), [ORCID](#)), 3) support searching through keywords as well as controlled vocabulary schema appropriate to disciplines, 4) incorporate abstracts, 5) facilitate full text searching and web crawling, 6) support metadata harvesting ([OAI-PMH](#)), 7) establish relationships through semantic web standards ([RDF](#)), 8) support usage tracking ([COUNTER](#)), 9) support description of related data ([DataCite Metadata Schema](#)), 10) support data exchange standards ([JSON](#)), and 11) document IP rights.

It's especially important for metadata to support the capacity for machines to access and analyze both the publications themselves and the underlying data that support them – in those instances where that data can be made openly accessible.

Comment 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?]

The benefits of public access policies to taxpayers will be realized to the extent that publicly funded research results are made openly accessible. The history of the development of the NIH Public Access Policy demonstrates conclusively that a broader federal public access policy (or policies) must be mandatory. The rate of compliance with the NIH policy increased dramatically following the end of the voluntary policy and the adoption of the current mandatory policy. Average manuscript submissions have grown from approximately 1,000 per month prior to April 2008 (the date of adoption) to current levels that are well over 5,000 per month (for the most recent twelve-month period). See:

<http://www.nihms.nih.gov/stats/>

A broader federal policy must be consistent across all agencies in its requirements and mandates. Uniform requirements and procedures across all agencies will reduce burdens on researchers (who often hold grants from multiple agencies) and on the institutions that support their compliance. Uniformity will reduce complexity and that in turn will reduce the time needed to educate researchers about policy requirements, to deposit articles, and to deal with deposit and compliance problems. Uniformity will also work to increase compliance rates. Publisher interests, for example those related to embargo periods and any deposit of final published versions of articles, are also best served by a uniform approach.

Procedures should include standard criteria for what should be deposited as well as clear instructions for the deposit process. Existing grant management systems should also be integrated into the deposit process to facilitate agency and public accountability.

Many researchers work with various deposit mandates. For example, most COAPI institutions expect faculty to deposit works in their institutional repositories and many faculty receive funding from multiple extramural sources that have

deposit requirements. Agency policies should leverage existing protocols to facilitate deposit of manuscripts to multiple repositories in a consistent, standardized manner.

Comment 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 vast majority of scientific knowledge resulting from federal funding appears in the form of peer-reviewed journal articles, the primary mechanism for scientific communication. As noted above under Comment 1, dissemination of the results of federally funded research is severely hampered by limitations on access to journal literature – to the detriment of science, economic growth, and the general public interest. For those reasons, agency policies should focus on peer-reviewed journal articles. As a second priority, policies should address related supporting materials that document the research process (data, protocols, survey instruments, etc.) and facilitate replication of results. Specific requirements for supporting materials will vary across disciplines.

Comment 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?]

As noted above under Comment #1, not providing public access to federally funded research incurs significant opportunity costs. The scientific, economic, and public benefits of providing access – the return on our nation’s investment in research – diminish to the extent that access is delayed or denied. Immediate access at the time of publication is therefore ideal in terms of overall policy goals. In any case, embargoes should be as short as possible.

To protect publishers from possible financial harm due to loss of subscriptions and licenses, a maximum embargo period of up to six months could be allowed, if publishers (or others who advocate for embargoes) can provide empirical evidence demonstrating the need. Members of COAPI are not aware of any data demonstrating that the NIH Public Access Policy, with a one year embargo, has led to subscription or license cancellations or otherwise been harmful to publishers. The libraries of COAPI member institutions have not considered cancelling subscriptions due to public access and public access has also not been a factor in instances where journal cancellations were necessary due to budget reductions. In addition, COAPI members are not aware of any evidence that academic and research libraries either have considered – or would in the future

consider – public access to federally funded research to be an adequate substitute for journal subscriptions or licenses.

It is important to note that some publishers who have expressed concern in the past that public access would result in loss of subscription revenue have changed both their views and their practices. In addition, many journals, such as those of Highwire Press, open up retrospective access to their content following embargoes of 12 months or less. Embargo periods of six months or less are also the norm for biomedical research funders worldwide.

If it is demonstrated through empirical evidence that embargoes are necessary, members of COAPI believe that a uniform embargo period of six months or less should apply across all funding agencies. Such an approach has the benefits related to consistency discussed above under Comment #6; it would speed research access while also taking into account publisher interests.

If a decision is made to adopt different embargo periods for individual disciplines or sub-disciplines, shorter embargo periods (less than six months, for example) should apply to rapidly changing fields and those where research results often lead directly to commercialization.

We would emphasize that the burden of proof for the need for embargoes should rest on those who believe they are necessary. The benefits of public access are clear. In the absence of empirical evidence clearly demonstrating the need for embargoes, immediate public access should be the norm, since it is the best way to foster innovation, competition, economic growth and scientific progress.

Final Comment [Please identify any other items the Task Force might consider for federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.]

Members of COAPI believe that public access involves a public good. Federally funded research information (in the form of final peer-reviewed author manuscripts) is made possible through taxpayer dollars and should therefore be made accessible to the public in ways that maximize the taxpayer's investment in research.

At the same time, we recognize that private parties contribute to the creation of federally funded final author manuscripts. While peer review is provided gratis by fellow researchers, publishers do assist in coordinating the peer review process. In view of that contribution, publisher interests do need to be taken into account in the development of public access policies. But publisher interests should not be allowed to outweigh the interests of the public in accessing such information, the interests of federally funded researchers in seeing the widest possible dissemination of their work, or our national interest in scientific and economic development that will clearly be furthered through an optimal policy approach.

Publisher interests should be protected in rough proportion to their contribution to the full process of research production and dissemination in the form of the final author manuscript. Given all that is involved in the process of creating research and producing final manuscripts, the publisher contributions are relatively small. For that reason we wish to reaffirm our conviction that publishers or others who advocate for embargoes that delay access should demonstrate through empirical means the need for such embargoes.

In conclusion, we urge the development of an optimal public access policy approach that is as uniform as possible for all major federal research granting agencies, that is mandatory for all researchers funded in whole or in part by those agencies, that results in access to final author manuscripts that is as rapid as possible (with embargoes only where need is empirically demonstrated), and that allows for flexible rights of reuse. That approach will maximize the outcomes of the taxpayer's investment in research to the benefit of science, the economy, and the general public.

For information or questions regarding this submission, contact:

Ray English
Azariah Smith Root Director of Libraries
Oberlin College
Oberlin, OH 44074
440-775-8287
renglish@oberlin.edu

or

Laura Ada Emmett
Associate Librarian for Scholarly Communications
University of Kansas Libraries
Lawrence, KS 66045
785-864-8831
aemmett@ku.edu



American Society of Plant Biologists

12 January 2012

Submission for the Record: **Response to November 4, 2011, Federal Register Notice of Request for Information, OFFICE OF SCIENCE AND TECHNOLOGY POLICY, Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research; FR Doc No: 2011-28623**

Submitted by: Crispin Taylor, Executive Director, American Society of Plant Biologists
Phone: 301-296-0900; email: ctaylor@aspb.org

Electronically submitted to: publicaccess@ostp.gov

The American Society of Plant Biologists (ASPB) appreciates this opportunity to submit comments and would be delighted to continue working with the Office of Science and Technology Policy (OSTP) and other federal partners through a process of active engagement.

About ASPB

ASPB is a 501(c)(3) not-for-profit membership corporation created in 1926 and headquartered in Rockville, MD. Today, ASPB is an organization of approximately 5,000 professional plant biology researchers, educators, graduate students, and postdoctoral scientists with members in all 50 states and throughout the world. A strong voice for the global plant science community, the Society's mission—achieved through work in the realms of research, education, and public policy—is to promote the growth and development of plant biology, to encourage and communicate research in plant biology, and to promote the interests and growth of plant scientists in general. The Society publishes two of the most widely cited plant science research journals: [The Plant Cell](#) and [Plant Physiology](#).

As a publisher, ASPB plays a central role in the process by which plant biology research is developed, validated, communicated, disseminated, and ultimately accepted by the scientific community. To publish its two top-ranked journals, ASPB expends millions of dollars annually on peer review, editorial management, production, printing, shipping, distributing, and hosting its online journals on a fully digital, highly reliable platform.

Whether an article is read online or in print, high-quality peer review, page composition (XML), copyediting, and the listing and linking of bibliographic and reference data must be managed, necessitating considerable human capital investment in staff, in addition to scores of editors around the world. Our editors maintain the quality and reputation of our journals, utilizing the well-established system of peer review, whereby independent experts review submitted articles.

Accepted articles are those that pass muster based on established criteria, including novelty and significance of the research findings. Managing peer review for ASPB's journals is a complex undertaking. It requires sophisticated electronic resources, associated support personnel, and help from thousands of referees. Each year ASPB makes such necessary investments to fulfill its public nonprofit mission, generating an intellectual return through the dissemination of scientific research.

Introduction

ASPB aims to achieve the widest possible dissemination of the research results it publishes in its journals. Enabled by Internet technologies, ASPB in 2012 disseminates more information, more widely and more affordably, than ever before in its history. This accomplishment requires heavy investments in technology and infrastructure (such as an online platform) and business acumen to develop sustainable free and low-cost access models, whether by pay-per-view, article rental, or as a benefit of membership. But it is not just the cost of producing the articles that is important in driving the development of novel business models; it is their value to the community.

ASPB believes that it would be in the best interest of the United States government and all other stakeholders to strike a balance between public access and the needs and interests of the scholarly publishing industry because of the impact and value the latter brings to the progress of science and its contributions to American society and the national economy. Such a balance can be achieved based on shared principles, including the importance of peer review, the recognition of economic realities, the exploration and adoption of adaptable and viable publishing business models, the need to ensure secure long-term archiving and preservation of scholarly information, the increasing need to establish connections among disparate information sources and repositories online, and the desirability of broad access. One way to achieve this balance is for government to adopt a sensible, flexible, and cautious approach to drafting and revising public access policies—an approach that engages all concerned parties, including federal agencies, scientists, university administrators, librarians, publishers, and the public.

Indeed, it is ASPB's position that government agencies should develop flexible public access policies through voluntary collaborations with nongovernmental stakeholders, including researchers and publishers. Policies should be guided by the urgent need to foster interoperability of information across multiple databases and platforms. Agencies' efforts and resources could then be directed toward facilitating cyberinfrastructure and collaborative programs with and among agencies and other stakeholders to develop robust standards for the structure of full text and metadata, navigation tools, and other applications to achieve interoperability across the scholarly literature and other information sources.

ASPB Responses to RFI Questions

(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? According to trade association and other industry surveys of US publishers, both the nonprofit and commercial sectors already serve a robust, innovative global market for the access and consumption of peer-reviewed publications. Academic, corporate, and governmental research and education communities constitute primary segments of the market. Global revenue from scholarly journal

publishing was estimated at \$8.0 billion in 2008^{1,2}, with approximately \$3 billion attributed to the US market. The enterprise employs approximately 110,000 people worldwide, with 30,000 in the US. New publishers, journals, and business models evolve or emerge constantly, signaling a healthy, competitive marketplace. There is, to our knowledge, no evidence that the current system is in any way inimical to maximizing US economic growth, and there is no indication that the productivity of the American scientific enterprise is inhibited by it. So, ASPB's position is that there is no role or need for agencies to seek to grow existing or new markets related to peer-reviewed publications and no robust economic arguments for pursuing policies aimed at making articles publicly accessible.

Indeed, the combination of investments in digital and online technologies (by publishers and others) and the formation of library consortia in the US and around the world has accelerated and broadened access to peer-reviewed literature, and it has dramatically decreased the cost of such access. ASPB currently serves over 2,000 research institutions, and every person affiliated with these institutions has instant access to ASPB journal content online.

Furthermore, current conditions in the scholarly communications market already support a growing diversity of business models, as well as continuous innovation. It is our belief that the US government should support and encourage this diversity through its actions and policies, for example, by developing partnerships with publishers aimed at seeding further innovation and by providing funding support for experimental and innovative approaches toward increasing interoperability. (For more specific suggestions regarding partnerships and pilot projects that would meet mutually beneficial goals and conserve precious federal research funds for the agencies' primary mission of funding research, please see ASPB's responses to Question 5 later in this document. These recommendations for partnerships and pilot projects with federal agencies were developed in collaboration with a number of scientific publishers as we engaged over the past year in productive discussions with subject matter experts within the NSF and DOE, two US federal agencies that fund substantial research in the biological and physical sciences and engineering.)

As stated in the 2010 *Scholarly Publishing Roundtable* report³, many publishers have made the decision to move toward increasingly open structures and archives⁴ as enabled by Open Access business models and new solutions to associated permissions, such as Creative Commons⁵ licenses. These licenses provide a means for exercising certain rights regarding the re-use of an item. For example, these licenses could provide reuse rights if the resulting new works are also made available to the public. The *Roundtable Report* also notes that the number of journals making a change in business model is appreciable but small within the universe of more than 25,000 scholarly peer-reviewed journals⁶. ASPB echoes the *Roundtable Report* assertion that no existing digital business model has demonstrated its viability to the satisfaction of all, and we caution against de facto government endorsement of any single approach.

As part of the market's evolution and scholarly publishers' commitment to community and dissemination of peer-reviewed information, an increasing number of all types of journal publishers are electing to make their articles freely available to academics and others in 100 or more developing countries. Some well-known programs include the United Nations' HINARI, AGORA, and OARE Research4Life programs, in which ASPB's journals participate; HighWire Press's Developing Economies Program; and JSTOR's Developing Nations Initiative, in which the ASPB journals also participate. For descriptions of these and more, see www.library.yale.edu/~llicense/develop.shtml.

To meet the market's increasing demand for easily accessible quality information, ASPB invests considerably in new technologies for viewing and sharing its journals. For example, within the past year, ASPB has deployed a mobile phone reader for *Plant Physiology* and *The Plant Cell*. Such ongoing investments in existing products and services and the development costs for new products are funded through subscription fees and author payments. ASPB and many other scholarly publishers offer an immediate free access option for authors, and ASPB's journal *Plant Physiology* currently offers this option at no cost to corresponding authors who are members of the Society.

The ability for scientific publishers, large and small, for-profit and not-for-profit, to experiment with different publication, business, and access models is paramount and assures the vitality, diversity, and effectiveness of scholarly communication, leading to scientific and technological advances. Rather than mandate business models and de-incentivize market efficiencies, a more effective approach by government would be to incentivize the continued growth and vitality of the scholarly communication market for the benefit of the scholarly community and, in turn, the nation's competitive position. To that end, working with publishers, libraries, and other stakeholder communities, research agencies should identify specific needs of particular user groups and collaborate with publishers to meet those needs most effectively. Obviously, researchers, professionals, funders, and various segments of the general public (e.g., patients) have different information needs. ASPB is collaborating with other scholarly publishers to identify and address any existing access gaps through initiatives such as the low-cost article rental scheme pioneered by DeepDyve and the Research4Life consortium for developing countries (mentioned above).

To maximize the effectiveness of its efforts, government does have an important role to play in convening stakeholders to develop standards for data and metadata, thereby helping to make research more readily searchable and discoverable. Publishers are already working in partnership to develop standardized information and collections through initiatives such as CrossRef⁷.

With a relatively straightforward implementation of existing policy, government could make the funder-collected and maintained outputs of taxpayer-funded research, such as grant reports and research progress reports, freely available to the public⁸. Furthermore, to incentivize open access publishing, funds could be made available specifically to support payment for open access to published articles as pilot projects. Several research funders have already adopted this approach (e.g., Howard Hughes Medical Institute, Wellcome Trust, and Max-Planck Institutes).

In the same vein, government funding could be provided to license content from publishers in order to make it available to specific audiences. (Publishers license content to customers of many kinds, including government agencies, and have the ability to ensure its continued availability with existing infrastructure.)

ASPB has been a participant in working groups that are proposing and planning partnerships with NSF and DOE on access, linking of grantee reports to publications, data mining across agency and publisher databases, tools and methods for identifying publicly funded work, and potential pilot projects in these areas.

Government mandates for public access come at a significant cost to the US economy and to the scientific enterprise. Data from the National Institutes of Health's (NIH's) PubMed Central (PMC) repository indicate that two-thirds of PMC's users are from overseas. This suggests that critical export opportunities for the industry may be compromised, potentially resulting in the loss of US jobs⁹. Significant economic value added by the publishing industry could be wasted if revenue

derived from sales in the global market is compromised or eliminated because mandates require that articles appear for free on government-owned or operated websites. ASPB is actively involved in efforts to grow its business in Europe, Asia (including China), Latin America, and here at home. Government mandates that would require the ASPB journals to post content for free under a limited embargo period are bound to cut into those efforts and harm the Society's mission – including its capacity to continue to disseminate the peer-reviewed information published in its journals.

PubMed Central adversely impacts the US scientific enterprise in another way: by consuming financial resources for a duplicative and unnecessary repository that might otherwise go toward directly supporting the scientific enterprise.

In summary, ASPB believes that publishers should continue to be free to experiment with various business models in the marketplace of ideas and economics. ASPB endorses the Roundtable Report recommendation that "Agency policies should encourage the development, in a competitive landscape, of new value-added information products and services that take advantage of a scholarly environment in which articles are increasingly interoperable and available through licenses that support creative reuse. Such development should be carried out on a level playing field among all those who would devise such products and services." We believe that it is essential that any public access policies developed by the government do not undermine the ability of the market to create and sustain peer-reviewed journals.

(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?

ASPB and other scientific publishers rely heavily on the reputation of their journals to compete in the marketplace. Copyright protection reinforces the motivation for sustaining managed peer review, thereby protecting a journal's reputation. Any policy decisions regarding the publication and dissemination of peer-reviewed scholarly publications resulting from federally funded scientific research must respect US copyright law as it presently exists. Under the law, these works meet the criteria for copyright protection. It is a constitutional right granted to the copyright holder to exercise the exclusive rights attached to a work. In its role as the guardian of those rights, government must seek to strike the appropriate balance for all stakeholders through fair interpretation of the law.

It is ASPB's position that agencies should provide free public access to final research reports and link them directly to any peer-reviewed journal articles that are derived from the funding, regardless of the access mechanism via which those articles are available. This solution would drive the standardization of information reported on publicly funded research, promote rapid dissemination (rather than waiting for an article to be authored and subsequently peer reviewed), and ensure preservation of intellectual property rights, which provide the incentive for producing, distributing, and preserving all forms of intellectual property.

ASPB encourages agency policies and actions that work to ensure copyrighted materials are protected from unauthorized dissemination and piracy. Copyright is an essential ingredient in

promoting creativity, innovation, and the continued integrity and reliability of the scholarly record. There is some evidence that the NIH policy undermines intellectual property rights and promotes piracy of intellectual property. As noted in response to Question 1, the NIH public access policy and availability of articles through NIH's database, PMC, undermine an important US export market. Furthermore, copyrighted material downloaded from PMC appears on rogue Internet sites, resulting in significant annual losses to US publishers.

Nearly all scholarly publishers adopt liberal copyright policies, allowing authors to post copies of their manuscript on their individual and institutional websites with very little restriction, share copies with colleagues, and use their manuscripts for other educational and research purposes. Only commercial use is restricted and enforced by the industry.

(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?

A defining feature of the Internet is that information is dispersed and widely distributed. It is, nevertheless, readily discoverable. So, although a centralized data platform may have some potential advantages related to simplicity of operation, the use of a centralized, government-controlled platform for a large corpus of scholarly content has many significant downsides, not the least of which is increased and unnecessary costs to the government. A centralized approach discourages innovation by driving traffic away from innovators, including publishers, thus minimizing scientific and commercial opportunities.

However, an important role for government in this arena would be to drive and fund the development of interoperability standards that would facilitate and enable ever richer connections among journal articles and other types of scholarly information available online and promote the widespread adoption and use of such standards.

ASPB supports the recommendation of the *Roundtable Report* that states that government policies should be guided by the need to foster interoperability and encourage "additional multiagency programs supporting research and development to expand interoperability capacity and to develop and promote additional interoperability practices and standards." The *Roundtable Report* further notes that the NSF, DOE, and other agencies provide important funding for the development of interoperability capacities through their cyberinfrastructure programs.

In developing public access policies and procedures, agencies should carefully consider international cooperation with a larger vision that includes building standards and fostering distributed systems that are global in scope and go far beyond the work funded by US federal research dollars. In the Internet age, research and research resources are distributed globally. US federally funded research is only one part of the entire universe of information on any given topic, and in some disciplines, research is increasingly non-US government funded. A centralized repository such as PMC is not a model that is universally applicable or necessarily the best model for the future. Indeed, the success of the Internet is its evolving capability to connect an exponentially growing array of highly distributed information resources and databases. Any successful and optimized scientific publishing

system will incorporate effective incentives to implement and expand interoperability and reuse across internationally distributed databases.

It is ASPB's position that stewardship of publications in the Internet age should be the collaborative responsibility of the publishing, library, and research communities. US government involvement in the long-term stewardship of publications is best addressed as part of the copyright system and through the Library of Congress digital preservation initiatives primarily as a promoter of standards, as noted above, and as one of many stewards of specific data platforms that need to be linked across public and private boundaries.

What constitutes a publication and the nature of publication is changing with technology. A publication is no longer just a chunk of text fixed in time forever but a fluid representation. Publications can include supplemental material, multimedia files, software, and links to resources on the web and can be revised and corrected over time by the authors and publishers, hence the emergence of new community initiatives such as CrossRef's CrossMark¹⁰ service, which electronically watermarks an article's Version of Record (VoR), and DataCite¹¹, which extends the CrossRef-promoted Digital Object Identifier (DOI) to datasets. Any plan for the future should recognize that the static aggregation/library model is not likely to hold up well in the distributed and dynamic Internet milieu.

ASPB believes that it is unlikely that one optimal procedure for preservation and stewardship will emerge to become applicable across all of scholarly publishing. For now, ASPB strongly recommends that agency policies embrace diversity, decentralization, and interoperability. In the long term, systematic collaborations among stakeholders (government, publishers, universities and their libraries, and other not-for-profit participants in the scholarly publishing system) will be necessary to achieve maximum benefit. We note that libraries, in partnership with publishers, have established entities for preservation of digital documents that are already in wide use, for example, Portico¹² and CLOCKSS¹³.

Long-term stewardship of content comes at significant cost that is being borne by publishers and others. In an era of dwindling federal resources, central federal repositories are arguably duplicative, an unnecessary expense, and a recurring burden that may not be viable in the short or long term. Long-term stewardship might be more suitably carried out by the private sector or through collaborative stakeholder projects. There are productive ways to define appropriate roles of government and nongovernmental participants in the system, and ways that government agencies and nongovernmental stakeholders can collaborate as equal partners to their mutual benefit in strengthening the scholarly publishing system and expanding public access to its outputs.

(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?

Yes, please see detailed response to Question 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?

To facilitate public access and drive and support scholarship, agency databases should be able to communicate with each other. Each agency's policies should include at least a minimal set of common core properties that promote access to and interoperability among the content in all public access databases. Specifically, ASPB encourages agencies to develop collaborations and partnerships with scientific publishers to develop and implement:

- Standards and persistent identifiers to enhance the discoverability of research results and to promote interoperability among agency, publisher, and any third-party databases and platforms;
- Discovery tools to facilitate journal content mining; and
- Pilot projects that would drive access, use, and innovation from research results.

Specifics on these items are discussed below.

Beyond common properties, agencies should have the flexibility to manage and modify their policies in response to evolving circumstances. Each agency should fully engage researchers, institutions, and publishers working in fields that coincide with that agency's missions, both in establishing initial public access policies and in modifying those policies as appropriate over time.

Many scholarly publishing organizations, such as ASPB, were founded by scientists for scientists and fully embrace providing publishing and other services as their primary mission. As part of this objective, ASPB's executive director was an active member of the Scholarly Publishing Roundtable, and he has subsequently remained involved in working groups of nonprofit and commercial publishers that have proposed implementing joint projects with both the DOE and NSF with mutually agreed-upon goals.

Standards and Identifiers: Agency Funding Information

Most funding agencies currently require researchers to acknowledge in publications the support that they have received. There are no standards, however, on how this should be done. Consequently, agency funders find it difficult to know what publications have arisen from the research they have funded. ASPB supports the recommendation that publishers develop, in collaboration with funding agencies and CrossRef, means for standardizing funder information and making that information available to funding agencies and the public. We believe that a community-wide solution of this type will be easier and far less expensive to deliver than for each agency to develop its own response to the problem. This is because publishers are in the best position to provide a simple way of ensuring that journal articles are accompanied by standardized, high-quality metadata providing information about the agency, program, and even the specific grant that funded the research. It would be very expensive for agencies to obtain this information through data mining of existing publisher databases.

This proposal has been endorsed by CrossRef and a number of major scientific, technical, and medical (STM) publishing trade associations, including the Professional and Scholarly Publications Division of the American Association of Publishers (PSP-AAP) and the International Association of Scientific Technical and Medical Publishers. Related to this proposal, the DOE's Office of Scientific and Technical Information (OSTI) has agreed to maintain a registry of standard nomenclature for funding agencies and the associated naming and numbering system for grants. OSTI already

houses technical reports and data sets for more than 40 federal and international funding organizations.

With the successful implementation of this funding identity proposal by STM publishers, CrossRef, and the DOE, agencies would have access to standard metadata from published articles. By displaying this information on agency websites, visitors—from the research community to the general public—could follow the link (enabled through the DOI) to the publisher's platform where article abstracts are freely available and the full VoR (maintained by the publishers) is made available through a variety of access mechanisms, including innovative rental access models that give the public instant access for a modest fee. More than 40 scholarly publishers, including ASPB, are currently testing this particular access mechanism.

Standards and Identifiers: Promoting Interoperability

ASPB is seeking to collaborate with operators of a prominent knowledge base in plant biology that incorporates a rich array of genomic information from a wide variety of plant species to establish mechanisms for algorithmically connecting journal articles to database entries upon publication. Specifically, the collaborators propose to enable the retrieval of functional gene annotations and molecular annotations from ASPB journal articles using data-mining tools such as Textpresso¹⁴ and BioCreative¹⁵, both of which make use of Natural Language Processing and are organized around robust and highly structured ontologies. The collaborators plan to create a reference library that includes known and predicted gene names, symbols, functions, phenotypes, and pathway annotations in three target plant species. Together with the ontologies, which will play a key role in structuring data annotation, the library will also help establish data capture architectures that the ASPB journals would implement with their authors as manuscripts are being submitted, thereby directly, immediately, and algorithmically connecting published journal articles with the underlying datasets and knowledgebase. Both collaborators envision developing proof-of-concept data-mining methodologies that would be broadly applicable in other fields of research.

Standards and Identifiers: DOIs for Data Sets and Supplementary Material

Increasingly throughout the world, investigators are being asked to share or provide plans regarding how they will share with other researchers the primary data, samples, physical collections, and other supporting materials created or gathered in the course of their work. Grantees are expected to encourage and facilitate such sharing. Scholarly publishers are already participating in a number of initiatives designed to facilitate the voluntary sharing of data or to foster interoperability among data sharing repositories, and they would be willing to work with NSF, DOE, and other database/repository operators to develop recommended practices for assigning DOIs to data sets and supplementary material.

For data policies, publishers would draw on their experience with initiatives such as Opportunities for Data Exchange (ODE; see www.alliancepermanentaccess.org/current-projects/ode), which aims to gather and promote best practices on the way scientific data are treated, and CoData, a partner of the International Council for Science (ICSU) World Data System (www.icsu-wds.org). The goals of the relatively new ICSU World Data System (WDS) are to create a global federated system of long-term data archives and data-related services covering a wide spectrum of natural sciences, thereby encouraging interdisciplinary scientific approaches. For supporting information, publishers would draw on their involvement with the joint NISO/NFAIS Working Group on Supplementary Journal Information (see www.niso.org).

Standards and Identifiers: Author Name Disambiguation

Name ambiguity and attribution are persistent, critical problems embedded in the scholarly research ecosystem. ASPB encourages all federal agencies to work in collaboration with publishers as well as universities, funding organizations, and corporations from around the world to eliminate this problem through Open Researcher and Contributor ID (ORCID). ORCID is a recently established nonprofit organization whose goal is to establish an open, independent registry of researchers that is adopted and embraced as an industry-wide standard to resolve systemic name ambiguity by means of assigning unique identifiers linkable to an individual's research contributions. Researchers will be able to create, edit, and maintain an ORCID ID and profile free of charge and will define and control the privacy settings of their own ORCID profile data. Participants expect that accurate identification of researchers and their work will facilitate emergence of new services and benefits for the research community by all types of stakeholders in scholarly communication, from commercial actors to nonprofit organizations, and from governments to universities.

Discovery Tools: Content Mining

Content mining can be especially useful to the scientific community in driving interdisciplinary research and supporting the identification of new areas of discovery, and publishers are committed to managing content in modern digital formats to ensure that users gain maximum benefit. Scholarly publishers should work with funding agencies to develop pilot projects for journal content mining that would create thesauri, perhaps building on the ontologies that are used to define architectures for some types of databases, using their expertise to identify, organize, and analyze content to create conceptual links within and between highly technical subject matter. Although there are various ways to perform this type of processing, certain elements are common to all methods, including an automated way to process all sizes and types of content in which to identify relevant information and facilitate its extraction and analysis.

Such pilots should focus on goals such as the following:

- Structuring input text, deriving patterns within the structured text, and evaluating and interpreting the output;
- Extracting semantic entities from publisher content for the purpose of recognition and classification of the relations among them; and
- Enabling developers who wish to design and implement applications to analyze publishers' content, or test applications, as part of their research within publishers' content.

Consensus approaches within the community could also be explored for developing better standardized, mining-friendly content formats, a shared content mining platform, and common permission rules for content mining. The Publishers Research Consortium recently completed an instructive study on article-level content mining based on a broad survey of ongoing or planned activities among nearly 30 STM publishers or associations (see www.publishingresearch.net/documents/PRCSmitJAMreport20June2011VersionofRecord.pdf).

Pilot Projects: Sponsored Access to Published Research

The "Gold" Open Access dissemination model, whereby an author or their institution pays an article processing charge to the publisher, delivers immediate and unrestricted online access to the VoR. ASPB suggests that agencies could work with publishers to set up experiments in specific scholarly communities to answer the following questions dealing with the cost, benefits, and sustainability of the Gold Open Access model, as well as investigate how such a model should be funded and administered:

- How much would it cost an agency to fund Gold Open Access in the aggregate and on a per-article basis?
- What is the most effective method to provide Gold Open Access funding for authors? The ability to use grant funds for sponsorship? A separate pool of funding reserved solely for Gold Open Access sponsorship? Other means?
- Should authors be required to expend grant funds on publishing articles derived from that funding? If not, how can authors be encouraged to utilize the available funds?
- How can agencies best administer a Gold Open Access program?
- Does Gold Open Access offer agencies new opportunities to showcase the productivity of their funding activities to the American public and federal oversight committees?

Pilot Projects: Linking to/from Research Reports

ASPB encourages federal agencies to fund a pilot project that would seek to determine whether and how publisher content derived from agency-funded research could be mapped against agency research reports and other content. Specifically, the project might send users from publisher websites to the agency website to view free government-sponsored research reports and would, likewise, send users from the agency websites to publisher sites to view free abstracts and links to the VoR of articles connected to a particular research report or funded project.

If successful, this would result in interoperability between online agency content and publisher platforms. This is of interest to scholarly publishers because they would like to work with major research funders to identify, organize, evaluate, and highlight published results from federally funded research, as well as identify relationships, projects, and offerings that might be applicable to other research funders.

Possible outcomes of such a pilot might include:

- The ability to identify all agency-funded research within publisher offerings and the ability to deliver associated metadata to agencies
- The ability to establish mechanisms and approaches that could be implemented (for all research funders) across the industry
- A capability to report to major funders on the impact of the research they fund, for example, through bibliometric and other tools
- A "research dashboard" capability or the ability to contribute to one already in existence, for example, <http://rd-dashboard.nitrd.gov/>
- A mechanism for low-cost content rental access to the VoR of published articles and a mechanism to explore its impact
- Subject area content portfolios of agency-funded research articles for internal agency use (e.g., study sections)
- The possibility to use the DOE-OSTI platform (the <http://www.science.gov>) to extend this pilot to other federal funding agencies, and
- Models to illustrate how traditional publishing systems can coexist with self-archiving, including the posting of content on individuals' websites or in institutional repositories.

(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?

An excellent mechanism to ensure public access to federally funded research results is by providing access to final agency reports. Every federally funded research project is required by law to provide a detailed final report. The research reports are a condition of the government contract. These reports should be archived and made accessible to the public. Some science funding agencies make these reports freely available via the web, others do not. Making all such reports available and accessible in a comprehensive and systematic way would solve an essential public access problem. One leading example is DOE's OSTI, which publishes final reports online in a portal called Information Bridge. These reports are not journal articles, but the final reports are often much longer than the resulting journal article (if such article exists—researchers typically publish only positive results and then have to meet the publication standards of the journals in their field), more timely, and provide more information.

Moreover, NSF instituted a new reporting requirement as a result of specific legislation in the America COMPETES Act (Section 7010: Reporting of Research Results), which required that "all final project reports and citations of published research documents resulting from research funded in whole, or in part, by the Foundation, are made available to the public in a timely manner and in electronic form through the Foundation's Website." For several years, publishers have proposed working with authors to develop short abstracts for a lay audience to accompany each research report.

Publishers are partnering with federal agencies to develop policies that maximize public access to research results and provide easy links between research reports (detailing research results, perhaps including lay summaries) and the peer-reviewed VoR, including complete access to the abstract or summary. Such projects would result in interoperability between funder and publisher content, ensuring access and better reporting on the results of funding.

(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?

No. Publishers also invest in these other types of content used by researchers, often by conceptualizing the project, commissioning the content, and investing heavily in its development. Any kind of mandated access to that content is an expropriation of that content.

(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?

There is no "appropriate" embargo period after publication before the public is granted free access to the peer reviewed scholarly publications. Embargo periods should be consistent with the mission and business needs of publishers. ASPB believes strongly that a uniform access policy or mandate for scholarly publications would be an ineffective approach. Any overarching government-wide policy or embargo period would fail to accommodate such key factors as the specific needs of any given agency, the rapidly changing nature of scholarly publishing, and the unique considerations of the various fields of science and the journals that serve them.

References

¹ Cox, J. and L. Cox, *Scholarly Publishing Practice: Academic Journals Publisher's Policies and Practices in Online Publishing*, 3rd ed., ALPSP (2008),

<http://www.alpsp.org/Ebusiness/ProductCatalog/Product.aspx?ID=44>.

² Outsell, "An Open Access Primer-Market Size and Trends" (2009),

http://www.outsellinc.com/contact_us/open_access_primer_2009.

³ Report and Recommendations of the Scholarly Publishing Roundtable, January 2010, available at www.aau.edu/WorkArea/showcontent.aspx?id=10044. Referred to throughout this document as the *Roundtable Report*.

⁴ Morris, S., *Journal Authors' Rights: Perception and Reality* (London: Publishing Research Consortium, 2009), <http://www.publishingresearch.net/documents/JournalAuthorsRights.pdf>.

⁵ Creative Commons (<http://creativecommons.org/about>) is a nonprofit corporation that provides free licenses and other legal tools to mark creative work with the freedom the creator wants it to carry, so others can share, remix, use commercially, or any combination thereof.

⁶ Ware, Mark and Michael Mabe, *The STM Report: An Overview of Scientific and Scholarly Journals Publishing*. September 2009.

⁷ CrossRef (www.crossref.org) is a not-for-profit group founded by publishers in 2002 that now maintains 50 million items. Almost 1,000 publishers participate, assigning Digital Object Identifiers (DOIs) to published content items. Development of the CrossRef service has resulted in seamless navigation of the research literature by users so that researchers using the bibliography in one article can link from a reference to the full text of the referenced article.

⁸ This would ensure readability to the broadest audience. NSF is already pursuing such a policy, see <http://www.nsf.gov/pubs/policydocs/pofaqs.jsp>, and DOE through its Office of Scientific and Technical Information provides public access to nearly 300,000 DOE-funded research reports, see <http://www.osti.gov/bridge/>.

⁹ See, e.g., 2009 U.S.-China Joint Commission on Commerce and Trade (JCCT) Factsheet. Available at <http://www.ustr.gov/about-us/press-office/fact-sheets/2009/october/us-china-joint-commission-commerce-and-trade>.

¹⁰ CrossMark (www.crossmark.com) is a current pilot project of CrossRef to that will allow readers to easily determine whether they are looking at the publisher-maintained, stewarded version of a journal article.

¹¹ DataCite (<http://datacite.org>) is a not-for-profit organization established to facilitate easier access to research data on the Internet, increase acceptance of research data as legitimate, citable contributions to the scholarly record, and support data archiving that will permit results to be verified and re-purposed for future study.

¹² Portico (<http://www.portico.org/digital-preservation/>) is a digital preservation service provided by a not-for-profit organization with a mission to help the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways. It is among the largest community-supported digital archives in the world, working with libraries, publishers, and funders to preserve e-journals, e-books, and other electronic scholarly content.

¹³ CLOCKSS (*Controlled LOCKSS*) is a not-for-profit joint venture between the world's leading scholarly publishers and research libraries whose mission is to build a sustainable, geographically distributed dark archive with which to ensure the long-term survival of web-based scholarly publications for the benefit of the greater global research community (<http://www.clockss.org/clockss/Home>).

¹⁴ <http://www.textpresso.org/>

¹⁵ <http://biocreative.sourceforge.net/>

Adrian Pohl
Thu 1/12/2012 4:20 PM
RFI on Public Access to Peer-Reviewed Scholarly Publication

Dear people at the OSTP,

below are my answers to your questions on Open Access. I am responding as an individual working in an institution which provides information (research tools as well as licensed content) to academic libraries. Also, I am coordinating the Open Knowledge Foundation's "Working Group on Open Bibliographic Data". I make all my academic publications accessible on the web under a CC-BY license.

All the best
Adrian

(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 most efficient way to translate basic research findings into innovations that grow the economy is to allow as many innovators access to these findings as easily as possible. Thus, making scientific content openly available under open licenses such as CC-BY is exactly the kind of measure to enable and further research-driven innovation. (For a clear definition of the term "open" see <http://opendefinition.org/okd>.) The traditional entities to archive and make accessible scholarly works have always been (university) libraries. As such, federal agencies can and should support libraries to resume this task which has temporarily been outsourced to commercial publishers, which have prevented access by a subscription and copyright model which prevents innovators from accessing the latest research findings and has generated a rise in cost manifold beyond inflation as measured by the consumer price index. Thus, it is imperative to reduce the costs of public access to publicly funded research, while allowing as many innovators access to research findings as possible. Diverting those funds who are currently being accumulated with the shareholders of commercial publishers, towards libraries will cut publishing costs by orders of magnitude (through eliminating many middle-men) and enlarge the circle of potential innovators by orders of magnitudes, thus potentiating the current cost/benefit ratio exponentially.

(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?

I would dispute any actual intellectual property rights of publishers over publicly funded research. I am aware that legally there are such rights as researchers most often hand over their copyrights to these

publishers. However, since commercial publishers do usually add little to no value to the published research results (even peer-review is performed pro-bono by researchers themselves), this practice needs to end. Publicly funded research has been bought by the public and belongs to the public. Ensuring open publication licenses such as CC-BY for literature about publicly funded research is one of the possibilities to ensure the public retains its intellectual property on the research it funded.

(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?

Multiple private sources are always a suboptimal choice for long-term archiving: few private entities survive 'long-term'. Some libraries, in contrast, have been around for centuries, significantly longer than most private entities. Centralized access, however, suffers from as many cons as any monopoly. Ensuring the libraries of every research institution are sufficiently equipped to maintain long-term archiving of scholarly literature allows for a federated, decentralized archive of scholarly literature beyond any short-term financial fluctuations and allows for international collaboration for maximum safety through world-wide redundancy (following the motto "Lots of copies keep stuff safe"). A small fraction of the current subscription costs paid for by libraries would ensure such a long-term, publicly accessible archive.

(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?

Not to my knowledge. Given the instability of private sources and the decades-long history of price-gouging, I would argue that this would not be a good idea, either.

(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?

What is required is an evolving metadata and protocol standard that grows with the scientific enterprise and is under the control of scientists and developed in close cooperation with librarians. A recent innovation in this direction is BibJSON (<<http://bibserver.okfn.org/bibjson/>>). At best, this standard would build on Linked Open Data technologies, as this technology for exposing (meta)data on the web allows interlinking and thus makes the aggregation of citation data and indication of publication usage in university courses and analysis of this data technically easy. The actual location of the publications is, of course, irrelevant, as long as proper long-term archiving is ensured (see above). It is crucial that metadata for all scientific publications is made fully accessible on the web under an open license. See the Principles on Open Bibliographic Data for more detail: <<http://openbiblio.net/principles/>>.

(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?

The benefit is maximized by minimizing the costs associated with access. The costs are minimized by preventing third parties from adding costs to the process. One way to establish a short and thus cheap supply line is to have scholars deposit their work directly at their libraries, avoiding the costs of intermediaries such as publishers. The process of this deposition would still be identical to the current process (i.e., peer-review), albeit without intervening entities which withdraw funds but add little to no value.

(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?

Most definitely, yes, all of them. The public bought them and thus owns them. In some ways, scholarly work is nothing but commissioned by the public.

(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?

As the public already own these scholarly works, it is difficult to understand why there should be any embargo period to allow private entities, which have not contributed to the work, to profit from it. Every scholarly work that has been paid for by the public should be available to the public for immediate re-use and enter into the economy. It is hard to understand why there should be a waiting period for innovation to enter the market.

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

1/12/2012

This is a public and open document intended to draft a collective response to the request of information posted by the Science and Technology Policy Office (OSTP), on whether peer-reviewed publications resulting from federally funded research should be required to be made publicly available.

Dear Office of Science and Technology Policy,

Kitware applauds the initiative of the OSTP on seeking public feedback on these matters of high relevance to the scientific community and to the American public. However, please note that this is not an official Kitware response.

In order to contribute to this process, we reached out to our many collaborators and invited them to join us in writing a collective and thoughtful response to the insightful questions of the RFI. The result is the document attached to this submission letter. The names of the contributors and those in favor of this response are found at the end of the document.

Please find below our response to the RFI on “Public Access to Peer-Reviewed Publications from FFSR”. NOTE: In the responses below we use the following acronyms:

FFSR: Federally Funded Scientific Research

License of this Document: **CC0:**



To the extent possible under law, The Authors contributing to this Document have waived all copyright and related or neighboring rights to RFI Response. This work is published in: United States.

<http://creativecommons.org/publicdomain/zero/1.0/>

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 publicly 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?

Response:

Grow Existing Markets Related to Access:

A vibrant market of open access publishers has developed in the past ten years, with thousands of journals covering many different fields. The Directory of Open Access Journals [<http://www.doaj.org/>] currently lists 7,372 journals. Furthermore, several traditional publishers - such as [Springer](#), [Wiley](#), and [Nature Publishing Group](#) - have implemented Open Access options that enable authors to choose to publish their articles openly, either in Open Access journals or by way of hybrid Open Access schemes. In the latter, authors can choose to pay the publisher to allow their papers to be freely downloaded by readers, or use the traditional approach by which readers are expected to pay the publisher in order to get access to the papers, either via subscriptions or one-time payments for a given article.

Federal agencies should support the adoption of “Open Access” as the standard way of publishing the results of federally funded research. What has been termed the “author pays” model should be understood as a “funding agency pays” model for publishing fees in Open Access journals. The NIH, as part of its Public Access policy, has already stated that publication fees [can be charged to grant funds](#). This is a standard practice of funding organizations such as the Wellcome Trust and the Howard Hughes Medical Institute.

It is also fundamental to produce a clear definition of what “Open Access” means. We strongly propose that the definition of “Open Access” for articles resulting from federally funded research must be articles that are distributed by the copyright holder under the Creative Commons by Attribution License 3.0: <http://creativecommons.org/licenses/by/3.0/>. Under this license, anyone is permitted to copy, distribute, and create derivative works of the article, with the only requirement of providing attribution. We propose that “proper attribution” be defined as citing the uniform resource identifier ([URI](#)) or digital object identifier ([DOI](#)) of the original article. No further requirements for attribution should be demanded, and particularly, the attribution methods must not be left to be defined by authors on a case-by-case basis. Instead the federal agency must specify this standard method of attribution to ensure a reasonably low bar of effort that will lead to compliance.

Previous experience with the NIH Public Access policy has demonstrated that the open access policy is ineffective if it is not enforced. Federal agencies should therefore implement a system for verification of compliance, which should be reported as part of the “past performance” section of future funding applications. In this way, researchers’ compliance with public access policies will benefit applications for new funding.

An interesting and detailed set of suggestions for paying Open Access publication fees can be read here: <http://blogs.law.harvard.edu/pamphlet/2011/11/16/how-should-funding-agencies-pay-open-access-fees/>

Grow the Economy:

As an economic model, the goal of the scientific enterprise is to gather knowledge and information and to disseminate it in a usable form. The public availability of content will be the most important way of increasing efficiency and productivity of the scientific enterprise. Removing the barriers to scientific publications creates opportunities for developing new collaborations and for investigating previously unforeseen avenues of scientific research.

The public availability of articles will facilitate their access by academic institutions, companies, and citizens, and will reduce the amount of time it takes for research to impact small businesses and start-up endeavors. On a global scale, developing countries will profit immensely from free access to information as this will give them a chance to develop their own economies. Indirectly, the US and other countries will profit from the opening up of new markets.

The traditional process by which publishers request unpaid copyright transfers from authors, and then use those same copyrights to put articles behind toll-gates that restrict access to information for the 95 years awarded by copyright laws is detrimental to the further development of the scientific enterprise. Worst of all, they are an impediment to the education and appreciation of scientific research by the general public, as well as to the participation of the public in furthering those research efforts.

Examples of successful efforts for engaging the public in the practice of scientific research include the Polymath project, the NASA Galaxy Zoo, the Moon Crater Zoo, and the regular involvement of amateur astronomers in comet discovery. Michael Nielsen, in his book "[Reinventing Discovery](#)," goes into more detail on how massive participation of regular citizens, endowed with online collaboration tools, are transforming the practice of science. The need for and benefits of a networked society with free and unrestricted access to knowledge are dealt with in depth in the recent book "Too Big to Know" by David Weinberger.

Costs:

The current publication process is inefficient and not cost effective. For example, articles submitted by authors are unnecessarily deconstructed, retyped, and re-edited by the publisher to recompose a final version with only minimal incremental refinements. This process would be far more efficient and cost effective if publishing software was widely available for researchers to write initial versions of their articles in a collaborative fashion with their colleagues. In turn, articles can be passed directly to publishers using open standard file formats.

Federal agencies can help increase the efficiency and reduce the cost of publishing by supporting the development of open and royalty-free standards for scientific publications, and encourage commercial applications to implement these standards.

Types of Access Necessary:

The status of "public availability" must be defined in terms of:

- (a) Placing the articles in Public repositories, (without requiring registration or fees)
- (b) Distributing the articles under Creative Commons by Attribution License

Other licenses that impose restrictions such as "*non-commercial*" or "*only for research*" are not useful in practice, given that U.S. courts consider most universities and non-profit organization

activities to be of a commercial nature. Authors of papers must be required by federal agencies to retain the copyright of their articles (or at least be precluded from transferring copyright away) and in this way remain empowered to make licensing decisions about the articles. [Princeton University](#) and Harvard University faculty members have stopped the common practice of transferring copyright of articles to journals. The practice of not transferring copyright must become a policy of federal agencies to be applied to recipients of federal funding.

[See for example: <http://dx.doi.org/10.3897/zookeys.150.2189> “Creative Commons licenses and the non-commercial condition: Implications for the re-use of biodiversity information”]

[See Princeton University report <http://www.cs.princeton.edu/~appel/open-access-report.pdf>]

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 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?

Response:

In addition to the stakeholders listed in this question, it is critical to note that the general public is the primary stakeholder to be considered here. Given that in the context of FFSR it is the public’s tax dollars that are paying for the scientific research, the public’s interest should be considered foremost when considering trade-offs between available options.

In order to have a productive discussion on intellectual property, it is important to first deconstruct the term “intellectual property” and clarify its meaning in the context of current U.S. laws. We do this in **Appendix A** and conclude that **copyright** is the only concept of intellectual property that is relevant to this RFI.

Copyright is originated by the authors of articles, when they put the expression of ideas required to disseminate the outcome of their research in a tangible medium. In the context of federally funded research, authors are performing this work as part of their job duties. Therefore the articles are the outcome of “*work for hire*” and it is the employer of the authors who holds the copyright of the resulting articles. It is commonly the case that universities and other research institutions assign that copyright to the authors themselves, but this is a matter of policy choices by the institutions. It is a common commercial contractual practice that when one organization contracts another to develop creative works, the paying organization will retain some of the copyright rights (if not all) of the resulting creative work. In the context of federally funded research, it will be then consistent with common commercial practice of requiring awardee institutions to return the copyright of the articles resulting from federally funded scientific research (FFSR) to the federal agency. The U.S. government does not originate copyrights, but it can hold the copyrights of creative works when they are transferred to it.

For example, the Federal Acquisition Regulations: *FAR Subpart 27.4—Rights in Data and Copyright*. https://acquisition.gov/far/html/52_227.html

“(1) Data first produced in the performance of this contract.

(i) Unless provided otherwise in paragraph (d) of this clause, the Contractor may, without prior approval of the Contracting Officer, assert copyright in scientific and technical articles based on or containing data first produced in the performance of this contract and published in academic, technical or professional journals, symposia proceedings, or similar works. The prior, express written permission of the Contracting Officer is required to assert copyright in all other data first produced in the performance of this contract.

(ii) When authorized to assert copyright to the data, the Contractor shall affix the applicable copyright notices of [17 U.S.C. 401 or 402](#), and an acknowledgment of Government sponsorship (including contract number).

(iii) For data other than computer software, the Contractor grants to the Government, and others acting on its behalf, a paid-up, nonexclusive, irrevocable, worldwide license in such copyrighted data to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly by or on behalf of the Government. For computer software, the Contractor grants to the Government, and others acting on its behalf, a paid-up, nonexclusive, irrevocable, worldwide license in such copyrighted computer software to reproduce, prepare derivative works, and perform publicly and display publicly (but not to distribute copies to the public) by or on behalf of the Government.”

It has also been clarified that, for the purpose of Federal Acquisition Regulations, computer software falls into the category of data.

Federal agencies could therefore introduce requirements by which the copyright of articles resulting from FFSR should be transferred back to the federal agency.

As described in **Appendix A**, the economic logic behind copyrights does not hold for the case of articles resulting from funded scientific research. This is because in the context of scientific publishing, copyright does not have an economic role in “encouraging the creation of works of authorship” (since authors do not get paid by publishers), nor does it play a role in protecting creative works as scientific articles are expected to have a minimal amounts of “creative, invented” content.

The weak level of copyrightable content that a serious scientific research article should have, combined with the rationale that the purpose of copyright is to benefit the public by making available the result of creative works, and that the public has provided monetary compensation to produce those creative works, leads to the conclusion that copyright protection is not really required in the context of FFSR scientific research.

It is also important to clarify that in the standard practice of scientific publishing, publishers are not originators of intellectual property. The real creators are the researchers who write articles intended for publication. Publishers acquire intellectual property on those articles through the practice of requiring authors to sign copyright transfer agreements as a condition of publication. In these transactions, publishers do not provide any monetary compensation to authors. In economic terms, the transfer of copyright from authors to publishers is essentially a donation.

As a result, there is no need to provide any protection for publishers, as they are already acquiring for free a product in which they have not invested any significant financial resources to produce. The contribution of the publisher is limited to coordinating the work of associate editors to compose the final collection of articles for publication, to host the digital documents online, and to provide the gates that regulate access to the publications to paid subscribers only.

In most cases, associate editors and reviewers who contribute the bulk of the peer-review process are volunteers who are not paid by the publisher. Therefore, it should be questioned why publishers benefit for free from the work that scientific researchers as authors create under the support of FFSR, rather than the taxpayers, who are paying for the bulk of the research enterprise and are the rightful copyright holder.

This doesn't mean that publishers do not need to be paid for providing the service of disseminating articles. As providers of commercial services, publishers certainly deliver a valuable contribution to the scientific enterprise and must be compensated for such services. However, that compensation does not have to be achieved at the price of restricting access to FFSR articles. A variety of business models that make Open Access a viable financial endeavor have been demonstrated in the past ten years.

Our point is that intellectual property, particularly copyright, is not needed in this economic transaction that compensates the publishers for their services; under modern business models of open access publishing, there is no need for publishers to hold the copyright of the articles. Instead, publishers simply need to be the recipients of a license given by the copyright holder, allowing the publisher to copy, distribute, create derivative works, and perform public displays of the articles. The prime examples of licenses suitable for this purpose are the Creative Commons by Attribution license, and the Creative Commons Share Alike license.

Note that other Creative Commons licenses, such as the CC Non-Commercial license and the CC No Derivatives license, will not be suitable for allowing publishers and other institutions to productively use the articles resulting from FFSR.

More can be found on these topics at:

- <https://svpow.wordpress.com/2011/10/22/economics-of-open-source-publishing/>
- <http://blogs.lse.ac.uk/impactofsocialsciences/2011/11/09/functionality-academic-publishing/>

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?

Response:

Centralized Option

Pros

- It may facilitate the creation of uniform methods of accessing and searching for articles.
- As with Apps platforms such as the Android and iPhone operating systems, a cohesive platform can lead to innovative development of different means of accessing, searching, deconstructing, and analyzing articles.

Cons

- It tends to result in bottlenecks, delays, congestion, and a lack of flexibility and agility.
- Centralization results in creating a single point of failure, where the entire system depends on a critical piece to be working all the time.
- It will take longer to be put in place.
- It makes it very difficult to innovate over time and to introduce new functionality that can transform the way that the data is used.

Decentralized Option

Pros

- It spreads and distributes the load of the system across multiple archives.
- When combined with smart redundancy, it provides protection against potential loss of information.
- It provides an open market for innovative methods to evolve, which enables researchers and the public to consult and data-mine the content of scientific publications.

Cons

- It requires a concerted and coordinated effort to define standard mechanism for
 - Replication of data
 - Federated search
 - Interoperability

Conclusion:

The system should be decentralized based on agreed standards and interoperability. Federal agencies should host archives of the published materials, but those archives should be commonplace and be replicated in different institutions (for example, university libraries).

Wide replication is the best way of ensuring continuous availability. This method is the essential mechanism used by the Internet itself, and has also been demonstrated by large scale source code repositories in Github (<http://github.com>), where some of the most popular code repositories have been replicated thousands of times (for example, see <https://github.com/popular/forked>).

Rich environments of replication, combined with SHA1 hashing that makes it possible to verify differences between multiple copies of a resource, guarantee the perpetual availability of a digital resource. To be more specific, a worldwide cataclysm would be the only way to wipe-out all copies of the “rails” repository, for which 2,569 copies have been made available worldwide: [\[https://github.com/rails/rails/network\]](https://github.com/rails/rails/network). Distributed replication, versioning, searching, and indexing are standard features in peer-to-peer software applications, of which several open source implementations are available.

A decentralized storage solution, however, must be paired with a federated system of indexing and searching for content to ensure ease of search and access to the publications. Such systems are widely available and have been used to support many legal applications of peer-to-peer networks.

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?

Response:

University libraries, archives, and public libraries already archive articles and provide long-term stewardship of the results of FFSR. It has been only due to the recent publishing practices of copyright control, such as DRM and online-only-licensing for access to articles, that libraries have been prevented from playing their natural role of long-term stewards of published content. This has been true for many years in the case of books.

Long term preservation of published materials is the job of libraries and archives, not the job of publishers. Libraries and archives have a much better guarantee of longevity than publishing businesses and societies that provide the services today.

Once federal agencies implement policies that preclude researchers from transferring copyrights to publishers, and that require researchers to make articles available in public repositories using appropriate licenses, libraries and archives will be able to regain their historical role as long-term stewards of these published materials.

No individual organization can be a sole, reliable provider of long-term, fail-safe storage for the large body of articles resulting from FFSR. As the Internet itself has demonstrated, only a distributed, decentralized system built upon light and open standards can provide reliable, long-term, and innovative support of the public dissemination of information.

The adoption of permissive practices on copyright and licensing for FFSR publications will be of fundamental importance in enabling the unfettered replication of articles in any medium, including digital ones. It will therefore empower decentralized systems to host replicated archives of the articles, along with experimenting with innovative technologies for maximizing the dissemination and collective exploitation of the information contained in the articles.

There are open access journals and publishers that are examples of successful models of publishing innovation and stewardship, such as the [Insight Journal](#), [PLoS](#), and [BiomedCentral](#). These journals foster accessibility to the results of scientific research and are creating a new paradigm for scientific publishing. More importantly, they are reviving the support for verification of reproducibility, which should be the hallmark of scientific research.

Question 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?

Response:

Adopt standards of publishing technology that:

- Do not rely on proprietary formats and
- Are not subject to proprietary restrictions (patents or copyrights).

There is an abundant body of publishing technology that is openly available to the public. Examples include RTS, Latex, HTML, Wiki formats, ODT, ebooks.

All the adopted formats must be machine readable (digital) to facilitate indexing and large-scale data-mining of the literature. The Library of Congress, in collaboration with the National Library of Medicine should define a minimalistic schema of metadata, and it should be done in less than six months. Much of this work is already done by PubMed and Medline.

Standards of unique resource identifiers such as the ones provided by “handle.net” should be required.

Question 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?

Response:

1. Streamline the process of paying for publishing services. For example, a standard R01 grant should have a pre-specified budget for paying for publication charges in open access journals based on the “authors-pays” model (<http://www.plos.org/publish/pricing-policy/publication-fees/>). Publishers should negotiate their rates with the government in the same way that all other contractors (from service providers up to manufacturers of military equipment) do.

Negotiated rates are already a requirement for all universities and companies that receive grants from and do contracted business with the federal government. There is no reason why publishers shouldn't be subject to the same conditions of rate negotiation when they are providing services to the federal government.

[For more details, see: https://www.acquisition.gov/far/html/Subpart%2042_7.html]

2. Implement a distributed system of repositories that provides redundancy of storage along with a distributed system for indexing and search that can be easily navigated without a single point-of-failure or bottlenecks.

This technology is already available in the form of peer-to-peer networks, for which multiple free and open source software implementations are available.

3. Define a specific set of copyright licenses that will be admissible for labeling articles as “publicly available,” and then require that all articles resulting from FFSR be made available by distributing them under one of these accepted licenses.

In particular, these licenses must not have any restrictions on the commercial use of the content, and must allow for modification and redistribution of the copyrighted content. Ideally, this would be the Creative Commons by Attribution license 3.0, and the Creative Commons by Attribution Share Alike 3.0 licenses. The Creative Commons Non-Commercial license should be excluded. This is consistent with what open source communities did for the open source definition, which requires that licenses allow for modification, redistribution, and commercial use of content.

[For more details, see: <http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0030097>]

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?

Response:

Yes, from both an economic point of view and that of promoting the progress of science and technology, all non-classified information that derives from FFSR must be made publicly available. This must include book chapters, conference presentations, articles in conference proceedings, audio recordings, podcasts, video recordings, and training materials whose content is **substantially** based on FFSR results.

If the public has paid for the development of any of these materials, then the public must have unrestricted access to them. Note that this is not “free access,” given that the taxpayers have indeed *already* paid for those materials beforehand. This is simply returning to the public what the public has paid for.

More on this topic at:

[\[http://www.nytimes.com/2012/01/11/opinion/research-bought-then-paid-for.html?_r=2\]](http://www.nytimes.com/2012/01/11/opinion/research-bought-then-paid-for.html?_r=2)

Aside from classified and export controlled materials, the only other exception that should be made is the protection of the privacy of human subjects participating in medical research. Note however, that once medical datasets have been anonymized properly, they should fall in the category of public dissemination. Federal agencies, particularly the NIH, should create an easy option for patients to consent to share their medical information if they wish to do so once properly informed of the implications, both in the sense of risk and in the sense of the potential benefit for the advancement of scientific 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?

Response:

No embargo period should be required or allowed. Articles should be made publicly available immediately after being published. In a typical FFSR project, the American taxpayer has paid for the research with one or more years in advance by dutifully paying their contributions to the federal budget in the form of taxes. There should not be any further delay in making the results of the research available to taxpayers, as they have already paid for it.

Publishers' business models must be restructured in such a way that they are no longer an obstacle to the public dissemination of scientific information. The viability of such models has already been demonstrated by open access publishers, and with the options offered by hybrid publishers; that is, publishers that offer authors the choice of processing their articles as open access articles, or as traditional closed access articles.

The cost of publishing is about 1% to 2% of the cost of performing research. This cost is already paid by the federal agencies through the indirect channels of overhead (indirect costs) that goes to finance the operation of research institutions, including their libraries, and particularly the subscriptions that the libraries pay to publishers. It would be a lot more efficient to clearly incorporate the cost of publication upfront into the preparation of research proposals and utilize such a fraction of the budget to pay for the publication fees of open access publishers. Notice that this doesn't at all diminish the peer-review process that is required to ensure the high quality of content, given that this activity can continue to be performed on a volunteer basis, as it is done today. Publishers today do not pay authors, reviewers, or associate editors for the work they contribute to the endeavor of preparing and reviewing articles for publication.

Publishers should be paid up front from grant funds, so they will not need to engage in the practice of using copyright to implement toll-gates that restrict the public's access to FFSR results. Instead, publishers just need to receive a license to publish the FFSR articles from the authors or their institutions. In this way, articles can be made immediately available to taxpayers and the general public, the rightful copyright holders of the articles content. Having been compensated for their services, publishers will not need to further restrict access to readers.

For more suggestions on how to pay for open access, please see:

<http://blogs.law.harvard.edu/pamphlet/2011/11/16/how-should-funding-agencies-pay-open-access-fees/>.

This reference above discusses the topic of balancing library budgets with an open access payment fees taken from grants. It also includes discussion about funding libraries to play the role of archives, and as nodes in a decentralized system that facilitate access to FFSR results. This is after all, what Web technology was invented for at CERN.

[\[http://public.web.cern.ch/public/en/about/web-en.html\]](http://public.web.cern.ch/public/en/about/web-en.html)

Signatures

Name	Title	Institution
Luis Ibanez	Technical Leader	Kitware Inc.
Katie Osterdahl	Communications Specialist	Kitware Inc.
Katie Sharkey	Communications Specialist	Kitware Inc.
Arno Klein	Asst. Prof. Clinical Neurobiology	Columbia University
Tom Olijhoek	Scientific Consultant	Dutch Malaria Foundation
Pascal Christoph	Developer	hbz-nrw.de
Cameron Smith	Computational Scientist	Rensselaer Polytechnic Institute
Wesley Turner	Technical Leader	Kitware Inc.
Raphael Ritz	Scientific Officer	International Neuroinformatics Coordinating Facility (INCF)

Appendix A - Intellectual Property in Scientific Data.

The term of “intellectual property” is commonly used as an aggregate of the concepts of

- Copyright
- Patents
- Trademarks
- Trade secrets

In order to understand how these concepts apply to the challenge of maximizing access to the results of scientific research funded by the federal government, it is important to analyze the concepts independently.

Copyright is a government-awarded monopoly given to the creators of works of art. This monopoly awards creators the exclusive right to (1) reproduce the work, (2) prepare derivative works of it, (3) distribute copies of it, (4) perform it publicly and (5) display it publicly. The duration of copyright is: (a) the lifetime of the authors plus 70 year, (b) 95 years for works created by a corporation, or (c) 120 years for unpublished works created by a corporation. The goal of copyright is to provide an incentive to the creators of works of art by giving them exclusive rights on the exploitation of the works for a limited time

In the context of dissemination of scientific data, the economic bargain of copyright bears very low or no relevance, given that researchers (those who acquire and process the data) do not get paid when publishing that data. Instead, they get funded proactively for performing the research that leads to gathering information that is later published. Therefore, a very concrete economic incentive has already been provided and delivered to the researcher in the form of funding that American taxpayers have invested in the acquisition of the data.

As opposed to a novelist, whose income is purely based on the sale of copies of her/his book, the salary of a researcher is based on their performing the duties of scientific research. Granted, publishing datasets is part of such duties, but it is not equivalent to the creative activity of writing works of art (such as novels, music, or poems). Given that, in the context of FFSR, researchers are already paid by the public beforehand and so there is no need for the economic incentive of copyrights to address any “market failure” on the production of public goods (in the economic sense of non-rival and non-excludable goods), as is the case for novels, poems, and music. On the contrary, once the FFSR data has been acquired, every day that passes without this data being publicly shared is a day in which economic waste takes place and the economy at large performs less efficiently. It is also a day in which American taxpayers do not get anything back from the funds that they provided to the research enterprise.

Additionally, the nature of scientific research requires that the content of scientific datasets must be measurements of facts and should be devoid of any “creative elaborations”. In other words, the more “scientific” a dataset is, the less “creative artistic content” it should have in it; therefore, the less it deserves the protection that copyright is intended to provide to creative works of

authorship. The creativity of the researchers lies in the definition of the acquisition protocols, the experimental design, and in the specific apparatus or software used during the data acquisition, which sometimes are made especially for a specific dataset. The dataset itself, on the other hand, shall not include any creative content. A high quality scientific dataset must be a concise collection of facts, measurements, and computations on those measurements. Datasets with high levels of “creative content” are by definition not scientific datasets, and should not be produced as the outcome of federally funded research, or any other process that aspires to be called “scientific”.

Patents are government-awarded monopolies on the commercial exploitation of an invention. This 20-year long monopoly is awarded to the inventors in exchange for the public disclosure of the invention, and its eventual delivery (at the expiration of the patent term) to the Public Domain. Given that public disclosure is a requirement of the patent economic bargain, for awarded patents there is no concern about including information in articles intended for publication. The full information about the invention should already be publicly available at the U.S. Patent Office at the time that the patent is awarded to the inventors. Data is not “patentable subject matter” given that it is not the result of a creative process and is not useful, non-obvious, or novel. Datasets collected in the course of scientific endeavors are expected to be a collection of factual data, and therefore, they are as far as they can get from the type of “creative” work that patents are intended to protect.

Trademarks are symbols, designs, and terms that identify a product, service or company in the public marketplace. They are intended to prevent confusion in the marketplace, to protect the reputation of the producers of goods and providers of services, and to reduce the transaction cost that consumers have to invest in finding good and services that satisfy their needs. In the context of dissemination of scientific data, trademarks play a minimal role given that datasets are not supposed to be mechanisms of marketing goods and services. It is actually contrary to ethical standards in the scientific research field to use dataset publication as a venue for promoting goods and services in the context of commerce.

Trade Secrets refer to information that organizations keep confidential. For a piece of information to be considered a trade secret, it must have some value and derive part of its value from the mere fact of being secret. Trade secrets are managed via contracts, typically established between organizations in the form of non-disclosure agreements and between organizations and their employees in the form of confidentiality clauses that are incorporated in employment contracts. It is the responsibility of the institution to take affirmative steps to prevent its confidential information from becoming public.

In the event that a piece of confidential information is leaked publicly, there is no legal protection that can prevent the further dissemination of such information, except from forbidding an intruder to make use of data that was acquired illegally (e.g. by trespassing into private property). Therefore, in the context of dissemination of scientific data, trade secrets are only relevant as a context in which institutions should establish policies and verification mechanisms that prevent confidential information from being included in any dataset that is submitted for public release. It is the responsibility of the institution and its employees to protect such confidential information. Once data is published, the institution has relinquished its claim for such data to be considered a trade secret.



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

Office of Science & Technology Policy

January 2012

Response from Enabling Open Scholarship (EOS)

A. Opening statement

Open Access to federally-funded research will advance American science and technology, speed up research, reduce duplication, increase the usage and impact of research, facilitate interdisciplinary research, improve the quality of research because greater scrutiny will be possible, enable the deployment of new semantic technologies to create new knowledge from existing research findings, provide the wherewithal for better, smarter research assessment and management and provide greater payoff for the US taxpayer from the funds invested in research across all federal agencies.

A number of constituencies will benefit: as well as the research community itself, which will have immediate and untrammelled access to the information it needs to do its work, the professional, practitioner and lay public communities will also benefit. These things will lead to wealth creation, improvement in the quality of life and a better informed populace in an increasingly scientific/technological world. Access must be made available in ways that permit full re-use of research results and through services that maximise ease of use and convenience for the relevant user constituencies. We elaborate on these points in our detailed response below. Our response is organised by answering the questions listed in the *Request For Public Comment*.

B. Enabling Open Scholarship (EOS)

EOS is an organisation of universities and research institutes worldwide whose managers have come together to discuss, shape and promote the principles of open scholarship. EOS has members on six continents, from the largest, broad-based universities and research institutes to some of the smallest, most specialized research-based institutions. As well as universities and research institutes, EOS also has government departments and research councils (analogous to the NSF or NIH) in several different countries as members. Board members are listed at the foot of the document.

C. EOS' responses to the questions in the Request For Information

(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 argument that there can be economic benefits from public access is now substantiated by evidence.

First, the economic modelling work of the Australian economist, John Houghton, on the economic benefits of moving to a fully Open Access scholarly communication system has shown that there would be savings for all national economies studied (Australia¹, United Kingdom², The Netherlands³, Denmark⁴, USA⁵). Most importantly in the context of this response to the OSTP RFI, the US study – which modelled the economic effects of the Federal Research Public Access Act – showed that the incremental benefits of a mandatory Open Access policy over a period of transition of 30 years would be valued at some 8 times the costs of implementation, with the proportion accruing to the US itself of some 5 times the implementation costs.

The methodology used by Houghton has been criticised by some publishers for misrepresenting their costs, though his model is available for anyone to use and these publishers have never populated it with their own data to show what they believe is the true situation. At least they have not done this and made the findings publicly available. Nor have they made what they claim to be the true cost data available so that others might use them to populate the model. The criticisms from the publishers have been dealt with in a public statement by the sponsor of the UK study, the UK's Joint Information Systems Committee (JISC)⁶.

¹ Houghton J, Steele C and Sheehan P (2006) Research communication costs in Australia: Emerging opportunities and benefits: report to the Department of Education, Science & Training. <https://digitalcollections.anu.edu.au/handle/1885/44485>

² Houghton, J *et al* (2009) Economic Implications of Alternative Scholarly Publishing Models: Exploring the costs and benefits [United Kingdom study] <http://ie-repository.jisc.ac.uk/278/>

³ Houghton J, de Jonge J & van Oploo M (2009) Costs and Benefits of Research Communication: The Dutch Situation. http://www.surffoundation.nl/SiteCollectionDocuments/Benefits%20of%20Research%20Communication%20April%202009_%20FINAL_logos2.pdf

⁴ Houghton J (2009) Costs and benefits of alternative publishing models: Denmark. http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fdownloads%2fDK_Costs_and_benefits_of_alternative_publishing_models.pdf

⁵ Houghton J (2010) Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs [US study] <http://www.arl.org/sparc/publications/papers/vuFRPAA/index.shtml>

⁶ JISC Response to: Some comments prepared jointly by The Publishers Association, the Association of Learned and Professional Society Publishers and the International Association of STM Publishers on the report "Economic Implications of Alternative Scholarly Publishing Models: Exploring the costs and benefits"

Second, there is the problem of access to research information for companies and the effects that this has on innovation. The European Commission's own *Community Innovation Survey* has shown that there is a 'weak link between innovative enterprises [mainly small- and medium sized businesses, SMBs] and public research institutes/universities' and that 'innovative enterprises find the information they need more easily from suppliers or customers than from universities or public research institutes'⁷. Another study on accessibility of university research to SMBs showed that while 71% of respondents in innovative companies find accessing articles fairly/very easy, 66% of respondents pay for access in the form of subscriptions or society memberships which is costly. Moreover, there is 'by definition, a minority (29%) for whom access was fairly or very difficult'⁸.

There is now some early evidence of the actual economic costs and benefits to SMBs from access problems. Work carried out in Denmark on behalf of the Danish Government [disclosure: one of us was a co-author on this study] showed that 79% of small-medium sized innovative businesses had problems accessing the basic scientific research information they need. Difficulties in accessing research articles costs €73 million (circa USD 94 million) *per annum* to Danish firms. Product development is delayed or abandoned without access to research articles. The value of academic research to sales is around €2.1 million (USD 2.7 million) per company *per annum* and the value of delays, in lost sales of new products, is around €4.8 million (USD 6.2 million) *per annum*⁹. The businesses surveyed for this study ranged from biotech companies through engineering, construction, software and environmental services to horticulture and plant breeding.

This issue has been explicitly acknowledged in the UK: the Minister for Science, David Willetts, has set up a working group to study how to broaden access to research article sand himself concludes that 'Research stimulates and fuels innovation and economic growth. So, to maximise UK innovation we need to maximise access to and the use of research findings'¹⁰.

Finally, a recent study has also underlined the benefits to the private sector in the UK from access to research results¹¹.

There is no reason to suppose that the need for scientific information is any less for similar innovative US companies than it is for Danish or British ones and anecdotal evidence indeed

<http://www.jisc.ac.uk/media/documents/publications/responseoneiaspmreport.pdf> by Houghton et al. & Oppenheim et al., commissioned by JISC (published January 2009)

⁷ Parvan, S-V (2007) *Statistics in Focus: Science and technology*, 81/2007.

http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-07-081/EN/KS-SF-07-081-EN.PDF

⁸ Ware, M (2009) Access by UK small and medium-sized enterprises to professional and academic information <http://www.publishingresearch.net/SMEaccess.htm>

⁹ Houghton J, Swan A & Brown S (2011) Access to research and technical information in Denmark (Adgang til forskningsresultater og teknisk information i Danmark) <http://www.fi.dk/publikationer/2011/adgang-til-forskningsresultater-og-teknisk-information-i-danmark>

¹⁰ <http://nds.coi.gov.uk/content/Detail.aspx?ReleaseID=421232&NewsAreaID=2>

¹¹ HOST (2011) Benefits to the private sector of Open Access to higher education and scholarly research. http://open-access.org.uk/wpcontent/uploads/2011/10/OAIG_Benefits_OA_PrivateSector.pdf

suggest they share the same problems¹². Maximising access to research information for these sectors enables them to do their innovative work more easily, with economic and social benefits that result for society at large.

(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 is assigned to authors by law (unless the employer opts to assert rights over work produced by employees) with the expectation and recognition that authors will benefit from the opportunity for their work to be used and to have impact and benefits for others. Creators of scholarly works are rewarded in terms of career advancement and personal achievement by building on their findings themselves and by having others build on them also. There is no conflict with public access principles: indeed, increased access and use of scholarly outputs results in greater impact in academic terms (citations) and in other measures of societal value. Federal policy should acknowledge these points and either explicitly require authors to retain sufficient rights to make their work publicly available under any terms laid down by federal policies, or require that authors transfer sufficient rights to the relevant federal agency to enable the agency to make the work publicly accessible. This is not a novel position. It is the basis of the policy currently in existence at the NIH and, moreover, many universities around the world retain rights to make their researchers' work publicly accessible, or are formally assigned that right by their researchers by agreement.

Since copyright is always a bundle of rights rather than one entity, the right to publish the work and make money from that can be transferred to publishers through a Licence To Publish (LTP). Thus publishers' interests are also legally protected. A number of publishers do NOT require transfer of the full copyright bundle and are happy with an LTP¹³: there is no reason why this should not extend to the majority where prior policy conditions upon authors make this the reasonable and workable solution.

(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

¹² "With a small oncology company ... it is imperative that I have access to the literature. But small companies do not have the "deep pockets" necessary... The for-profit journal publishers have effectively barred access to key scientific information except to those who can afford their outrageous fees. Much of the most innovative work is being done at companies like mine that cannot afford to pay \$30+ per paper or pay per-search charges in abstracts or journal collections." Terence Dolak, SDR Pharmaceutical, New Jersey. http://blogs.openaccesscentral.com/blogs/ccblog/entry/unemployed_retired_might Lose_touch

¹³ <http://users.ecs.soton.ac.uk/harnad/Hypermail/Amsci/7801.html>

government can ensure long-term stewardship if content is distributed across multiple private sources?

There are reasons why a Federal agency should keep custody of all published content. The most important are: (i) for internal research management and monitoring purposes (ii) for preservation and curation (iii) so that the contents can be enhanced (better metadata, improved mark-up) to enable science to work better.

We suggest, however, that it is not necessary for content to be *deposited* centrally. If the right metadata schema is embraced by all relevant institutional repositories, it is technically simple to harvest the content appropriate for the relevant Federal agency's own archive.

This is the model recommended some years ago for national-level Open Access collections¹⁴. It is also the one now adopted by the European Commission for its own-funded research: the Commission-funded OpenAIRE repository is harvesting European-funded works from institutional repositories across the European Union, and the Commission's policy requires those works to be deposited locally wherever there is a suitable institutional archive. Other national Open Access collections have also adopted this model¹⁵.

The advantages to this model are that institutions are already equipping themselves with repositories, so the basic infrastructure is already being put in place and, importantly, institutions can be partners to funding agencies in monitoring and policing mandatory policies (both their own and on behalf of funders). Indeed, evidence shows that so far, notwithstanding the much-improved compliance rate for the NIH public access policy, the highest rates of compliance with mandatory policies are still seen at universities that have such policies and conscientiously support and monitor them locally¹⁶.

(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?

Long-term stewardship of research outputs is undertaken by libraries and by publishers as well as specific preservation services that may be public or private concerns. University and research institution libraries have the relevant expertise to provide preservation and curation services at least into the medium term: the academic library community as a whole has the organisational wherewithal and capabilities to determine that policies and practices are put in place to ensure the safe keeping of scholarly material into the long term.

While there is always scope for encouraging public-private partnerships to create better services, the overall goal remains that access be free for both current and past literature. There are examples where public-private partnerships work to this end, such as in the case of UKPMC, the UK site for PubMed Central. This is funded by both public (UK research

¹⁴ Swan, A., Needham, P., Proberts, S., Muir, A., Oppenheim, C., O'Brien, A., Hardy, R. and Rowland, F. (2005) Delivery, Management and Access Model for E-prints and Open Access Journals within Further and Higher Education. Technical Report, JISC, HEFCE. <http://eprints.ecs.soton.ac.uk/11001/>

¹⁵ For example, Ireland, amongst others: <http://rian.ie/en/static/Aboutus>

¹⁶ Gargouri, Y., Hajjem, C., Lariviere, V., Gingras, Y., Brody, T., Carr, L. and Harnad, S. (2010) Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research. *PLOS ONE*, 5 (10). e13636. <http://eprints.ecs.soton.ac.uk/18493/>

councils, the British Library) and private (medical charities) funders. All parties are committed to providing public access for the long-term, in the interests of the public, research and the missions of the sponsors.

Where legacy literature is in the hands of private publishers, as is the case for the publishers' archives of versions-of-record of journal articles, there is certainly room for consideration of how public access might be provided to that material. However, the most important thing is that policy ensures that this is not the *sole* means of preserving the literature since private publishers' interests are served by access restriction rather than access maximisation.

(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?

The research literature should be made available through OAI-PMH-compliant¹⁷ repositories or journal sites. The ideal is to have a metadata set that describes adequately the provenance and content of the article, and the funder and grant award information, so that searching for outputs of particular research programmes is enabled. For full interoperability, a machine-readable licence should also be part of this metadata set.

Federal agencies are highly likely to want to understand how the material whose production they have funded is used. To this end, attention should be paid to ensuring that the material is stored in repositories that can provide usage data. There are a number of initiatives and standards being developed that will enable usage data to be aggregated across repositories and, hopefully, across publisher sites too, though the latter is dependent upon publisher cooperation: a promising start in this direction in the form of the PIRUS project has successfully proved a concept but the follow-up will focus only on repositories unfortunately. Nonetheless, it is useful to be able to measure usage across repositories and federal agencies will benefit from this.

In terms of academic impact (citations) some nascent services that will work on the Open Access corpus are in development, and the recently formal launch of Google Scholar Citations means that there is now at least one useful alternative to inaccessible (commercial) services that works across the whole scholarly literature.

There is no doubt that further developments will occur in the area of technical interoperability but this will always be work-in-progress: the best options available now, and there are a number of good options, should be exploited now to bring forth public access, while a watching brief is kept upon new developments over time.

¹⁷ <http://www.openarchives.org/pmh/>

(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?

The benefit to all stakeholders who can use research findings is maximised by immediate, full, Open Access, delivered through a well-designed system which adheres to standards on technical interoperability, making the provisions and the finding and using of Open Access content as simple as possible and barrier-free.

The burden can be minimised by simplicity of policy and process. Policies across federal agencies should be coordinated – ideally, copied, but allowing for some minor differences where appropriate and really necessary – so that grant-holders, their institutions, libraries and publishers do not have to cope with a plethora of variations.

(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?

Peer-reviewed journal articles are the primary target for public access policies: the material is supplied free of charge by the authors who have conducted the research using public funds. Conference proceedings are the main publication channel for certain disciplines, notably most fields of engineering, including computer science. Papers submitted for peer-reviewed conference proceedings are generally provided under the same conditions as those to journals – that is, they are provided free of charge by the authors. Where this is the case, and the work is publicly funded, the same conditions of access should pertain in policies.

Books are usually distinguished from the above because they are written with some expectation of royalty payment to the author. The public usefulness, and the fact that most books are written about research that is publicly funded, make this a more difficult case for policy development. At the moment, policy should encourage book content to be made accessible as soon as possible, and it should be noted that there is plenty of evidence now to indicate that such a process frequently drives up sales. Also, it looks likely that Open Access monograph publishing will continue to grow from its current small base to become a significant part of the book market, though business models that work for the long term have yet to be fully worked out. There are, however, some promising initiatives in this area and they signal better access to monograph content in the future¹⁸.

(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?

¹⁸ See, for example: <http://www.youtube.com/watch?v=niyYVWa2w6w>

No embargo at all is the desirable goal for research papers. We have provided evidence above on the effect of delays in accessing research findings and we see no compelling reason for enshrining such detrimental effects in policy. The argument for embargoes is made by publishers on the basis that publishers must have time to collect revenues from subscriptions, yet a majority of publishers permit public access through institutional repositories to the author's final version of a journal article immediately after peer-review, indicating that it is perfectly possible to continue in business while permitting this practice. Indeed, there is still no evidence to conflict with the testimonies¹⁹ provided by the American Physical Society and the Institute of Physics Publishing (UK) in 2005, where both stated that no subscriptions losses can be attributed to the self-archiving of papers in the high energy physics Open Access repository, arXiv, despite the full contents of many journals having been made available through this route by authors since 1991.

Moreover, if publishers still fear a detrimental effect from lack of embargoes, they have a further option: there is now plenty of evidence that publishers can make a sound and sustainable business from flipping their business model to collect revenue in the form of article-processing charges and publish Open Access journals (which do not have an embargo). From independent start-ups through to bold 'flips' from the subscription model, publishers have demonstrated that Open Access publishing is a viable alternative to the subscription sales model.

The public benefit of immediate access to research findings is demonstrably high. Policies that accommodate embargoes reduce that public benefit in favour of a financial benefit to private concerns, one that can be derived anyway by a change of business model, leading to an outcome where the needs of all parties are satisfied.

This submission

Submitted to the Office of Science & Technology Policy by the Board of *Enabling Open Scholarship*, January 2012:

Professor Bernard Rentier (chair), Rector of the University of Liège, Belgium

Professor Tom Cochrane, Deputy Vice Chancellor, Queensland University of Technology, Brisbane, Australia

Dr William Dar, Director General of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Professor Martin Hall, Vice Chancellor, University of Salford, UK

Professor Stevan Harnad, Canada Research Chair, Université du Québec à Montréal (UQAM), Montreal, Canada

Professor Keith Jeffery, Director of IT and International Strategy at the Science & Technology Facilities Council, Swindon, UK

Dr Sijbolt Noorda, President of VSNU, the Association of Dutch Research Universities, The Hague, Netherlands

Professor Stuart Shieber, James O. Welch, Jr. and Virginia B. Welch Professor of Computer Science, School of Engineering and Applied Sciences at Harvard University and Director of

¹⁹ <http://users.ecs.soton.ac.uk/harnad/Hypermail/Amsci/4336.html>

Harvard's Office of Scholarly Communication, Cambridge, USA

Professor Ian Simpson, Deputy Principal for Research and Knowledge Transfer, and Professor of Environmental Science, University of Stirling, UK

Professor Peter Suber, Berkman Center for Internet & Society, Harvard University, Cambridge, USA

Professor John Willinsky, Khosla Family Professor of Education at Stanford University and director of the Public Knowledge Project at the University of British Columbia and Simon Fraser University

Dr Alma Swan (Convenor/Coordinateur), Director of Key Perspectives Ltd, Truro, UK

Matt Cooper

gsc@unt.edu

President, University of North Texas Graduate Student Council

Denton, TX

The Graduate Student Council (GSC) at the University of North Texas advocates on behalf of 7,784 graduate students on federal, state, and local issues, and supports their development through professional training and networking. The GSC supports public access efforts as tools to increase our nation's competitiveness in the 21st century and ensure that knowledge freely flows to those who seek to use it--both within the academy and elsewhere.

Comment 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?

As the United States continues its shift toward a knowledge-based economy, making publicly funded research available to all graduate and professional students is a direct investment in America's future. Open-access to publically funded research will facilitate the development of new entrepreneurial ventures by bright, innovative, and talented new graduates. Many graduate and professional students are already funded by federal dollars either directly through federal loans, or indirectly through NSF, NIH, USDA or other federal agency grants. The skills and talents that graduate and professional students develop are reliant upon having access to the most recent and up-to-date knowledge generated in their field.

Not having access to the most up to date research means that federal investments are being allowed to dull. Expanding the public's access to cutting-edge research will help graduate and professional students to enter the workforce running, allowing them to continue to develop new innovations and industries while they are still students, and after leaving their institutions of instruction. It is today's graduate and professional students in the humanities, arts, biological/health sciences, social sciences, engineering and computer sciences that will develop and found the Fortune 500 companies of the next century. Open access can help these new job-creators and job-holders to get their ideas and companies into the marketplace.

For example, a working paper from researchers at Harvard Business School indicates that dissemination of problem information to external researchers can increase the rate of problem resolution by 29%--even at firms who traditionally deal in science-driven R&D processes (Karim R. Lakhani et al., "The Value of Openness in Scientific Problem Solving," <http://www.hbs.edu/research/pdf/07-050.pdf>)

Comment 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?

The United States public would be best served by making public access available for all published peer-reviewed works or presentations that were funded by federal dollars. Many are already available freely in University repositories, where faculty may be encouraged to store pre-published versions of their manuscripts, or on federal agency and personal websites. Developing an open-access policy could move these databases and archives to a more easily searchable and centralized location similar to the current PubMed and Google scholar databases.

Developing such a database without infringing on copyrights could best be accomplished by ensuring that federally funded researchers be required to publish their findings through appropriate Creative Commons CC-BY licenses. High impact journals will always have the need and desire to publish high quality articles and research in order to keep their journals relevant. Researchers and scientists will continue to maintain their need to publish in high impact journals in order to remain relevant in their fields and ensure their knowledge is widely disseminated. Requiring the use of Creative Common CC-BY licenses would allow publishers and scientists to continue to publish the highest quality articles in the highest impact academic journals while still allowing for appropriate and legal dissemination of these works.

Comment 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?

Long-term stewardship would be best guaranteed by the hosting of databases by government agencies. The NSF, NIH, USDA, and other government institutions are the most appropriate stewards to ensure that publicly funded articles are permanently preserved, and made both accessible and usable by the general public. The hosting of such articles in a centralized database would best enable innovative companies and individuals to develop new services and companies. In order to accomplish this wide availability - approved repositories that meet conditions for public accessibility, usage rights, interoperability, and long-term article preservation could be maintained by third-parties and innovative public/private partnerships.

Comment 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?

No comment at this time.

Comment 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?

No comment at this time.

Comment 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?

Peer-reviewed conference papers and proceedings represent a significant portion of published literature and information that is relevant to their respective fields. Often, these papers and proceedings will contain additional or unique information on research that is otherwise unpublished. Due to both their impact and contribution, these types of publications should be included in the same category as journal articles with respect to public accessibility. Additionally, conference proceedings and papers can often represent first step towards journal publications, whose purpose keep the field's community of authors and researcher up-to-date on both current trends and current work being done in the field. For this reason, public access to these types of publications is important, allowing readers to remain apprised to both current research and current trends. Certain conference proceedings and papers may contain comprehensive reviews of published research to date, keeping both old and new authors informed on a comprehensive outlook in a particular field of interest. For this reason, public access remains important as it enables a wider audience of readers to both understand and perhaps enter a field of research. Conference proceedings and papers allow authors to share their research with the broader community as it progresses. Public access to such publications will enable others to keep up-to-date with current and future trends on specific subject, enabling a fast dissemination of knowledge throughout the research process. Book chapters that are derived in part from publicly funded research may represent a separate category from journal articles or conference proceedings, due not only to their publication medium, but the content contained within the chapters that may not derive from federal funding.

Comment 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?

An embargo period no longer than 6-12 months.

**Response to Request for Information: "Public Access to Digital Data Resulting from Federally Funded Research," November 2011
January 12, 2012**

Wendy Pradt Lougee
University Librarian
McKnight Presidential Professor
University of Minnesota Libraries
wlougee@umn.edu

Thank you for the opportunity to comment on "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research." These comments are submitted on behalf of the University of Minnesota Libraries. The University of Minnesota is one of the leading public research institutions in the United States, and a key contributor to the entrepreneurial economy of the state of Minnesota, as well as to scholarship both nationally and internationally. We strongly advocate for a policy requiring full public access to all publications resulting from federally-funded research as soon as possible after publication. We believe that such a policy would provide immeasurable public benefits far outweighing any costs or burdens such a policy might impose.

(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 publicly 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?

1.1 Costs and Benefits: The Challenges We Face Today

Simply put, the present environment of limited access to most research publications imposes many costs across many sectors of the U.S. economy. Increasing access to research publications may impose some burdens on publishers -- although it is by no means certain that a well-crafted open access policy will in fact damage publishers' bottom lines -- but it will unquestionably benefit many others. Here are a few concrete examples from the state of Minnesota of the challenges scientists, industry professionals, and members of the public currently face in accessing publicly-funded research:

- The University of Minnesota Chemical Engineering and Materials Science department runs a fellowship program in partnership with technology businesses. Enterprise partners sponsor fellowships for industry professionals (often scientists from their own research divisions) to collaborate with students and researchers at the University of Minnesota. Research fellows in this program perennially inquire about access to SciFinder, one of the most important subscription information resources in chemistry.

Unfortunately, SciFinder is one of the very few licensed resources for which even on-campus access is limited to University personnel by the terms imposed by its publisher, the American Chemical Society, and we are not able to provide access under any circumstances to these unaffiliated individuals. Since a large portion of the research available in SciFinder is the result of federal funding, a policy *increasing access to federally funded research would demonstrably improve the resources available to these researchers.*

- The University of Minnesota Extension program has a history of over 100 years of bringing the scientific knowledge and expertise of the University out into our communities, in partnership with federal, state, and county governments. The Extension program is vital to the health of individuals and communities, both urban and rural, across the state. It is also a key contributor to the success of the agricultural, environmental, and tourism industries throughout Minnesota.

Extension educators already make comprehensive use of education and information resources produced by federal agencies, knowing that these resources are free for all to use. However, Extension educators frequently wish to share research publications with their service communities. The growth of PubMed Central as a result of the NIH open access policy has been a great boon for Extension public health educators, but Extension faculty and staff in agriculture and environmental sciences are often limited to sharing abstracts or rough summaries of research data with their service communities.

Increased access to federally-funded research would allow Extension educators to get research into the hands of individuals who can put that research directly into practice in Minnesota's communities, and in some of Minnesota's most important industries. Moreover, many of the individuals, non-profits, and government programs with which Extension educators work have limited access to the Internet due to limited financial resources or remote rural locations, so re-use rights such as the ability for Extension educators to photocopy and distribute publications, or compile them into educational materials, would be immeasurably helpful to truly getting the research into the hands of practitioners.

1.2 Improving the Productivity of the Scientific Enterprise – Citation Impact

Scientists and scholars measure the productivity of the research enterprise primarily in terms of the spread of knowledge and the impact of their own research among their peers. Some of the best and brightest agree that “[b]road dissemination of research results is fundamental to the advancement of knowledge.”¹ Less altruistically, faculty across all disciplines report availability to peers in their disciplines as the most compelling factor in their choices of publication venues.² One of the best ways to measure whether research is available to other scholars and scientists is to track citations of publications – and numerous studies have documented that making works openly available increases the numbers of citations to each work. Wagner’s annotated

¹ An Open Letter to the U.S. Congress Signed by 41 Nobel Prize Winners. (2009, November 6). Retrieved January 8, 2012, from http://www.taxpayeraccess.org/supporters/scientists/nobelists_2009.shtml

² Schonfeld, R. C., & Housewright, R. (2010). *Faculty Survey 2009: Key Strategic Insights for Libraries, Publishers, and Societies*. Ithaka S+R. Retrieved from <http://www.ithaka.org/ithaka-s-r/research/faculty-surveys-2000-2009/Faculty%20Study%202009.pdf>

bibliography shows about 39 articles demonstrating an open access citation advantage (OACA).³ Another study (Gagouri, et al) responding to suggestions that OACA is simply a product of selection bias (i.e., that scholars only make works open if they are particularly likely to be cited), recently argued that there is a bias toward high-quality work in open access, but also noted that increased citation is in fact an independent phenomenon and real benefit of open access.⁴

1.3 Improving the Productivity of the Scientific Enterprise – Economic Impact

Another way to measure the productivity of the scientific enterprise is to consider the economic impact of scientific research. Commercialization is one valuable way to realize economic benefits from publicly funded research. However, researchers in intellectual property policy applaud the value of private research, but point out that the economic value of research cannot be measured solely in terms of commercial exploitation: the unquestionably hugely valuable Human Genome Project would have provided far fewer scientific *and* commercial benefits in private hands.⁵ Other researchers have demonstrated that opening access (i.e., limiting IP restrictions) to patentable products of bioengineering research both increased the volume of follow-on research and increased the diversity of uses to which the original advances were put.⁶ Increasing open access to research may provide a wide range of economic benefits. Economists John Houghton & Peter Sheehan suggest several specific areas in which the economic impact of open access to research might be felt,⁷ including:

- Speeding up research through faster access, potentially increasing return on both private and public investment in research.
- Reducing redundancy and duplicative efforts through wider access.
- Improving collaboration across disciplines and institutions through wider access, and potentially increasing the ability to recognize commercial applications.
- Reducing costs of education, producing a better future research workforce.
- Increasing access to individuals in health care, education, and smaller industrial enterprises, hence improving their productivity and service levels.
- Possible new industries developing around openly available content.
- Producing better informed citizens and consumers who can make more socially beneficial choices about their lives and the services and products they consume. (We would add that these better-

³ Wagner, A. B. (2010). *Open Access Citation Advantage: An Annotated Bibliography*. Issues in Science and Technology Librarianship. doi:10.3998/3336451.0009.202

⁴ Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). *Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research*. PLoS ONE, 5(10), e13636. doi:10.1371/journal.pone.0013636

⁵ Eisenberg, R. S., & Nelson, R. R. (2002). *Public vs. Proprietary Science: A Fruitful Tension?* Academic Medicine, 77(12), 1392-1399.

⁶ Murray, F., Aghion, P., Dewatripont, M., Kolev, J., & Stern, S. (n.d.). *Of mice and academics: Examining the effect of openness on innovation*. NBER Working Paper Series, (14819). Retrieved from <http://nrs.harvard.edu/urn-3:HUL.InstRepos:4554220>

⁷ Houghton, J., & Sheehan, P. (2006). *The Economic Impact of Enhanced Access to Research Findings*. Center for Strategic Economic Studies Working Paper Series, (23). Retrieved from www.cfses.com/documents/wp23.pdf

informed citizens may themselves contribute directly to research gains, in the form of the already-growing participation in “citizen science” efforts such as Stardust@home and Zooniverse.)⁸

Houghton’s economic research carefully models how increased accessibility and efficiency of research might affect the return on R&D investment in many different countries. In a 2006 article, his most conservative models predicted a \$1.5 billion annual gain in a move to open access; the middle-of-the-road models predicted annual gains of over \$16 billion.⁹ More recently, Houghton assessed the costs and benefits of the proposed FRPAA legislation within the U.S. (benefits approximately 5x costs) and overall (benefits approximately 8x costs.)¹⁰

The University of Minnesota research enterprise is currently estimated to contribute about \$1.5 billion to the Minnesota economy each year. University of Minnesota alumni have founded at least 10,000 businesses in the state.¹¹ Imagine how much greater that contribution could be if more of the research produced by the University was made available to the public. Although full access to publications resulting from federally-funded research may impose some costs on publishers, it would not significantly affect the publication process or impose new burdens on researchers, and the examples above demonstrate how such access could directly improve the work of innovators and industry professionals. Full access and re-use rights could also enable development of innovative tools by programmers in both commercial and open-source environments enabling new forms of search, analysis, and connectivity for published research.

(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?

2.1 The Intellectual Property Interests of Stakeholders in the Research Process

All of the named stakeholders have an interest in *making use* of the results of research. However, to the extent that publishers have an intellectual property ownership interest in the published results of research, it is coterminous with the copyright interests of researchers, and is usually acquired from them with no remuneration, and in fact sometimes at a cost to the researchers. Researchers *produce* intellectual property as a direct result of federal research funding; publishers *acquire* a copyright interest in research products after they have been created. It should be noted that publishers do contribute valuable editorial, promotional, and other functions to the publishing process, but only require distribution rights to fulfill their role.

⁸ Stardust@home <http://stardustathome.ssl.berkeley.edu>; Zooniverse <https://www.zooniverse.org>

⁹ Houghton, J., & Sheehan, P. (2006). *The Economic Impact of Enhanced Access to Research Findings*. Center for Strategic Economic Studies Working Paper Series, (23). Retrieved from www.cfses.com/documents/wp23.pdf

¹⁰ Houghton, J.W., Rasmussen, B. and Sheehan, P.J. (2010) *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*, Report to SPARC by Victoria University's Centre for Strategic Economic Studies. Retrieved from <http://www.cfses.com/FRPAA>

¹¹ Tripp Umbach. (2011). *The Economic and Societal Impact of the University of Minnesota*. Retrieved from http://impact.umn.edu/assets/pdf/Final_Report.pdf

Policy choices for federally-funded research publications can have little effect on the intellectual property rights of most participants in the process of scientific research, because no open access policy will change established intellectual property laws. No one has documented any increased risks of infringement on the intellectual property rights of any stakeholder under existing open access policies. Existing policies have admittedly already changed *practices* surrounding researcher management of their rights and the acquisition of rights by publishers from scholarly authors – but these changes have meshed quite well with the intellectual property interests and practices of authors and most other participants in the process.

2.2 Specific Intellectual Property Interests – Patent and Copyright

Patent rights are only indirectly implicated in the publishing process, as researchers generally do not publish or publicly comment on potentially-patentable innovations until the patent application process is well underway. A policy that requires open access to published research does not threaten the patent rights of researchers, funders, or supporting institutions - or businesses that build on these efforts - since the appropriate rights-management processes are already established, and equally applicable to all research publication, in any medium or any access mode. Upcoming changes under the America Invents Act will only reduce the effect that publications can have on the patentability of research; under the new regulations, patents will be awarded to the first party to file an application, regardless of the date of invention. Publication before filing may still create “prior art” that can undermine patentability, but these risks will be unchanged from the current system, and are well-managed by researchers.

Copyright rights, on the other hand, are directly implicated throughout the publishing process. However, no particular approach to distribution inherently affects the copyrights in research publications, since copyrights cannot be transferred except via a formal licensing agreement or written transfer. The copyright status of an article is the same whether it is published on paper, in a limited-access online service, or made freely available online to all. A policy which requires research publications to be made freely *available* to all does not affect the copyright in those publications. Open online distribution does increase the visibility of research, which can sometimes lead to increased opportunities for infringement. However, there is no evidence that open publications are more frequently infringed than limited-access publications – in fact, since openness reduces barriers to legitimate access, open publications may be less likely to be copied by questionable or illegitimate means. Some may fear that broader access will lead to increased copyright infringement, but there is no evidence of this with current open access publications – and where there are fewer limitations on access and use, there are fewer opportunities for infringement.

Enabling full *access* to research publications still falls short of enabling a number of uses that could be highly beneficial to scholars, industry, and members of the public. A work that is publicly accessible is still subject to all the limitations of copyright, which may present barriers to many productive uses. A policy enabling wide public *re-uses of* (rather than simply access to) publications would create additional value. Teachers would be certain they could reproduce the articles for their students. Scholars could reproduce the text for new and emerging forms of computational analysis. Entrepreneurs and developers could build new tools, services, and device applications related to these publications without worrying about reproduction or derivative work

rights. Existing open licensing tools such as Creative Commons licenses (specifically, a Creative Commons Attribution CC-BY license) would enable all of these uses, while ensuring full credit to researchers.

Unquestionably, such a policy would have a direct impact on the copyright in those publications. Wherever the copyright in the publications may lie (with the researcher, the publisher, or shared among multiple parties), the copyright holder(s) would necessarily have to cede some control under an open re-use policy. However, it is worth noting that *researchers usually do not currently control or receive remuneration* for any of these uses, and Federal agencies, research institutions, and other stakeholders currently usually must *pay for these uses*. While a policy requiring wide public usability for published research might require cession of some copyrights currently controlled by publishers, it would not materially change the rights that researchers currently control. It would also provide dramatically increased usability for many stakeholders in the system of scientific innovation.

(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?

Centralized repositories bring great benefits, such as providing an authoritative copy of a publication (and if necessary, including corrections or retractions), maintaining a single access point for direct searches across large bodies of research, and establishing interoperable access mechanisms. Centralized hosting is the best opportunity to maintain the integrity of the published research, and maintain public access over time. Centralized hosting will also result in greater innovation around research content. Decentralized access necessitates the development of tools to search across multiple repositories, which may sometimes result in beneficial innovations in search functionality, but usually simply results in less-than-optimal search experiences. Centralized access with open infrastructures, on the other hand, introduces numerous efficiencies, which can enable development of innovative third party search, analysis, and other tools. PubMed Central, the central repository under the NIH open access plan, has proven invaluable in improving access to federally-funded health research. Searchers know where to go, and know that the copy they are accessing is the copy of record.

However, we do support a managed strategy for redundant copies to ensure long-term access to authentic works. These copies could be stored in library and institutional archives, in subject repositories, and with publishers. A policy that enables extensive re-use by both commercial and non-profit users (such as under a Creative Commons Attribution CC BY license) would remove barriers to decentralized storage, and would maximize experimentation and innovation with published resources, regardless of their home.

We do not advocate for a policy that leaves the provision of enduring access solely in the hands of commercial publishers; ideal primary deposit is in repositories hosted by the government and/or non-profit institutions and organizations. The goals of most corporate publishers are to maximize profits for their shareholders – this is absolutely appropriate, but in pursuit of those goals archives often change hands. These are not ideal

conditions for preserving access to published research. Any library staff member who works with licensed electronic resources has numerous stories to tell about access lost – sometimes temporarily, but sometimes for long periods - when one publisher was acquired by another. The worst cases include permanent loss of data and are not easily repaired. It is not unheard of for a publisher to contact a library in search of back print copies of a journal to re-digitize to replace a lost (or never-created) archive. The federal government or universities, by contrast to publishers, are long-lived institutions, and the public-oriented goals of federal research funding are more stable over time.

(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?

Libraries, museums, and other cultural institutions throughout history are good examples of public-private partnerships that ensure access to and long-term stewardship of the items in our collections. Traditionally, libraries have acquired books by purchase and both preserved them and made them accessible to the public over time. As acquisition of materials for library collections has increasingly shifted to a licensed-access model, it has been increasingly difficult for libraries to leverage our expertise and experience in providing access and preservation. The terms of many licenses preclude libraries from storing and providing access to the materials, leaving the task of providing long-term preservation and access in the hands of publishers.

Some publishers have developed innovative tools for accessing published research, but they have been less successful at making these tools, or their separate article archives, interoperable. It is still a difficult programming task to build a tool that searches across multiple commercial databases because each is formatted differently. By contrast, the systems that libraries and other non-profit entities have built to host content generally have robust systems for access by outside programmers, and follow strong standards to ensure interoperability. The HathiTrust Digital Library, for example, has robust APIs that allow export and interoperability of all HathiTrust data, including the full text of public domain works. Using the API, many public domain works are now available for purchase as physical copies via Amazon.com. Unfortunately, due to rights limitations, libraries are often unable to apply these robust tools to the most current content. Third-party applications and innovations *can* be built via collaboration with publishers – much of the discoverability data in Google Scholar, for example, is provided directly by publishers. But many entrepreneurial developers need to engage in experimentation or proof-of-concept testing, and may not have the financial or social capital to negotiate with publishers; even Google Scholar was initially developed using only publicly available discoverability data. Policies that require enforceable and robust open standards for storage and access would allow improved collaboration between publishers and libraries and other cultural institutions. Such policies would also enable innovation by third-party actors large and small.

We believe that permanent storage in a public archive compliant with repository standards such as OAIS and the emerging ISO/DIS Standard 16363 for Trusted Digital Repositories is a necessary part of long-term stewardship of published research, and that partnerships between publishers, libraries, higher education institutions, and government agencies can most robustly support the long-term preservation and access to federally-funded research publications.

(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?

The best way to encourage interoperable search, discovery, and analysis is by using standardized metadata schema appropriate to the materials in question. There are many relevant existing standards to draw from: the protocols of the Open Archives Initiatives represent a reasonable minimum metadata standard across disciplines, and the National Information Standards Organization promulgates many relevant metadata and other standards. Community-based expertise should be used to develop standards and conventions for data structure and metadata management specific to a discipline's research output. Repositories should also be encouraged to explore further the use of semantic web technologies (RDF and URL-identified entity and relationship vocabularies) and linked data to leverage discovery. Emerging metadata standards will provide important improvements to access, interoperability, and use. For example, ORCID is developing a new approach to uniquely identifying researchers, and can not only enable improved discoverability and access to a researcher's publication output, but also provide improved function to institutional grant-monitoring systems and to funder review of output.

It is increasingly recognized in many scholarly communities that published research and the underlying research data on which the publication is based can and should be associated, and that scholars in the future may interact with published articles and associated research data. Thus, it is important to develop and maintain metadata specifications that are unified for both publications and research data, and recognize the relationships between these materials.

Open means of data exchange, such as APIs, are also essential to realizing the full potential of research repositories. Standard and open data exchange allows for greater interoperability, and also enables development of new resources, tools and applications built on repository contents. The EthicShare project at the University of Minnesota harvests citation data from various repositories and web resources, resolves to relevant licenses for an individual user, creating a robust discovery and collaboration environment for this interdisciplinary field. Currently, the project makes extremely productive use of information from PubMed, OAlster, and other open bibliographic resources, but is somewhat limited by the lack of public APIs for publication archives in related fields. An API requirement could facilitate the development of this project and many other non-profit and commercial tools and applications. .

(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?

One way that Federal funding agencies can maximize the benefit of public access while minimizing burdens and costs, is to standardize policy requirements, so that compliance can be streamlined across disciplines and institutions. Another way to minimize costs is to build on expertise that already exists in communities and organizations experienced with providing public access to published works. Established archives such as

PubMed Central, arXiv.org, and HathiTrust can be looked to as models for governance, infrastructure, and standards.

Tools that automate the process of depositing and distributing published research already exist and can be integrated into the workflow of authors and publishers. Excitingly, SWORD is currently exploring how best to enable deposit of research data as well as research articles.¹² Similarly, integration with grants management and researcher profile tools already in use at many institutions would ease burdens on grants-receiving institutions, while enabling greater transparency and accountability for federal research funds and improved grant reporting to funding agencies.

(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?

Publications resulting from federally-funded research that do not take the form of scholarly journal articles often convey equally valuable information and knowledge. All researchers should be encouraged to make the results of their research freely available as soon as possible after publication in order to realize the greatest possible public benefits. However, the processes for distributing non-article publications differ in fairly significant ways from the processes for distributing scholarly journal articles, and are even less uniform across disciplines. For these reasons, the full implications of a public access policy for these kinds of peer-reviewed publications are as yet unclear. Conference proceedings do usually resemble journal publications, but often present research at a more nascent stage of development than articles. Many conference proceedings are already made publicly available, so it is clear conference proceedings are not inherently unsuitable for open access, but more flexible provisions, perhaps based on date of final publication rather than date of presentation, might need to be adopted.

Monographs have a slower publication cycle, and sometimes remaining commercially saleable for several years. Rights ownership and remuneration are also often dealt with differently in monograph publications, and individual authors may have direct economic interests in their own monograph publications. There have been few systematic efforts to ensure public access to the full text of newly-published scholarly books, so the economic implications of doing so are not well understood. However, several authors, such as Harry Lewis, James Boyle, and Ted Striphas have made their monographs available simultaneously in commercial print publications and via free electronic copies with good success. Studies of scholarly monographs released under similar hybrid commercial print/free electronic distribution models have not conclusively shown any consistent harm to sales, and in several cases it appears that the free electronic copies have in fact driven sales of print copies.¹³

Of course, the benefits of providing public access to research are not limited to direct profits. The National Academies Press has been experimenting with providing free public access to electronic versions of its publications for years, and this past summer announced that from now on, all books published by NAP will be available as free PDFs. Their intention of this program is to widen the distribution and increase the impact of

¹²

¹³ Hilton, J., III, & Wiley, D. (2010). The Short-Term Influence of Free Digital Versions of Books on Print Sales. *Journal of Electronic Publishing*, 13(1). Retrieved from <http://dx.doi.org/10.3998/3336451.0013.101>

NAP-published research, with an avowed goal of increasing downloads from 700,000 per year to over 3 million in 2013.¹⁴ For this non-profit publisher, the increased access to and use of the materials they publish is a clearly beneficial effect.

(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?

Scientific research is most valuable when it is available for other scientists, innovators, entrepreneurs and businesses to learn from and build upon in a timely manner. The sooner publications are made publicly accessible, the sooner the information therein can be put to use. We advocate for immediate public access, but in no case advocate for a general embargo period longer than twelve months from publication.

The NIH public access policy allows for a twelve-month embargo period, and has not been shown to have significant detrimental effects for publishers in the field. Libraries, institutions, and organizations with a timely need for these publications still pay for early access – but organizations whose budgets cannot support paid access are still able to access the research. Many publishers have adopted shorter embargo terms (the *New England Journal of Medicine* and many other biomedical publications make all their contents freely available after six months¹⁵) without apparent harm. Certainly shorter embargo terms would be a good thing for researchers hoping to increase the reach and impact of their research, for individual taxpayers researching a health condition affecting their families, and for workers and researchers at institutions who cannot subscribe.

It is possible, though not proven, that embargo terms shorter than a year could have an impact on publisher profits. However, publishers have not demonstrated any harm from current embargo periods of twelve months or less, and many publishers have voluntarily adopted shorter embargo periods. Moreover, publishers are but one of many stakeholders in the systems of scholarship and scientific exploration. The impact of embargo terms cannot be measured solely by hypothetical damage to publisher revenues, but also by the value that free access creates for other stakeholders.

¹⁴ National Academies Press. (June 2, 2011). National Academies Press Makes All PDF Books Free to Download. Retrieved from <http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=06022011>

¹⁵ *About NEJM Past and Present*. <http://www.nejm.org/page/about-nejm/history-and-mission>



January 12, 2012

The Honorable John P. Holdren
Assistant to the President for Science and Technology and
Director, Office of Science and Technology
New Executive Office Building
725 – 17th Street, NW
Washington, DC 20502

Comments in response to Office of Science and Technology Policy Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research
Federal Register Doc No 2011-28623
<http://www.gpo.gov/fdsys/pkg/FR-2011-11-04/html/2011-28623.htm>

Dear Dr. Holdren:

Northwestern University is a private research institution with 16,377 students and approximately 3,000 full time faculty. In 2010-11, Northwestern researchers attracted total awards and grants of approximately \$511.7 million. Northwestern's libraries hold more than 5 million volumes, 4.6 million microforms, and provide access to 110,341 current periodicals and serials. In addition, the library system boasts more than 700 databases and 6,000 electronic journals. 56% of the libraries' \$14 million collection budget is devoted to these e-resources.

Northwestern is recognized both nationally and internationally for the quality of its educational programs at all levels. *U.S. News & World Report* consistently ranks the University's undergraduate programs among the best in the country.

Among graduate programs, the Kellogg School of Management regularly ranks among the top five business schools in the country for both its traditional curriculum and its executive master's program. *U.S. News & World Report* rankings placed Northwestern's School of Law 11th, and the Feinberg School of Medicine in the top 20.

(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 publicly 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?

Making peer-reviewed scientific publications freely available after publication, with minimal restrictions on use, will accelerate scientific discovery and expand opportunities for entrepreneurs to develop new services and products. Lowering or removing barriers to access to new research results

will increase opportunities to identify new partnerships with industry, complementing the goals of university patent and technology transfer processes, and the goals of federal programs like the Small Business Innovations Research/Small Business Technology Transfer (SBIR/STTR) <http://www.sbir.gov/> and the recently announced initiative to speed commercialization of university research (National Advisory Council on Innovation and Entrepreneurship, Department of Commerce, 2011). Closer academia-industry links and shorter cycles between research, dissemination of results, and commercialization accelerate the public's return on its investment, creating new markets, new jobs, and new tax revenue for local, state and federal governments.

It has been conservatively estimated that expanding an NIH-type post-publication open access policy to other federally funded research will result in improvements in research efficiency and accessibility, and yield for the American taxpayer a return approximately 8 times larger than the initial research investment (Houghton, 2010, p. 8). These projections align (and yet, they pale in comparison) with the measurable economic impacts of the Human Genome Project. That massive public project, whose results were made immediately available for both public and commercial use, yielded a return on investment of approximately \$141 per \$1 of public funding (Battelle Technology Partnership Practice, 2011, p. 6). In contrast, it is estimated that the IP restrictions temporarily placed on genes sequenced by Celera in its competing project have had a lasting negative impact on subsequent research and innovation. Genes first sequenced by Celera have fewer scientific publications and are less likely to be used in genetic tests (Williams, 2010, p. 2).

Likewise, providing immediate free access to research articles removes barriers for researchers, who can more quickly and effectively incorporate up-to-the moment findings into new research, accelerating scientific productivity. Even in university environments, researchers still report some difficulty gaining access to all of the scholarly material they need to conduct research, and these effects will be more severe for smaller businesses and worse yet for the general public. Open access publications, available through models ranging from fully open access journals to self-archived publications in university, disciplinary or funder repositories like PubMed Central, are downloaded more and cited more frequently than publications for which a subscription is required. Citation rates are significantly higher for immediate open access articles even when controlling for factors such as mandated vs. self-selective archiving, journal impact factor, and number of references cited (Gargouri et al., 2010, p. 8). Some studies have shown increased citation rates as high as 600% for open access publications (Swan, 2010, p. 17), though this varies significantly by discipline, and ranges from 40% to 90% are more common. Across social science, science and humanities disciplines, providing open access to published literature, particularly if the access is granted immediately after publication, will increase the impact of research. Most importantly, respected open access initiatives have succeeded in providing this broad access while maintaining and sustaining a robust peer-review process and continuing to provide many valuable services such as editorial enhancement, error checking, citation mining, and indexing and linking services.

Large databases of freely accessible scientific literature can also spur development of new knowledge exploration tools that aid researchers facing the daunting task of finding relevant publications amongst the hundreds of thousands of new articles published each year. Software like IN-SPIRE™ <http://in-spire.pnnl.gov/> and the Action Science Explorer (Ferrante & Zgorski, 2011) and projects such as the Large Knowledge Collider (LarKC) <http://www.larkc.eu/> give scientists powerful new tools for finding connections between previously unconnected research, using machine learning, automated reasoning, and network science to make new inferences and suggest new pathways for research. Tools such as BioXM(Maier et al., 2011) combine assertions drawn from published

literature with data about genes and other objects to yield new insights. These powerful computational tools depend on access to both metadata and full text for published articles, and constructing the new data sets and indexes on which they operate requires that the articles be free of downstream use restrictions, including prohibitions against commercial use.

Another example of machine-aided exploration may be found in the small but vibrant community developing around research networking (RN) tools. Both open source (VIVO <http://vivoweb.org/>, Harvard Profiles <http://profiles.catalyst.harvard.edu/>) and commercial tools (SciVal Experts <http://www.info.scival.com/experts>, Thomson Reuters InCites <http://researchanalytics.thomsonreuters.com/incites/>) demonstrate the power of constructing author and concept network visualizations atop metadata and full text of research publications. These tools give universities and funders more accurate pictures of research output and ease the burden of publication tracking and reporting, but can also facilitate new collaborations and suggest new directions for exploration. However, RN tools will be limited by the quality and breadth of their inputs. In implementing a research networking tool at Northwestern, we have found that commercial database providers can be reluctant to make metadata or full text available for these non-consumptive uses, particularly if a commercial competitor developed the RN tool. The promise of research networking tools and other machine-aided inference systems will be severely constrained without access to large, freely reusable collections of research publications.

The National Institutes of Health (NIH) experience implementing a public-access policy and a large central database of results clearly show that this is a cost-effective approach to supporting open access to research. The article system's annual maintenance costs are approximately \$3.5 - \$4 million dollars, or roughly 1/100th of 1% of the NIH's \$30 billion budget (Lipman, 2010).

(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?

To gain the greatest societal benefit from publicly funded research, preserve the ability to redistribute publications for teaching and other purposes, prepare derivative works, and promote development of new discovery tools and products, a public access policy should seek to make publications as broadly usable as possible, as close as possible to publication time, with few to no restrictions on reuse. Licenses or policies that permit publishers to restrict uses of open access copies to single reader use only must be avoided, or at the very least a phased approach considered that will permit some restrictions on use during an embargo period, but release the works for full reuse afterwards. Any open access, whether green, gold, gratis or libre (Suber, 2008), is better than none, but the Creative Commons CC-BY Attribution license is most conducive to use by readers and machine reading systems.

Policies that permit publishers to compel authors to sign over their copyrights must also be avoided. Nonexclusive licenses to publishers should become the norm, rather than a surrender of the author's copyright. There is sufficient leeway in composition of such licenses to permit publishers to recoup costs associated with provision of publication services without restricting self-archiving or productive

and socially beneficial derivative uses of scientific publications. Some publishers have developed paid options (“SHERPA/RoMEO - Publishers with Paid Options for Open Access,” n.d.) for selective open access, often a hybrid model that mixes free and paid access content in the same publisher-hosted journal site. In some cases, as with the recently revised Taylor & Francis iOpenAccess service (“Taylor & Francis Author Services - iOpenAccess & NIH policy,” n.d.), articles are portable, and may be posted to any institutional or disciplinary repository, but carry with them additional terms and conditions to prohibit, as in the Taylor & Francis example, certain uses including commercial uses. It is understandable that publishers are leery of repackaging and reselling, but blanket prohibitions on commercial use, particularly when authors have paid several thousand dollars—as high as \$3000 per article in the case of T&F—for open access, may be unnecessarily restrictive, and cripples innovative, value-added and highly productive uses as well as simple reselling.

(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 federal government has already demonstrated through the NCBI systems, including PubMed Central, that it is capable of efficiently mounting a large scale, trustworthy and robust repository for both publications and data. Furthermore, it is appropriate for the U.S. government, as a major funder for scientific research, to also accept the responsibility for permanent stewardship of these important assets, to preserve them, and to continue to provide broad public access. Centralizing management of publications achieves economies of scale and eliminates the need for federated search tools, metadata or full text harvesting services, and other linking or mirroring systems to tie distributed archives together. Consistency and uniformity for publishers and authors will be the result. A disadvantage of a centralized approach may be that it minimizes the role of disciplinary and institutional repositories, and reduces capacity to provide specialized services and description tailored to the data and publication types specific to certain domains. A decentralized approach may also facilitate better access to research that is not funded by the U.S. federal government, but is available on an open access basis.

(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 research networking and knowledge extraction tools discussed above present very compelling cases for the potential of effective public-private partnerships. Many commercial database providers contract with open access publishers to include their full text in added value products that greatly enhance university researchers’ information discovery experiences. These offerings can and should continue to flourish in an open access environment, and can provide publishers and scholarly societies with additional revenue streams, greatly expanding the number of resources they can index, mine, and provide access to, and significantly enhancing their value. Publishers could also act as contracted service providers to provide open access repository services, provided they are able to meet conditions for trustworthiness, accessibility, reuse, and openness.

However, better models for partnership may exist between funders and universities, particularly with libraries that have amassed significant experience with digital repositories over the past decade. Northwestern has an internal digital repository system based on the Fedora Commons software, but is also a founding partner in the HathiTrust shared digital repository system. HathiTrust has satisfied a Trustworthy Repository Audit and Certification (TRAC) assessment and currently houses some 10 million digitized volumes. The partnership plans to expand support for other content types, and to pilot digital publishing services through the HTPub project <http://www.hathitrust.org/htpub>. This development and others like it, such as the California Digital Library's Merritt repository and eScholarship system, could dovetail with plans to expand federal open access requirements and accelerate scientific publication archiving programs. Should the U.S. government decide not to expand with NCBI-like central repositories, a promising model is partnerships with large university digital repositories or large multi-institutional repositories such as HathiTrust. Likewise, disciplinary repositories such as arXiv and the Social Science Research Network (SSRN) have succeeded in developing scalable, reliable solutions to open access archiving and could be logical partners in a distributed or shared/mirrored archive model. A network of distributed repositories, like the European DRIVER project <http://www.driver-repository.eu/>, would build on existing investments and disciplinary customizations.

(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?

The current NCBI databases and the NLM DTDs are an admirable model and an excellent starting point for an expanded network of open access repositories. Standardization of the form and content of metadata and full text (JATS XML, for example) will be critical to successful interchange, will promote use by machine readers, and can facilitate automated deposit and other unmediated or minimally mediated activities. Existing metadata standards such as Dublin Core and machine exchange via APIs or OAI-PMH are well established and accepted components in data linking and exchange. Custom metadata schema for most domains can be crosswalked to Dublin Core in the absence of common element sets for cross-searching, which is easily extended through application profiles or qualifiers. In addition to predictable forms for descriptive metadata, standard approaches must be devised to express rights and provenance beyond authorship (version, lifecycle events, etc.) in a machine-readable and machine-interoperable manner. PREMIS is an accepted standard for provenance and rights metadata in the library domain, and may be suited for extension for these purposes. The SWAN provenance, authoring and versioning ontology specification may also be a useful model. Emerging standards such as ORCID, for disambiguating author identities, and I-2 for consistent identification of institutions will also be important and must be supported, and collaborations with NISO, the Library of Congress and other groups involved in standards development and maintenance will be invaluable to consensus building.

The systems must fully enter the linked open data ecosystem, and must be capable of supporting semantic description and enhancement, either natively in the database or by exposing sub-article information through durable URIs so that inferences drawn from published research can be banked separately and linked to the evidence in the underlying articles. Ideally, funder databases can be

expanded to include direct storage of RDF-based statements formally asserting relationships between concepts or objects. Robust support for RDF and concept linking can enable formalization of statements, sometimes referred to as ‘nanopublications,’ as recognizable contributions to the scientific discourse (Mons & Velterop, 2009).

(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?

Adopting uniform requirements across all funding agencies will greatly simplify the burden on universities and their researchers. Scientists are likely to have grants from multiple agencies, and a single set of deposit requirements reduces complexity and simplifies compliance, reporting and monitoring.

(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?

Expanding public access policies to include other products of federally funded research may help to meet the goals of broad and timely sharing of results and address the lag between discovery and formal publication. Educational materials, including conference proceedings, technical reports, books and book chapters, should all be eligible for coverage under a policy, but the specifics of the policies may differ from those developed for journal articles. Conference presentations and technical reports may lie on one end of a spectrum, where rapid deposit is a reasonable expectation, but the economics of book publishing are more complex, and longer delays may be necessary to recoup author payments or other publication costs. Access and reproduction could also be significantly more complicated and cumbersome with books or proceedings where it could often be the case that some chapters or sections are Federally funded but not all. The full book or proceeding might not be able to be distributed as an integral product with consistent pricing or rights management. Public access policies should evolve in keeping with the norms and practices of academic disciplines and scholarly societies, preservation of high quality peer review, and the types and forms of publications natural to academic discourse. There may be opportunities to encourage sharing of other types of research results, e.g. negative results, which can also increase research efficiency. In all cases, policies should apply to results that the investigators have decided to disclose through publication, presentation or other means, thus avoiding potential conflicts with technology transfer processes as well as risks to national security or patient privacy. A public access policy should not develop new, alternative forms of publication, such as final project reports, as a substitute for the forms of scholarly communication that already exist and that serve the goals of research dissemination.

(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?

We are not aware of any data or studies showing that the NIH-permitted embargo period of 0-12 months has harmed publishers in any way, and ample evidence exists that the sooner an article is open access, the greater its research impact (Swan, 2010). Publishers of all kinds have been able to sustain the high quality peer review critical to a reputable scholarly communication system without suffering economic harm. Although embargo periods prior to the use of the publisher PDF vary (“SHERPA/RoMEO - Publishers allowing the deposition of their published version/PDF in Institutional Repositories,” n.d.), there is no clear pattern along disciplinary lines, and indeed, many publishers (225 according to the SHERPA/RoMEO lists) are willing to allow immediate self-deposit of this version. Many publishers, such as the American Chemical Society, are now willing to deposit the final version of an article on behalf of the author. The Houghton study into likely economic impacts of a broader federal open access policy states: “These estimates assume a six-month embargo period between publication and open accessibility. If there were no embargo, we estimate that incremental returns might be closer to \$1.75 billion. Hence, a six-month embargo reduces the returns by around \$120 million (NPV) (Houghton, 2010, p. 8).”

Thank you for this opportunity to comment.

Sincerely,



Daniel Linzer
Provost

Works cited

Battelle Technology Partnership Practice. (2011). *Economic Impact of the Human Genome Project*.

Battelle Memorial Institute. Retrieved from

<http://www.battelle.org/publications/humangenomeproject.pdf>

Ferrante, E., & Zgorski, L.-J. (2011, December 7). A New Visualization Method Makes Research

More Organized and Efficient. *National Science Foundation (NSF) Discoveries*. Retrieved

December 18, 2011, from

http://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=122509&org=NSF

Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-

Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research.

PLoS ONE, 5(10), e13636. doi:10.1371/journal.pone.0013636

- Houghton, J. (2010). *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs, report to SPARC*. Centre for Strategic Economic Studies Victoria University.
- Lipman, D. J. (2010). *Testimony, Public Access to Federally-Funded Research*. Washington, D.C. Retrieved from <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>
- Maier, D., Kalus, W., Wolff, M., Kalko, S. G., Roca, J., Marin de Mas, I., Turan, N., et al. (2011). Knowledge management for systems biology a general and visually driven framework applied to translational medicine. *BMC Systems Biology*, 5, 38. doi:10.1186/1752-0509-5-38
- Mons, B., & Velterop, J. (2009). Nano-Publication in the e-science era. Presented at the Workshop on Semantic Web Applications in Scientific Discourse (SWASD 2009). Retrieved from http://www.nbic.nl/uploads/media/Nano-Publication_BarendMons-JanVelterop.pdf
- National Advisory Council on Innovation and Entrepreneurship, Department of Commerce. (2011, April). Letter to Secretary Locke: Recommendations to facilitate university-based technology commercialization. Retrieved from http://www.eda.gov/PDF/NACIE_Letter-University_Commercialization.pdf
- SHERPA/RoMEO - Publishers allowing the deposition of their published version/PDF in Institutional Repositories. (n.d.). Retrieved December 20, 2011, from <http://www.sherpa.ac.uk/romeo/PDFandIR.php?la=en>
- SHERPA/RoMEO - Publishers with Paid Options for Open Access. (n.d.). Retrieved December 20, 2011, from <http://www.sherpa.ac.uk/romeo/PaidOA.html>
- Suber, P. (2008, August 2). Green/gold OA and gratis/libre OA. *Open Access News*. Retrieved December 20, 2011, from <http://www.earlham.edu/~peters/fos/2008/08/greengold-oa-and-gratislibre-oa.html>

- Swan, A. (2010). The Open Access citation advantage: Studies and results to date. Retrieved from <http://eprints.ecs.soton.ac.uk/18516/>
- Taylor & Francis Author Services - iOpenAccess & NIH policy. (n.d.). Retrieved December 20, 2011, from <http://journalauthors.tandf.co.uk/beyondpublication/iopenaccess.asp>
- Williams, H. L. (2010). *Intellectual Property Rights and Innovation: Evidence from the Human Genome* (Working Paper 16213). Cambridge, MA, USA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w16213>

Brandon Locke – brandontlocke@gmail.com
Student, University of Nebraska-Lincoln
Lincoln, NE

My name is Brandon Locke and I'm a graduate student in American History at the University of Nebraska-Lincoln. I have written my letter to address the specific questions in the RFI that most directly pertained to me. In short, I believe that it is unfair for private companies to capitalize off of publicly-funded research, forcing taxpayers to pay to fund the research and buy it back for public universities. It is also fundamentally counterproductive to innovation and education, by preventing students, researchers, and entrepreneurs from seeing and using the newest research. It is for these reasons that I believe publicly-funded research should be open to everyone and have no restrictions on use, besides proper attribution.

Question 1:

Agencies can best grow markets by doing everything possible to allow for open access and free use of research publications. Steps to increase access and usability of peer reviewed publications allow for publicly funded research to reach more people and be used in new and innovative ways to create economic growth and improve higher education. Wider and more productive analytical methods mean a much higher return on investment for research publications, lessening the burden on taxpayers, who currently fund both the research and the acquisition of scholarship. Agencies can grow and improve access and analysis markets by instituting rules that require publicly-funded research to be freely available and completely open to use.

By allowing individuals, companies, and organizations to freely access and use scholarship, the publicly funded research can reach a much wider audience and can be used in a wide variety of ways. Restricted access and use of publicly funded content means that public funds are not being used to the best of their abilities, and the return on investment is reduced. Full and open use means that students, entrepreneurs, and businesses can innovate without expensive barriers, making them better prepared to compete in the global marketplace. Full access and use allows readers to be much more productive with the information by using new techniques such as data mining and machine reading, and creating a new infrastructure for research. New pathways and connections can be made with open data and citation mapping. Under the current structures, information is locked into silos and users are not able to foster communication between research. Research is only as good as its reach and availability, and the current system is built to hinder access and use. Research can only be used by teachers and students if they have access and the more research that is available to students and teachers, the better and more up to date the education can be. Open Access also means that research will be available to the general public, increasing accountability for researchers and improving public information structures. Open access also fosters interdisciplinary application and greatly increases the value of established research.

Research publications can best be archived by making them immediately accessible and completely open to use in a centralized repository. Faster commercialization spurs economic growth, creating new jobs and advancing American businesses. Companies can also build upon public data and

improve services analytical and finding structures, like Google Scholar and goPubMed. By allowing entrepreneurs, scholars, and students to access them without restrictions, the entire data base can be used for data mining or derivative works, and can make the sum greater than the parts.

It is essential that research is available to the public immediately upon publication. It would not be conducive to innovation and cutting edge research if students, entrepreneurs, and researchers were forced to depend on old research when newer and better research is available. For taxpayers, immediate access to new research best utilizes public resources, and provides immediate benefits to universities and businesses, Open Access has been proven to increase citations, promote a diversity of sources within research, increase new research pathways, and make research immediately available for use in both application and further research.

In the current economy, graduate students will likely have a period between graduate school and employment, and the development of their research and their ability to keep up to date will be essential in job placement and success in their job. Once students leave school, they are met with expensive barriers that make it impossible for them to keep up with their field. This barrier also hinders entrepreneurs' businesses, leaving them at a disadvantage in the global economy. By making the most recent and advanced publicly funded research available, new graduates and small businesses stand a much better chance to utilize their skills and compete for a job or a share of the market.

Research can also be best utilized through storage and maintenance in a centralized repository, similar to the current NIH model. The benefits of an NIH-style access policy and infrastructure are estimated to be approximately eight times larger than the costs, and can be instituted at a relatively small cost. The NIH spends about \$3.5 – 4.6 million annually to provide access to all public-funded research, which is about 1/100th of 1 percent of their overall budget. Because of this policy, research is widely accessed and used by a broad population, with the majority of users outside of education. Full open access is ideal to making all of these ideas come to fruition. Restrictions on use of research also limit the possible value from research investment, and means that less money needs to be spent on duplicate research. It is also important that students be taught the most up-to-date information possible to best prepare them for the job market, and to make them best prepared to compete in the global market.

Question 2:

Publicly funded research can respect the intellectual rights of researchers and allow for the most complete utilization of research by implementing licenses like Creative Commons's CC-BY license.¹ The NIH currently allows articles to be used under "fair use," which protects authors but restricts some of the usefulness of the research. Full use would allow for scholarship to spread around the web and would facilitate open data mining and search methods, allowing researchers to find publications more easily. Full use of research allows taxpayers to get the largest and most complete return on their investment by spurning more innovation. The CC-BY license, like copyright, requires attribution of the work to the

¹ Creative Commons' CC-BY allows licensees to copy, distribute, display, and perform the work and make derivative works on it only if they give the author or licensor the credit.

author, so citations and impact ratings can only be increased in comparison to the current model. To further protect scholars' intellectual property; there could be an embargo period, where fair-use is applied, with the research moving to CC-BY or a similar open license. Again, this is not the best way to get productivity out of the research, but it does provide the author more rights over their work.

Question 3:

The federal government should provide permanent stewardship of research in a central database because it ensures that research is permanently preserved, made accessible, and most efficiently usable. By pooling all research together in a centralized location, everything is easily available and searchable in one place, and it's possible to build derivative databases that encourage communication between different publications (and possibly data as well), rather than different collections of research stuck in a number of separate silos, where integration is difficult or impossible. It is essential to the growing semantic web to have the underpinnings of the scientific and intellectual world interacting with the rest of the web. Federal stewardship is also very cost-effective, as stewardship for the NIH is only 1/100th of one percent of their budget.

Question 6:

Uniform requirements and mandates are necessary for consistent creation of publicly-funded research in universities. Because institutions often have researchers who hold grants from multiple agencies, all agencies should establish the same standards to smoothly implement research and allow institutions to focus on research rather than compliance. Uniformity amongst agencies means lower costs for institutions and an increased rate of compliance. Policies should take advantage of existing protocols to facilitate automatic deposit of manuscripts, and encourage the development of additional tools by interested agencies. Additionally, policies should integrate articles with grants management systems to improve agency accountability and provide information to the public.

Policies to increase tools and other finding methods should work to increase bibliographies and principal investigator profiles to better raise the connectivity of research and raise the profile of those researchers whose works are used and cited the most. These methods would allow universities to better measure research output and impact ratings, and would create better pathways to locate better research and allow universities and libraries to use repositories as teaching tools.

Question 7:

Educational materials such as book chapters, class notes, texts, syllabi, and conference proceedings should also be made readily available to the public, but may require different policies than those directed at journal articles. These types of unpublished works, most notably peer-reviewed conference papers and proceedings, represent a large portion of research and teaching materials that are very relevant to other scholars, as well as the public at large. Feedback from these kinds of papers is integral to the research process, and a wider audience can significantly improve research, as well as keep others informed on current trends and burgeoning research.



January 12, 2012

National Science and Technology Council
Task Force on Public Access to Scholarly Publications
c/o Office of Science and Technology Policy
Attn: Open Government Recommendations
725 17th Street
Washington, DC 20502

Re: ***Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research; Request for Information [FR Docket No. 2011-32943]***

Dear Task Force Members:

The American College of Rheumatology, representing over 8500 rheumatologists and health professionals, welcomes the opportunity to comment on the Administration's approach to public access components of the scientific research enterprise.

Rheumatologists treat patients with arthritis and other rheumatic and musculoskeletal diseases. These conditions can be painful, debilitating, life threatening and costly. Biomedical research plays a pivotal role in advancing diagnostics, treatments, and prevention strategies for patients with chronic diseases. Advancements in arthritis-related research have helped to prevent disabilities, allowing patients to continue working or return to work and contribute to their communities and the economy.

The ACR believes that scientific research publishing, like all other publishing, is a business governed by the copyright laws of the United States and most other countries. Unfortunately in some of the dialogues surrounding research publishing there is a conflation of the terms "public" and "free." We believe that the global journal corpus already provides a robust public access model for the dissemination of the peer-reviewed results of taxpayer funded research and other research. Government agencies that dispense funds to support taxpayer-funded research may wish to collect and publish free-of-charge reports generated by the recipients of those funds. However, we support the argument that these agencies do not have rights to the research articles written for and published by journals, nor is such a claim justified by any notion of an absence of access. Further, there is no evidence that making the current broad public access to the journal literature free will improve research productivity or the public welfare. We believe that free access, like copyright piracy, will be more likely to have the opposite effect.

(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 ACR is aware of no studies that support an argument that free access to the research literature will increase research productivity or economic growth. We do not believe it to be true that access to the research literature by those able to make use of it is rate limiting, and that there is untapped creativity that will be released if access is made free. The modern research enterprise is complex and requires major investments. Access to the research literature is not a constraint on this enterprise.

We do not accept the premise that because government funds scientific research, it is entitled to full access to and control of manuscripts reporting on this research. Publishing peer-reviewed research is expensive and has to be paid for. While the government pays for research, it cannot lay claim to the final publication. Having each funding agency open its database of funded projects, including research project reports and lay summaries, best serves the public interest and protects the scientific research enterprise.

Society today depends on a system of research communication that provides extremely broad access and strong quality controls. Research publishers are custodians of this system today because of the essential role that they play in the communication of scientific, technical, and medical research results. While it is the case that peer reviewers are generally not paid for performing the work of peer review, peer review is not free. Publishers invest hundreds of millions of dollars in end-to-end software tools to manage the peer review process and often also financially support the editorial groups who manage and perform peer review of submitted articles.

Government should not impose unfunded mandates that pertain to the outputs of the publishing process, including accepted author manuscripts and published journal articles. Such policies would not be justifiable, warranted or productive. Government-imposed public access policies would violate fundamental copyright principles by allowing the government to diminish existing copyright protections for private sector journal articles. Publishers make ongoing capital investments and incur significant operating expenses in carrying out these value-added activities. These are not paid for by taxpayer dollars. Any unfunded mandate has the potential to limit our ability to create the peer-reviewed literature in the first place. Nobody questions the considerable scientific value of peer-reviewed publications.

(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?

The federal government should not take accepted or published articles from publishers or learned societies, either directly or via a mandate placed on grantees, and make them freely available. However, several steps could be taken, including the following:

- Make funds available for the purchase of open access to published articles - these costs are a small fraction of the investment in the research.
- License content from publishers and learned societies and make it available to specific audiences.

- Make the funder-collected and maintained outputs of taxpayer-funded research, including grant reports or research progress reports, freely available to the public; private sector publishers could help make that content discoverable and linked to the journal literature.

(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?

Scholarly journal articles have been published for hundreds of years by a combination of society, not-for-profit, and for-profit publishers. This infrastructure has provided access to the literature for scholars and researchers, and the multiplicity of publishers has not prevented broad public access. In fact we strongly believe that the complexity of the system has promoted competition which has driven development of increasingly sophisticated platforms to deliver this content.

Publishers over the past decade have developed the Digital Object Identifier, a unique identifier for each piece of content, in this case a journal article. CrossRef, a not-for-profit group founded by a group of publishers, maintains 50 million DOIs. Almost 1,000 publishers and societies participate and assign DOIs to their published content items. Development of the CrossRef service has resulted in seamless navigation of the research literature by users, so that researchers using the bibliography in one article can link from a reference in the bibliography to the full text of the referenced article.

For many reasons including government budget constraints, we do not believe that the federal government is the best provider of these services, in particular as that development would involve using taxpayer dollars to duplicate an existing, well-functioning service.

(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?

There are a number of projects underway or envisioned for public-private partnerships. Please see our comments below.

Funding agency information

Most researchers acknowledge in their publications the research funder support they have received. However, there are no standards on how authors should do this. Consequently, funders find it difficult to know and track what publications have arisen from the research they have funded. Publishers are developing a means of standardizing funder information so that this information could be made easily available to funders. We believe that a community-wide solution of this type will be easier and far less expensive to construct than each agency developing its own response to the problem. Publishers are in the best position to provide a simple way of ensuring that journal articles are accompanied by standardized, high-quality metadata providing information about the agency, program, and even specific grant that funded the research. This proposal has been endorsed by CrossRef and major STM publishing trade associations.

With the successful implementation of this proposal, research funders would have access to the standard metadata from published articles that have arisen from the research they have funded. By displaying this information on their funder websites visitors will be able to follow the link to the publisher's platform, where article abstracts are freely available and the Version of Record, maintained by the publishers, is available through a variety of access mechanisms.

DOIs for data sets

Increasingly, investigators are being asked to share, or provide plans regarding how they will share with other researchers, the primary data and other supporting materials created or gathered in the course of their work. STM publishers and societies make significant amounts of this material available as supplementary material to published articles and are already participating in a number of initiatives designed to facilitate the sharing of data. Scholarly research publishers are willing to work with funders and database/repository operators to develop recommended practices for assigning DOIs to data sets and supplementary material so that datasets can be linked to primary research articles.

Author disambiguation

Name ambiguity and attribution are persistent, critical problems embedded in scholarly research. STM publishers are working to eliminate this problem through an initiative called the Open Researcher & Contributor ID project. ORCID is a newly established non-profit organization working to establish an open, independent registry of researchers that is adopted as an industry-wide standard to resolve systemic name ambiguity by means of assigning unique identifiers linkable to an individual's research contributions. Researchers will be able to create, edit, and maintain an ORCID ID and profile free of charge, including defining and controlling their own privacy setting.

Such a standard will not only enhance the scientific discovery process but also improve the efficiency of funding and collaboration. Participation in ORCID is open to any organization that has an interest in scholarly communications. All software developed by ORCID will be publicly released under an open-source software license approved by the Open Source Initiative. ORCID is governed by representatives from a broad cross-section of stakeholders including publishers, societies, libraries, and other institutions.

Content mining

Content mining has the potential to be useful to the scientific community in driving interdisciplinary research and supporting the identification of new areas of discovery. Publishers and their society partners are committed to managing content in digital formats to ensure that users gain maximum benefit. Publishers should work with research funders to develop pilot projects for journal content mining that would identify, organize, and perform analysis to identify and create conceptual links within and between that content that are not obvious to initial human inspection. Although there are various ways to perform this type of processing, certain elements are common to all methods, including an automated way to process all sizes and types of content in which to identify relevant information, and facilitate its extraction and analysis.

Such pilots would focus on goals such as:

- Structuring input text, deriving patterns within the structured text, and evaluating and interpreting the output;
- Extracting semantic entities from publisher content for the purpose of recognition and classification of the relations among them; and

- Enabling developers who wish to design and implement applications to analyze our content or test applications as part of their research within publisher content.

Consensus approaches within the community could also be explored for developing better standardized, mining-friendly content formats, a shared content mining platform, and commonly agreed permission rules for content mining.

Linking to/from research reports

Publishers of scholarly research should collaborate with research funders to determine whether and, if so, how publisher content could be “mapped” against research reports and other funder content. The goal would be to make connections between content items that would add value and richness to both groups’ digital offerings. Specifically, this collaboration would send users from publisher websites to the funder web site to view free government-sponsored research reports, and would send users from funder sites to view free abstracts and links to the Version of Record of articles connected to a particular research report or funded project.

If successful, this will result in interoperability between funder and publisher content and would enable publishers to work with research funders to identify, organize, evaluate, and highlight published results from their research funding and identify relationships, projects, and offerings.

Possible outcomes of the pilot could include:

- The ability to identify all agency-funded research within publisher offerings and the ability to deliver associated metadata to that funder;
- The ability to establish mechanisms and approaches that could be implemented (for all research funders) across the industry;
- A capability to report to major funders on the impact of the research they fund, e.g. through bibliometric and other tools;
- A “research dashboard” capability or the ability to contribute to one already in existence – e.g. <http://rd-dashboard.nitrd.gov/>;
- A mechanism for low-cost content rental access to published articles (VoR);
- Subject area content portfolios of NSF-funded research articles for internal NSF use (e.g. study sections); and
- The opportunity to use the <http://www.science.gov/> and <http://www.research.gov> platforms to extend this pilot to other federal funding agencies.

(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?

Please see comments for question four related to public-private partnerships.

(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 awardees institutions, scientists, publishers, Federal agencies, and libraries?

As surveys of researchers show, those who can benefit from access to the peer-reviewed journal literature already have access. We also believe that the publishers own and have copyrighted journal articles which are published in the journal titles they publish.

Both publishers and learned societies are committed to the wide dissemination of content. We support any and all sustainable access models that ensure the integrity and permanence of the scholarly record. This includes 'gold' open access, where publication is funded by a publication fee or article processing charge. Many publishers now offer open access options and/or publish open access journals, and work closely with funders, institutions, and governments to facilitate these developments. We believe that authors should be able to publish in the journal of their choice, where they feel their work will be best reviewed by their peers and where its publication will have the greatest potential to advance their field. Research funders could provide a fund to publishers to cover gold open access publishing fees.

(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?

They should not. Publishers also invest in these other types of content used by researchers, often by conceptualizing the project, commissioning the content, and investing heavily in its development. Any kind of mandated free access to that content is simply an expropriation of that content.

(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?

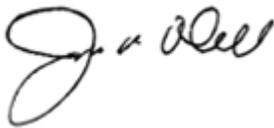
We believe that peer-reviewed papers should not be made public within the duration of the article's copyright without the copyright holder's permission. For accepted author manuscripts and published journal articles, both of which publishers have invested in heavily, we believe that publishers – and learned societies – themselves should be able to determine the business models under which their publications operate. This should include the time, if any, at which the final peer-reviewed manuscript or final published article are made publicly available.

Peer-reviewed papers are not the direct result of the Federal Government's investment. They should not be made freely available to the public unless the copyright owner authorizes the government to do so. Since the mid-1990s, the science journal publishing industry has been a key player in the truly dramatic digital revolution in the sciences, investing heavily to drive the shift of published research from print-only to "E-only." Rapid innovation in the publishing industry has dramatically improved functionality and efficiency for doctors and researchers, who can now perform complex searches of journals, immediately retrieve and print full text articles, link instantly to other cited articles, export text to other databases and programs, and receive e-mail alerts when new journal issues are released. Mandating free access will stifle innovation in what is

now a rapidly changing environment, both by decreasing the amount that publishers are able to invest and reducing their incentive to try new approaches.

The ACR appreciates the task force's review of recommendations for ensuring long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally-funded scientific research. Particularly in these challenging financial times, we believe the task force should strongly consider the recommendations we have provided in these comments. We stand ready to assist you further on these issues that affect the conduct of scientific research related to rheumatology and the broader rheumatology community, including the health and quality of life of our patients. If we can be of assistance to you in any way, please contact Adam Cooper, ACR director of government affairs, at acooper@rheumatology.org or (404) 633-3777.

Sincerely,

A handwritten signature in black ink, appearing to read "J. R. O'Dell". The signature is fluid and cursive, with the first letter of the first name being a large, stylized 'J'.

James R. O'Dell, MD
President
American College of Rheumatology



The Association for Research in Vision and Ophthalmology
1801 Rockville Pike, Suite 400 ■ Rockville, Maryland 20852-5622
arvo@arvo.org ■ +1.240.221.2900 (Tel) ■ +1.240.221.0370 (Fax)
www.arvo.org

January 12, 2012

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

On behalf of The Association for Research in Vision and Ophthalmology (ARVO), I submit the following comments in response to the RFI issued on November 3, 2011. ARVO is the largest and most respected eye and vision research organization in the world. Our members include more than 12,600 researchers from over 80 countries. ARVO encourages and assists research, training, publication and knowledge-sharing in vision and ophthalmology. ARVO publishes two medical/scientific research journals which are published online only and are hosted at HighWire Press which is considered by libraries internationally as a trusted site and archive. In mid-2012 ARVO will launch a new online-only journal on the topic of translational ophthalmic science & technology, which will also be hosted at a trusted site. In addition, ARVO voluntarily deposits complete articles of all NIH-funded research published in its journals in PubMed Central on behalf of authors and at no charge to the authors.

ARVO supports the principle of providing the public with access to the federally funded scientific research. However, we believe that releasing the peer-reviewed research articles in direct competition with scholarly publishers undermines the ability of associations and societies to maintain the high quality standards of selection, review, production, and publication as well as protection of the scientific record.

Scholarly publishers provide essential services that ensure the quality and integrity of journal content. Through peer review publishers and the scientific community identify scientific shortcomings and inadequacies which continue through the revision and re-review of articles. Over 50% as for some journals as much as 75% of submitted articles are ultimately rejected because of these inadequacies. The continuous feedback to authors through review and editing immeasurably improves the final published product. Publishers also serve as guardians of scientific ethics and standards to ensure accuracy, reliability, ethical treatment of patients and humane treatment of animal subjects.

In addition, in our opinion, the current NIH policy confuses the community and the public regarding the completeness of the “public” record and who the actual publisher of the scientific material is. NIH has established itself in direct competition with private publishers while using public taxpayers’ funds to complete their redundant work. These activities jeopardize the financial viability of journals, particularly those published by learned societies

and associations that are dependent on subscription revenue and author charges to sustain their journals and educational activities.

In response to the specific questions listed in the RFI, please accept the following comments.

- (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?**

RESPONSE: ARVO supports the practice of access to all scientific research to scientists, practitioners, and the learned public, not just access to federally funded research. Like most learned societies and scientific associations ARVO provides access through subscriptions to small and large institutional, commercial, and academic libraries worldwide and provides free and open access to the public within six (6) months of publication. Also like most scholarly publishers ARVO's journals are hosted online in internationally recognized sites that are considered by the library and academic communities as "trusted" sites. ARVO's journals are hosted currently at HighWire Press. As with most non-profit, scientific associations, we are committed to the preservation and archiving of all of our content through participation in programs such as LOCKSS, CLOCKSS, and Portico, through which participating institutions may download and store as a deep archive all content from Volume 1, Issue 1, page 1 to current data. The deep archive can be used to restore data lost at the institution and to update files to meet then current technological standards. ARVO, through its electronic host, archives XML files and all metadata, as well as figures and tables in their native format so that the content can be repurposed or configured to meet future technological needs and formats. Internal links in articles to already published content and well tagged metadata allow robust search engine results that ensure discoverability and, thus, increase productivity for those engaged in the scientific enterprise. For Federal agencies to create additional archives or access points appears to be a redundant and inefficient use of federal funds that could diminish funds available for ongoing scientific research.

For example, the NIH requires deposit of all federally funded, peer-reviewed scientific journal articles. It then reprocesses all files, using at least one **non-US vendor** to do so. This use of U.S. taxpayer funds to support non-US vendors does not meet the stated criteria. The development of these materials housed at PubMed Central is redundant to the online content hosted and archived at trusted resources.

- (2) What specific steps can be taken to protect the intellectual property interests of publishers, scientist, 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?**

RESPONSE: Having a single source of the definitive scientific content of a journal ensures the integrity and completeness of the scientific record; that source should be the original article published by in journal. For most journals that have an electronic publishing component the official article of record is the electronic, XML/HTML version of the article. For example, at trusted sites if an erratum (correction) or comments about a specific article are published they are linked to the original article in perpetuity. With multiple sites the likelihood of constructing and maintaining the complete record with correct links is more problematic and could even endanger the lives of patients if the record is incomplete. One of the hallmarks of scientific research is the ability to duplicate results; if future research determines a flaw in the original findings, it is of paramount importance that the original research be linked through references and use of DOIs to future findings. These links (called forward linking) are maintained by most scholarly publishers today. The use of embedded DOIs (digital object identifiers which are discrete for each article and include a publisher's identification) for all parts of articles, including figures, tables, and text units, would also ensure that intellectual property rights are maintained. DOIs are searchable and are associated with the original **publication of record**—the journal article and thus the authors and publisher. We suggest that policies encouraging the use and deposit of DOIs by all publishers for all articles be considered.

Conversely, establishing multiple deposit sites with varying times of free and open access to articles/content is counterproductive from a scientific integrity, content management, and financial standpoint. Publishers have already made the investment in time, scientific resources, and financial resources; multiple other site hosts is redundant and financially wasteful. In addition, sites such as PubMed Central (under the National Library of Medicine (NLM)) have on part of the body of knowledge in any given medical/scientific field—that content funded by a Federal agency, specifically the National Institutes of Health. The NLM also is selective in which journals are included in their indices. It usually takes two to three years after a medical journal begins publication for the NLM to consider adding it to the collection. In the case of new journals that would mean that up to three years of content would never be included in the Library.

- (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?**

RESPONSE: Given the technology that exists today there is one rationale for a centralized repository and that is the development of specific content databases and analytic tools that aid in data mining. However, the shortcoming to that rationale is that if the database only contains federally funded content then potentially greater than 50% of the body of knowledge in any given scientific area would not be included in the database. Is the government in a position to **at least double** the expense currently incurred to development and maintain the content? The private sector has the incentive and expertise to develop new tools to meet the needs of all stakeholders, including search engines that can identify all content that is appropriate for inclusion, including across disciplines. Given the nature of translational

research, it is highly likely that relevant content in a particular area may be published across journals, and not necessarily in journals included in NLM or PubMed Central and possibly not funded by a Federal agency, such as new developments in engineering fields that directly affect medical technologies. Again, leading to an incomplete picture in the government's centralized repository.

If a Federal agency maintains custody of all published content and controls access to that content, non-profit associations and societies would not be able to sustain a publishing program and several thousand journals would cease publication. The remaining commercial publishers could not and, we believe, would not be able or willing to absorb the content. If non-profit organizations lost their publications and the associated revenue it would significantly and negatively impact the educational and support services that organizations provide to their members. Journal revenue helps support all other facets of non-profit organizations, including scientific meeting large and small. In ARVO's case, at least three small, highly specialized meeting would be cancelled; the meetings lose money and are almost entirely supported by publication revenue.

The government can ensure long-term stewardship of decentralized content by working with publishers and organizations such as NISO to quickly develop standards that must be met. If hosting sites were evaluated in terms of these standards as well as for the technological reliability and discoverability of all content, then the sites could be deemed "trusted". A regular monitoring of and support of new host sites as well as established "trusted" sites could ensure compliance.

- (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?**

RESPONSE: There are three well-respected models of existing publisher-library-archive partnerships that exist and that archive ALL content, not just federally funded research: Portico (portico.org); JSTOR (jstor.org); and ITHAKA (ithaka.org). Portico has 135 participating publishers with over 12,300 individual titles and over 19.35M individual items preserved. Other groups that permit local library archiving of titles to which they subscribe include LOCKSS and CLOCKSS. Portico and JSTOR have robust tools and are committed to full archiving of all content as well as providing access to content if a publisher ceases to exist or ceases publication. A private company that allows subscriptions to content as well as indices is Thomson/Reuters Web of Knowledge. This is a very large database that is not always current and is very rigid in its requirements. It does not use the DOI despite repeated requests from large and small publishers.

- (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-review publications resulting from federally funded scientific research are publically available to ensure that these publications can be easily found and linked to Federal science funding?**

RESPONSE: The minimum metadata that needs to be made available is the full citation of an article. That is, the authors names, title of article, journal name/abbreviation, volume, issue, page, year. This would allow searching down to the article level, if needed, or at a journal level. It would be useful for all publishers to incorporate the article DOI into the article citation. Using any search engine the online article would be retrievable regardless of file location. This is standard metadata for all scientific journals, most of which now use the NLM DTD (a DTD is a document that defines how all elements in an article should be identified (tagged). The National Library of Medicine is the most commonly used DTD worldwide.) The metadata is currently available through the National Library of Medicine for most medical journals. However, please note, if someone uses a general term in a search engine, such as glaucoma, articles including that word will appear. Scientists, clinicians, and researchers are very sophisticated technology users and have little difficulty identifying relevant data across platforms, publications, or archives. Journals include funding information, Federal or otherwise, in the body of their articles and do include the citations of articles published (with links if required) in all grant and support applications.

(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?

RESPONSE: The best way to maximize the benefit to all stakeholders is to allow publishers to set reasonable times to full open and free access to their content based on the frequency of publication (monthly, bi-monthly, bi-weekly, etc.) and the timeliness of the content. In this way publishers can control their expenses and maximize the subscription revenue that results in 30-50% of revenue of most scholarly journals. Library organizations have stated that if a journal is published monthly, most librarians would not consider dropping the subscription if the journal became open and free at 6 months or longer after publication. That is, if a January issue is published January 1, and that issue became open and free on July 1 of the same year.

Since publication costs (charges to authors) are generally low (less than \$3,500) relative to the size of grants, these costs should be accommodated in all grants, with no expiration date of funds being paid to publishers. A common response to receipt of an invoice for publication charges is “my grant has expired and I don’t have the funds.”

Again, it should be mentioned that establishing government archives for federally funded articles is redundant and expensive. Publishers are already providing archives and are paying for them. Realistically, the average American citizen is not reading scientific journals and most do not have the training or experience to evaluate the content for relevance. Patient advocates claim that everyone should read the articles or be able to access them. In fact, most medical publishers will provide single copies of articles to patients or family members of patients with the recommendation that **they take the article to their treating physician to determine the relevance and to aid in asking questions.** These articles are not intended for lay people; they are written by and for experts in their fields.

- (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?**

RESPONSE: ARVO does not believe that book chapters should be covered by these policies. Publishing textbooks or scholarly tomes is as expensive, if not more so, as publishing a journal. If chapters will be given away publishers may stop publishing any books that contain federally funded material. Regarding meeting abstracts, ARVO's meeting abstracts are free from the day they are launched. As abstracts are usually partial or preliminary findings, they should not be considered the definitive information on the topic.

- (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 weight public and private benefits and account for external and 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?**

RESPONSE: This is an extremely difficult question to address. So much depends on the timeliness of the material, that is how important is the content to the community today, in 6 months, in 12 months, so that most researchers/scientists would have read the material by that date. ARVO has determined that the highest "reader" rate is achieved in the first 6 months of publication. While the readership has a long tail, that is, it will be read regularly for about 6 years, and then sporadically after 6 years but will continue to be read when it is 10 – 15 years old. ARVO spoke with subscribing libraries regarding its largest journal, which is the leading research journal in its field in the world, and learned that generally they would not drop the subscription even if the articles were not free and open for 12 months, and then ARVO decided that as a service to the community, to help stimulate research, and provide timely content, particularly to researchers in developing countries, it would open all articles to public access after 6 months in its largest journal. ARVO's second journal is open access immediately upon publication and has been since its launch 10 years ago. However, to sustain this model of open access there are author charges that must continue to increase in order to meet the costs of producing and hosting the journal online. ARVO is, therefore, comfortable with a 6-month embargo period.

Thank you for the opportunity to respond to this import RFI.

Submitted by:

Karen Schools Colson

ARVO, Director, Publishing Projects

On behalf of The Association for Research in Vision and Ophthalmology

1801 Rockville Pike

Rockville, MD 20852

January 12, 2012

Task Force on Public Access to Scholarly Publications
National Science and Technology Council
Office of Science and Technology Policy

SUBJECT: Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research

Arizona State University (ASU) appreciates the opportunity to comment on approaches for ensuring long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally funded research.

We wholeheartedly support the goal of providing timely, easier and less costly access to publications that result from federally funded research and point out that the public does have such access through numerous depositories including traditional libraries and, increasingly, through electronic sources.

Comment 1: Growing markets related to access and analysis and using those markets to grow the economy and improve productivity of the scientific enterprise.

Response: We echo the comments provided by the Council on Governmental Relations (COGR) that publications resulting from federally supported research are already available to the public. As a public institution conducting primarily fundamental research we caution regulators to avoid creating policies that may cause confusion and over regulation for researchers and the institutions that support them.

Comment 2: Protection of Intellectual Property

Response: The assignment of the ownership of the intellectual property as reflected in a publication, itself, as opposed to the intellectual property reflected in an invention and/or the associated research data is the responsibility of the author; in this case the researchers, scientists, etc. We can remind researchers to maintain their rights individually and to use the information included in a publication for educational and research purposes. We can provide researchers with proposed language to insert in copyright agreements to enable access but, ultimately, the responsibility falls to the researchers/authors of the publication.

Without the appropriate protections to the assigned copyright, publishers may be reluctant to include the work of federally sponsored researchers in their publications. Absent these protections and this outlet, public access will be profoundly undermined.

Comment 3: Central versus Decentralized Management of Access

Response: As the most obvious “central” repository we are curious why the Federal government wants to assume the responsibility of publishing or providing publication oversight to all the

published results of federally sponsored research. We do not see that as an effective or efficient use of Federal resources and we do not have confidence that it will provide for better stewardship of the scientific resources than the current system.

Comment 6: Maximizing Benefit while Minimizing Cost and Burden

Response: This issue presents a challenge for us. Of primary concern to us is the lack of any real controls in the process of research publication as we do not play a stakeholder role in the relationship between authors and the journals that publish the results of research. Rather the role we play as an institution is to remind researchers (who are the authors) to maintain their rights individually, to provide real public access and to ensure that the published version is available in the appropriate format for search and analysis. One area that we can participate in is to develop and provide model language to be included in copyright agreements that will facilitate more open access.

In the end, the responsibility falls to the researchers/authors of the publication. Publications that result in whole or in part from a federally sponsored award may appear several years after the completion of the funded research. This poses an additional challenge to us as an institution because the investigator/author may have moved to a new institution in the intervening period. Tracking publications from collaborative research with researchers/authors from more than one institution is a monumental task. Like COGR we expect that compliance with a government-wide policy will become a usual and customary practice in the research community and, as a result, researchers/authors will meet this obligation as a regular part of the publication process. But in the intervening period, the burden associated with a government-wide mandated process will be significant.

The costs involved in revamping the current expectations to the institution and/or investigator are real. As NIH has moved forward with its policy on public access, researchers have discovered a shifting of publication costs to the author. There are direct charges for the submission of articles – “article processing charges.” Journal charges to authors for public access for a single article have reached, in some cases, \$3,000. NIH has reminded the community that publication charges are an allowable expense against a grant. However, in many cases publications will be accepted after a grant has closed. As a result, we are expected to assist researchers in meeting these unexpected costs, putting greater strain on institutions like ASU. Charging these publication costs to a grant, if possible, will result in a real reduction in funds available to conduct the research itself. **Unless there is to be a government-wide investment to support the costs of publication, a government-wide policy requiring public access to publications becomes an additional unfunded mandate for the research community.**

Comment 7: Broaden Coverage to all written publications

Response: Expanding the current public access model from journal articles to book chapters, conference proceedings, etc., will only increase the costs and burden on all parties. Books are available in libraries; conference proceedings are often works in progress that may, eventually, be presented in print either in a journal or book. We do not believe that pursuing these research products will increase access to the ideas and data.

We recommend that the immediate emphasis should focus on the methods that researchers currently use to disseminate their results, primarily peer-reviewed journals.

Comment 8: Publishing Community Response

Response: Publishers are generally the holders of the copyrights to published articles and are the party responsible for providing public access. We are unaware of any evidence that the customary embargo of twelve months has prevented access to publications, hindered the growth of existing and new markets or undermined the productivity of the American scientific enterprise.

In closing, we support harmonization and coordination among the Federal agencies in order to streamline compliance with Federal mandates and regulations.

We appreciate the opportunity to comment on the questions posed by OSTP on the value of public access to peer-reviewed publications.

Sincerely,

Debra Murphy
Director, Office of Research Integrity and Assurance

Request for Information: Public Access to Digital Data Resulting From Federally Funded Scientific Research

The Zoological Society of London is pleased to respond to the Office of Science and Technology Policy (OSTP) request for information concerning long-term stewardship and broad public access to peer-reviewed scholarly publications resulting from federally funded scientific research.

The Zoological Society of London was founded in 1826 and has published scholarly zoological science since 1830. ZSL publishes three peer-reviewed journals, in partnership with Wiley-Blackwell. ZSL is a charity and learned society: revenue generated from publishing contributes to the science and conservation activities undertaken by the Society. Specifically, income generated from publishing funds ZSL's annual programme of science and conservation events, including conferences, seminars and workshops, which provide opportunities for knowledge sharing, mentoring and priority setting.

We support sustainable models of access. ZSL's journals use a subscription-based model, with an option for open access, where an author (via the institution or funder) provides payment to fund publication.

We appreciate the opportunity to participate in the Administration's consultation.

(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?

Publishers invest significant resources to establish new markets in developing economies to extend access to research journals: exploring these new markets is part of the 'journals business'. Publishers have pioneered new subscription initiatives that optimise access to academic literature, and support programmes that provide free or very low cost access to universities, research institutes, schools, hospitals, governmental offices and national libraries in countries with the lowest gross national income per capita through the Research4Life, eIFL and PERii schemes.

The US public should have access to research data resulting from federally-funded research. Project reports and summaries are the appropriate conduit for disseminating the outcome of federally funded research. These reports should be publically available via funding organisation databases, and can be linked to grant applications and agency auditing systems.

The process of scholarly publishing adds value (through editor expertise and peer-review). The journal article is a product in its own right, and not one that should be expropriated without compensation. Free access to papers resulting from federally funded research undermines copyright, intellectual property rights, jobs and exports. The value-added activities provided by publishers are not paid for by federal agencies and they are not free. In a rapidly changing publishing environment research communication needs to be operated under strict quality controls. Considerable investment is made to develop software (i.e. to manage the peer-review process) and the costs associated with supporting editorial groups are often met by the publisher. Policies that starve the system of resources will negatively impact the quality of published journal articles. Publishers should be able to develop and use appropriate business models to recover their investment.

Policies for archiving and public access that are tied to growth in the US economy may well serve to undermine the competitive advantage of the US scientific enterprise. A free access policy will allow the benefits of federally funded scientific research to be disseminated globally, giving an advantage to economies that have not made an investment in it. It could be argued that this is a poor return for US taxpayers. How can this be managed to promote growth in the US economy?

Government agencies should work with publishers to establish the most appropriate ways of meeting access requirements. This will provide an opportunity for all parties to develop systems for information provision which preserve the integrity of the scientific publishing process and give US taxpayers a return on the investment in research.

(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?

Providing global public access to the peer-reviewed published output from federally funded research can expose content to piracy and other unauthorized dissemination, which would undermine the income that scholarly publishers require in order to continue their investment for the benefit of the scholarly community.

The PEER project <http://www.peerproject.eu/reports/> into the behaviour of researchers has shown that the publisher-created Version of Record (VoR) from a peer-reviewed journal is considered by researchers to be the authoritative, definitive version (over versions in subject or institutional repositories). Publishers take seriously their role as the stewards of the research literature and version control is important to the integrity of the scientific record. The intellectual property interests of the publisher who have invested in the VoR should be protected.

The most efficient way to ensure the protection of intellectual property interests of all stakeholders would be to make the final research report freely available. This would permit broad dissemination of the research results obtained from federal funding. Subsequent research articles present far more than research data: they are a synthesis of the researchers' knowledge to date, and should be protected by appropriate intellectual property rights.

(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?

With advances in technology, centralization is not required. Any move towards centralization would require unnecessary duplication of effort and expense. Publishers have developed tools to ensure interoperability between different access systems. One example is CrossRef <http://www.crossref.org/>, an independent membership association, founded and directed by publishers. The mandate of CrossRef is to connect users to primary research content, by enabling publishers to work collectively. CrossRef is the official DOI® link registration agency for scholarly and professional publications and its citation-linking network covers tens of millions of articles and other content items from thousands of scholarly and professional publishers.

Publishers have an excellent record of developing discipline-specific tools to meet user requirements and have invested heavily in the development of tools to achieve interoperability between different access systems. Can this responsibility plausibly be the remit of government, or a good use of government funds?

(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?

Proposals have been put to NSF for collaborative projects to enhance the public access, utility and preservation of publications resulting from federally funded research. These initiatives include standardizing the collection, display and use of metadata to acknowledge the federal grant supporting the research from which a scholarly publication derived and linking back to the federal agency website. Also proposed is a project to understand the requirements for and benefits derived from content mining and to establish a methodology for overcoming current barriers, so that publishers can facilitate content mining with sustainable business models.

Other public-private partnerships that take advantages of existing publisher archives include:

Author disambiguation. STM publishers are working to eliminate author ambiguity through the Open Researcher & Contributor ID (ORCID) project. ORCID will allow researchers to create, edit, and maintain an ORCID ID and profile free of charge. Participation in ORCID is open to any organization that has an interest in scholarly communications. All software developed by ORCID will be publicly released under an open-source software license approved by the Open Source Initiative (OSI). ORCID is governed by representatives from a broad cross-section of stakeholders including publishers, societies, libraries, and other institutions.

Funding agency information. Acknowledgment is often given to the research funding organisations in the published journal article but this is not standardised. Funding organisations are tasked with tracking publications resulting from their funded research. A community-wide

solution to this problem is being pursued. This proposal has been endorsed by CrossRef and major STM publishing trade associations.

Content mining. Content mining has the potential to be useful to the scientific community by supporting new areas of scientific discovery. The aim is for publishers, their society partners and research funders to work together to develop pilot projects for journal content mining that would identify, organize, and perform analysis to create conceptual links within and between content that are not obvious to initial human inspection.

(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?

Publishers are involved in a project with CrossRef and the Department of Energy (DoE) to standardize the way funding information is collected and included in article metadata. This will allow agencies to easily obtain information about publications resulting from federally funded research.

Core metadata can allow users to find related information without the requirement of accessing the full text. The Dublin Core Metadata Initiative provides key specifications and best practice regarding the use of metadata for the description of various digital resources (including books and journal articles). It facilitates interoperability of different applications and vocabularies and optimizes the metadata for searching. In addition, CrossRef provides a cross-publisher linking network, allowing readers to easily link to other resources of interest on other publisher platforms. This works seamlessly through DOIs and metadata which are embedded in articles and other content as part of the value-added publication process.

(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?

Authors should have the freedom to submit the outcome of their research the most appropriate journal, which has the greatest impact and relevance to their field. This results in research papers being reviewed and read by research peers and furthers the advancement of the science. Research funding organisations could provide funds to publishers to cover the fees associated with the publishing process (Gold Open Access). The publication of research reports from funding agencies would provide open access to the research without compromising the integrity of the published journal article. Funds retained in the funding agency account could be credited on a paper-by-paper basis, which would give federal agencies and taxpayers an ongoing return on their

investment. Research published long after the termination of the agency grant can still be credited to the awarding agency.

(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?

No. Publishers add value to all types of content they produce. This investment is of enormous benefit to community and any government policy that mandates free access compromise future investment in high-quality publication, dissemination and preservation of the research.

(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?

Federal agencies should not impose inappropriate embargo periods on non-federally funded businesses. Any embargo period shortens the period of copyright protection and peer-reviewed papers should not be made public within the duration of the article's copyright without the copyright holder's permission. The embargo period established for each journal is evaluated according to the needs of the market. The journal half-life is an important means to establish the length of usage of an article within a given discipline. Foreshortening the time a publisher is able to recoup their investment has the potential to seriously damage publishers and the overall economy.

Publishers have invested considerably to optimise the speed and functionality of the online publication of research articles. This digital revolution provides rapid access and complex searching, access to datasets, supplementary information, and links to cited material. Furthermore, SMT publishing benefits from publisher activities which direct readers to content, alert readers to new research and marketing initiatives, Government agencies should collaborate closely with publishers, scholarly associations and universities to achieve the full potential of publicly accessible databases. Primary data and supplementary materials are increasingly being made available to the research community and publishers are at the forefront of promoting best practice in this area, for example, by presenting and repurposing data in formats to increase cross-referencing and reuse, and by linking datasets to primary research articles.

Policies which mandate free access will hamper future investment in these areas. In a rapidly changing publishing environment, continued development in functionality of the research article is critical. A co-ordinated response to the needs of the academic community and funding bodies can be met by publishers. The aim should be to produce a climate of equitable access while protecting appropriate intellectual property rights.



OFFICE OF SCIENCE AND TECHNOLOGY POLICY

REQUEST FOR INFORMATION:
PUBLIC ACCESS TO PEER-REVIEWED
SCHOLARLY PUBLICATIONS RESULTING
FROM FEDERALLY FUNDED RESEARCH
Published at 76 Fed. Reg. 68518 (November 4, 2011)

WRITTEN COMMENTS OF

COPYRIGHT CLEARANCE CENTER, INC.

January 12, 2012

Introduction

Copyright Clearance Center, Inc. (“CCC”), submits these written comments in response to the Request for Information of the Office of Science and Technology Policy published at 76 Fed. Reg. 68518 (November 4, 2011) regarding public access to peer-reviewed scholarly publications resulting from federally funded research.

CCC offers a host of different forms of licensing of text-based copyrighted materials, on behalf of publishers, authors and other creators (collectively, “rightsholders”), to users of all kinds, including academic, business and government organizations. These forms of licensing include (i) traditional collective licensing (one license covers all designated use of a repertory for a year), as well as (ii) both centralized (at CCC’s office and website) and decentralized (at the websites of participating rightsholders) licensing on an as-needed basis. The development of these different licensing models and modes of access was driven by the varying needs for use of content of many different types of users of scholarly (as well as trade, news and educational) publications. In the course of our business, we represent thousands and thousands of rightsholders of those copyrighted works, including thousands of rightsholders in peer-reviewed scholarly publications, and we sell licenses every year to thousands of users (virtually all of which – businesses, colleges and universities, and government agencies) are themselves organizations that together represent tens of millions of employees, students and associates.

Licensing is Part of a Larger Distribution System to Help Support Dissemination of Scholarly, Peer-Reviewed Articles and Books

Rightsholders use licensing services like those offered by CCC as part of what the Request for Information terms “long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally funded scientific research,” p. 68518. Such licensing helps contribute to the financial resources necessary to enable peer-reviewed publishing – and the awareness, curation and stewardship of scholarly output that such publishing represents – as well as to enable public access to the resulting books and articles. Licensing in the scholarly publishing industry fulfills this task not only by collecting royalties from users who pay for rights to re-use (such as the rights to reproduce and redistribute) published materials but also, in both the fee-paid and the large number of no-fee transactions managed by rightsholders and CCC, by providing feedback to rightsholders about the materials that users actually use. The financial contributions from the communities that use peer-reviewed materials for separate research (such as R&D-intensive commercial businesses) help pay for the overall publication effort; the feedback from both paying and no-fee licensees helps guide the direction of future publications.

In its licensing services, CCC serves both rightsholders and content users of all types and sizes. In CCC’s experience, the breadth of differing needs for licensing services arise from a host of distinctions not only among rightsholders’ business models, but also among the fields in which rightsholders research and publish, the user communities to which their materials are directed and the uses those users make of the materials, and their own funding sources. These differences underlie answers which the Office will receive to many of the questions posed in the Request for Information, including, for example, about the nature of interoperability among sources of materials or the need for, and the length of, embargo periods before public access is made free-of-charge.

In the past year or so, the scholarly publishing community – including both not-for-profit and for-profit publishers – has shown itself capable of addressing these distinctions while still serving those who need or want access to published materials. It has done so by expanding the distribution mechanisms available to different markets; one major example, assisted by CCC and its decentralized licensing facility (RightsLink[®]), is the development of an “open access” publishing model that enables interested authors and research institutions, rather than readers, to pay for publication and to direct the publisher to make public access to the articles they authored free-of-charge immediately upon publication – thereby supporting the traditional, and traditionally important, peer-access publishing model through a different funding mechanism. At the same time, varying distribution models – from annual subscriptions, to pay-per-copy or pay-per-article, to pay-for-access, to “rental”, of journal articles and books – are widespread in the publishing industry, with different models (and different pricing structures, including between those available to commercial, non-commercial and even individual users of the same material) of different utility to different users. Because these many models are available, the publishing industry is able to serve the scientific and other communities – both as creators and as

users of copyrighted material – in a market-sensitive way, drawing multiple sources of revenue together to support the entire system.

Any “one-size-fits-all” model for access to scholarly publications, like that enacted by the National Institutes of Health for public access, fails to distinguish the different needs of rightsholders and users, and even of the funders of research both inside and outside the government. By doing so, such a model upsets the balance among revenue sources that sustains science publishing and risks collapsing systematic dissemination of scientific research altogether. Recognition of such a risk has enabled private funders of research, such as the Howard Hughes Medical Institute and the Wellcome Trust, to work with rightsholders to sustain the balance. In contrast with the NIH policy, the America COMPETES Act, upon which this Request for Information is based, (i) established a public access policy for research funded by the National Science Foundation, (ii) provides a constructive model that can be replicated in a timely manner at other federal agencies, and (iii) is far more likely to support the long-standing and well-functioning scientific discovery and innovation system of publishing experimental results, maintaining the consequent economic benefits and employment, and supporting the Constitutionally-mandated system of intellectual property.

Conclusion

CCC strongly supports the continued vitality of the traditional peer-reviewed scholarly publishing system, with its wide variety of systems of distribution – including sales of copies, licensing of uses, and appropriate free-of-charge access as part of an overall system that ensures financial health for a system that has served science dissemination in the United States and world well. CCC encourages the Office of Science and Technology Policy to learn about the breadth and depth of alternative forms of access, and market-sensitivity to users and uses, that peer-reviewed scholarly publishing has developed in the United States and around the world and to take that into account in developing recommendations to the National Science and Technology Council for future government policy. CCC stands ready to be of assistance to the Office in any way possible.

CCC Contact Information:

Frederic Haber
Vice President and General Counsel
Copyright Clearance Center, Inc.
222 Rosewood Drive
Danvers, Massachusetts 01923
978-750-8400 telephone
978-750-4343 fax
fhaber@copyright.com

Response to Request for Information "Public Access to Peer-Reviewed Scholarly Publications from Federally Funded Research", November 2011

January 12, 2012

Clifford Lynch
Executive Director
Coalition for Networked Information
Cliff@cni.org

I am pleased to have the opportunity to submit comments to this request for information on "Public Access to Peer-Reviewed Scholarly Publications from Federally Funded Research" on behalf of the Coalition for Networked Information (CNI). CNI is a membership organization consisting of some 200 organizations, primarily but far from exclusively universities, who share a common commitment to advancing the intelligent use of information technology and digital content in support of scholarship. You can find more information on CNI at www.cni.org.

I want to be clear that while these comments are certainly informed by discussions with CNI's member organizations, they should not be viewed as representing the position of any specific member of CNI.

In general, CNI supports the analysis in the response to this call for comments already submitted by Prudence Adler on behalf of the Association of Research Libraries (ARL) on January 8, 2012, and available at <http://www.arl.org/pp/access/accessfunded/rfi-access-to-pubs-8jan2012.shtml>. It is clearly time to extend public-access policies to all federally funded research.

I want to supplement the ARL analysis with two additional points that speak to areas where CNI has focused some specific attention and expertise in recent years.

I believe that scholarly and scientific norms, as well as sound public policy, support the practice that the underlying data supporting published results need to be publicly available to facilitate replication and reproduction of those results. Also, their availability is important for additional scholarly analysis and re-use. These arguments have been widely presented in scholarly journals, governmental and scholarly policy reports, and studies by the National Academies. As journal articles and other forms of scholarly publication begin to move beyond the constraints of the historic printed page and exploit the affordances of the digital environment in which scholarly publications now exist, the boundaries between publications and underlying data will rapidly become much more fluid. A clear understanding about public access to publications will facilitate access to underlying data (as well as the understanding and reuse of this

data); similarly, barriers to public access to publications will create obstacles to public access to the underlying data.

My second additional point deals with the changing nature and use of the scholarly literature. For a number of reasons not just the size but the rate of growth of the scholarly literature is increasing steadily. It is increasingly unrealistic for unassisted human scholars to cope with this rate of growth; there's a new paper published every minute or two, every day of the week, every week of the year. We need to be able to apply information technology in more sophisticated ways to help scholars to deal with this flood of information; literally, to compute on the literature. As long as the vast majority of scholarly literature is scattered across an archipelago of proprietary, access-restricted silos, development and deployment of these computational tools to manage, navigate and mine the scholarly literature will be largely impossible. Public access to the publications from federally funded research - if access is appropriately defined to include these types of computational access - will substantially help in creating an environment that will facilitate the development and adoption of these computationally assisted discovery technologies, to the advantage of both scholarship and commercial exploitation of the body of scholarly knowledge.

I conclude with a few citations to work that explores these two points in more detail. Clifford A. Lynch, "Jim Gray's Fourth Paradigm and the Construction of the Scientific Record", *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Tony Hey, Stewart Tinsley, and Kirstin Tolle (Eds.), (Redmond, WA: Microsoft Research, 2009), pp. 177-183. Online at <http://research.microsoft.com>

Clifford A. Lynch, "The Shape of the Scientific Article in the Developing Cyberinfrastructure", *CT Watch* 3:3 (August 2007), pp. 5-11. Online at www.ctwatch.org.

Clifford A. Lynch, "Open Computation: Beyond Human-Reader-Centric Views of Scholarly Literatures", *Open Access: Key Strategic, Technical and Economic Aspects*, Neil Jacobs (Ed.), (Oxford UK: Chandos Publishing, 2006), pp. 185-193. Online at www.cni.org.

Request for Information on Public Access

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?

a. Growing Markets

The growth of existing markets and the development of new markets will both be accomplished most successfully by the provision of immediate, full access to and reuse of complete collections without commercial restrictions. This complete access will permit entrepreneurial members of the public to fully use these works to generate new services and products unencumbered by restrictions that might limit their innovative use. The sooner and more completely these works are available the more quickly and fully individuals and companies will be able to unlock their economic potential.

In particular, open works are more likely to be effectively commercialized because businesses will be able to operate without major restrictions that limit their creative application. Access based on limitations creates disincentives for use generally and may curtail the unexpected or inventive uses that a particular company might commercialize.

Where complete access is not made available companies will be forced to choose between a presumption of limited use or reliance on copyright exceptions such as fair use. As noted above, limited use reduces the ability of all users to pioneer groundbreaking applications and may make many potential investors hesitant to use works that carry such a limitation.

Though copyright exceptions play an important role in unlocking these works, statutory exceptions are generally aimed at existing communities and may not protect the innovative uses that supports the expansion of new and developing markets. Similarly, although fair use is a vital tool for users, those seeking to commercialize these works are likely to find the uncertainty inherent in this exception a significant deterrent to investment and entrepreneurship. Without

open access reliance on fair use would force a company into a legally uncertain position likely to make investors uncomfortable and force repeated legal analysis at significant cost.

Truly open public access also empowers more users and particularly new users to keep abreast of the latest trends in the research. Greater dissemination of the latest research can be expected to support greater innovation based on that research. Even in cases where the specific content is not applied in a market context expanding the base of knowledge will improve the intellectual marketplace and raise the quality of research across the board. Unexpected users in particular will be empowered to fully-engage with this content so as to discover commercial applications that would otherwise be missed.

Faster – ideally immediate – access is just as important. The sooner work is available to the public the more quickly all citizens will be able to apply ideas generated by the research thus leading to new products and services entering the marketplace more quickly.

Public access drives new industries and faster access facilitates new jobs across all segments of the economy. Knowledge-rich professions such as agricultural and biotechnical sciences, high tech professions such as energy, and information professions such as publishing all rely heavily on the sort of research at issue here and all of these professions are major source of sustainable, high-quality American jobs.

Access to this information will also incentivize private investment in technical solutions that build on government research – a traditional strength of the American economy that is badly needed today. We already have examples of IT infrastructure that aggregates and mines public information research such as Google Scholar and goPubMed. With full, immediate public access these companies can offer better services with increased commercial potential. New jobs and new companies can be developed to further capitalize on work that the government is already paying for.

b. Driver of Scientific Productivity

Scientific innovation and productivity also rely on open access to research funded by the federal government. We already have strong empirical evidence that open access research is read by more people. This means that open access research promotes more and faster follow-on

research as scientists use these works in innovative ways. Open access research also encourages a greater diversity of follow-on research as many minds explore different and unexpected angles, including research pathways that might otherwise have been missed. By letting all scientists incorporate the results of governmentally-funded research into their own work more quickly, open access will encourage faster, more thoughtful application of that research towards the next generation of scientific innovation.

Open access content is equally valuable for use by new scientific tools such as machine-readers and computational analysis. Fast access to all data is necessary for scientists to leverage these new tools so they can identify better content and scientific research can progress more quickly, and more intelligently.

Machine-reading in particular opens up entirely new scientific pathways, enabling the discovery of new connections across the body of research. This powerful new tool, however, relies on complete access to large bodies of data with no limitations. With complete open access new research pathways and semantic tools can be used to speed and transform scientific productivity. This, in turn, opens new avenues for commercial development that capitalizes on existing public investments.

Finally, open access permits unforeseen participants to join in the scientific enterprise. American history is filled with scientists and technological innovators who were not affiliated with established institutions and the recent rise of internet success stories in particular epitomizes the value of the unexpected innovator. Members of the public who might not otherwise have access to this research will be able to contribute in the tradition of amateur innovators such as Steve Wozniack, Bill Gates, and Mark Zuckerberg.

In the academy and the research laboratory, traditional scholars in related disciplines will also be empowered by open access to contribute to scientific progress across all disciplines. This paves the way for innovated interdisciplinary discoveries. It also increases the return on investment for all research since it will be able to be used across all contexts. By opening access to all citizens scientific progress can be driven by thinkers across disciplinary boundaries and beyond the walls of traditional scholarly institution to harness the American innovative spirit.

c. Costs and Benefits

i. Benefits: The benefits of open access have been demonstrated by several major governmental programs that are already in effect. The Houghton Reports on FRPAA make it clear that opening up access produced at least a fivefold increase in return on investment. The benefits of an open access policy similar to that of the existing NIH policy are estimated at approximately 8 times larger than the costs. The net present value gains of expanding an NIH-style policy to all other U.S science agencies is estimated to be on order of \$1.5 billion (net costs of running the archive). Of that figure, approximately 60% is estimated to accrue directly to the U.S. economy.

Open access provides the additional benefit of providing increased accountability for federal agencies. Outcomes of funded research will be easier to measure and Congressional budget drafters, appropriators, and authorizers will have better information to assess the value of existing expenditures and target funding on the most promising research. Policymakers will also have better information across the board based on the improved access and use of research.

ii. Cost: The NIH's open access policy provides a closely analogous example that illustrates the cost-effectiveness of open access policies. The NIH has proved cost-effective with between \$3.5 and \$4.6 million – or about 1/100th of 1 percent of the NIH's \$30 billion budget – providing access to better than 2.2 million articles. These articles are used by more than 500,000 users per day, most of whom come from outside of the traditional university environment. There is deep demand for this information across the public sector.

This use of NIH content underscores the cost-effectiveness of open access and provides an important base for expanding open access. By building on these existing programs existing infrastructure can be leveraged to avoid duplicative effort. This base can then be used to expand access to additional content at a minimal incremental cost.

d. Type of Access Needed

Free, immediate access that includes the right to reuse will have the greatest benefit for scientific progress, technological innovation and the American economy. Any restrictions on access to the material paid for by the public will limit the value of that information and significantly diminish the return on the public's investment. Full reuse will permit researchers to maximize the value of this work as well as unlocking additional value in the years to come. It

will also limit duplicative costs and build on the results of this research in sustainable ways that will continue to sustain scientific progress and commercial innovation for decades.

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?

Open access fits comfortably within the current copyright regime which balances the right of the public to use works and the intellectual property rights of the authors that create and the agencies that fund their work. Successful open access systems such as the NIH permit use of works based on existing copyright mechanisms such as fair use and the eventual entry of works into the public domain.

Along with these established copyright rules, greater utility should be enabled to permit use that supports the scientific and commercial innovation that public funding is designed to encourage. Mechanisms to enable full use of this material should be included in the policy so that users can engage in distribution, reuse, text mining, data mining, computation, and the creation of derivative works.

Adoption of a licensing system similar to the Create Commons “CC-BY” license will permit full use of this content by the public even in cases where existing copyright exceptions are not available. Licensing mechanisms such as the Creative Commons have been extraordinarily effective because they are simple to create and attach, easy for users to understand, and fit comfortable with the open use of content created to serve the public good.

An IP regime that balances the rights of all parties will best-serve the public. An embargo period will permit publishers to commercialize these works while users are able to rely on fair use for the comment and criticism needed to evaluate new articles. After this embargo period works should be available to the public subject to a standard CC-BY license that assures attribution but otherwise unchains works so that the public can maximize their value.

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 federal government is the appropriate entity to provide permanent stewardship of these articles, and is in a unique position to ensure that publicly funded articles are permanently preserved, made accessible, and useable. As such, any public access policy must give the government the rights to archive and distribute these works.

At a minimum, the government must maintain an accessible, mirrored version of all content so that the public can be assured of having access. We have numerous examples from other agencies such as the SEC and USPTO of the federal government maintaining large databases of information. The closest analogy, the NIH, has proven to be extremely cost-effective: NLM reports PMC costs less than 1/100th of one percent of NIH' s operating budget to run.

Distributions across multiple repositories is not a problem but all repositories must have the same conditions surrounding access and use to ensure genuine long-term storage and sufficient interoperability. Repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of articles, could be maintained by third parties. This would encourage innovative public/private partnerships and permit numerous companies to develop tools and search strategies that improve search efficiency much as companies such as Westlaw and Lexis generate millions of dollars every year by supporting access to legal documents in the public domain.

A “dark” archive that does not provide access to all parties is not an acceptable solution. Efforts to archive content must be measured in decades, not years, and library experiences have shown that regular access/use of digital materials is crucial element in effective long-term preservation. Without regular access/use, archival veracity cannot be ensured and public access may be limited by whatever institutions do make the content available if they push the boundaries of accessibility based on format, etc.

The federal government making this content available is not duplicative; it is necessary to ensure this public investment is protected and fully-leveraged. Current market attempts at archives are not adequate. For example, Cornell and Columbia report that only ~15% of their combined journal holdings are currently archived by LOCKSS and Portico combined.

Whether a centralized or decentralized model is chosen, all works must be made available to the public in such a way that preservation, access, and use are fully protected. The federal government has the infrastructure and the mandate to do this.

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 most successful models will be those that recognize all of the partners in the research enterprise. Publishers, libraries, intergovernmental organizations, scientific communities across national borders and of course scholars themselves all have a stake in this process. A narrow focus on “existing publishers” risks missing the important contributions of these other stakeholders as well as the invaluable advances made by the next generation of innovators.

Publishers – both established and forthcoming – can play an important role in this process by providing approved repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles. No single stakeholder, however, should be given a monopoly on these works financed with public dollars. Partnerships should permit multiple points of access for users and must be open to anyone in the marketplace who can improve on existing services or offer competing models for innovative use.

Public-private partnerships with academic stakeholders are another important opportunity that should not be ignored. Universities and libraries have extensive experience and existing archive infrastructure, and should be actively encouraged to partner with federal agencies. Empirically, None of the 50+ research funders who currently have public access policies are using publisher sites as the final archives. There are, however, good examples of funders partnering with academic and research institutions in this role.

We have several examples to draw on in this area. In Europe, the Digital Repository Infrastructure Vision for European Research (DRIVER) provides a test case for interoperation of both data network and knowledge repositories as integral parts of the E-infrastructure for research and education on a scale comparable to the United States.

At North Carolina State University we have had success with our own repository, as have our colleagues at partner institutions in the Research Triangle, an area where the broad dissemination of scientific information has led to internationally-praised innovation and substantially boosted the economy. We have also had success with larger partnerships such as the Hathi Trust and Open Library Initiative. As we have seen at NCSU, as well as through initiatives such as Mendeley, PLoS, and even Google, the most successful partnerships are those that recognize all stakeholders in the research enterprise, as well as the public good that they ultimately serve.

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?

Policy surrounding metadata should recognize that metadata is more than a simple description of an item, it is a means for enabling specific actions. As such, metadata should be designed to facilitate specific, desirable actions around use, reuse, and analysis of published works. To enable this use metadata should be machine-readable, particularly for use and reuse.

Creation of metadata should begin with existing standards. Standards such as Dublin Core for exchange, ORCID1/2 for controlled identifiers, and Counter/Sushi for usage tracking provide a base of established and tested models that can be evaluated and improved upon as time passes. Established metadata agencies such as NICO and LOC that have spent years developing

expertise on metadata interoperability should be involved in the ongoing development of new standards.

It is important that metadata be coupled with an API for standards-based data exchange. Published articles and data are distinct issues and metadata must be cognizant of these differences, particularly since articles can also be used as data. Metadata can be used to build bridges between these two through semantic relationships, unique identifiers, and similar coding. The most successful metadata will build on existing standards to enable the specific actions required to maximize access, use, and archiving of these important public resources.

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?

In order to minimize the costs open access policies must be based on consistency of requirements and mandates are essential across disciplines. Researchers often hold grants from multiple agencies and consistent policies will reduce inefficiencies for institutions and individuals. Specifically, open access policies should include uniform requirements for peer-reviewed literature, uniform deposit requirements that reduce complexity and cost. Uniformity can also be expected to increase compliance.

Maximizing the return on taxpayer investment can be accomplished through several principles. First, the policy should take advantage of existing protocols to make deposit in multiple repositories as efficient as possible. This can be accomplished with tools such as SWORD and additional tools that should be developed at the encouragement of the policy.

Articles should also be integrated with grants management systems. This will increase efficiency as well as agency accountability. Properly run, this open access system can be an important tool for providing better information to taxpayers about what they are getting for their investment.

Public access policies also offer an opportunity to enhance productivity management tools in the academy. Universities will be able to better measure research output. They can facilitate the creation of better bibliographies and PI tools, and universities/libraries to use repositories as teaching tools (i.e., teaching scholars more effective literature analytics, etc.

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?

All scholarly and educational materials created with taxpayer funding should be made readily available to the public. Comparable efficiencies support wide dissemination and similar benefits can be expected to accrue to the academic enterprise, scientific innovation and the American economy when these materials are made available.

It should be recognized, however, that different issues arise with different types of material. As such, policies for distinct materials may reflect the distinct nature of those materials. There are important differences between the ecology of journal articles, book chapters, and other educational material and these differences may require different policies. For example, text books are designed for a specific audience

Similarly, different types of educational material are created in the context of different existing models. As such, the policies for different materials may have to be adjusted so as to minimize disruption of those distinct models. The incentive structure for journal articles is built on reputation and prestige but not on financial rewards. Monographs, on the other hand, may be created with less concern about the reputation of the publisher and more focus on financial rewards.

Public access policies should be cognizant of the differences between different types of educational materials but not at the cost of core principles such as openness, access, and maximized efficiency of taxpayer 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?

Access that is complete and immediate best-serves the interests of the American public. The more quickly and readily works are made available the better-able all citizens will be to maximize the scientific and commercial potential of taxpayer-funded works.

If the decision is made to reduce the benefit to the public in order to support the current subscription model used by some academic publishers an author-defined embargo of no more than 12 months could represent an acceptable compromise. An embargo determined by the author of the work that runs between 6 and 12 months would permit publishers to commercialize works at the height of their value while still permitting relative quick access by the public so that these works can be used to grow the economy and drive innovation.

This 0-12 month embargo period has been used across most major disciplines with great success. It represents the norm for the industry and has been adopted by hundreds of journals. Despite concerns when embargoes were first adopted, no one has presented any data demonstrating that this policy has harmed publishers.

Indeed, early concerns about openness are increasingly being replaced by groups such as the Royal Society embracing open access. Royal Society, publisher of the world's first scholarly journal, recently opened access to their back file of articles with a 12 month embargo period, noting that this prestigious and heavily cited back file, dating back to 1665, accounted for less one half of one percent of their overall publishing revenue.

If an embargo is employed then calculation of the effect of the embargo must consider all factors. The assumption that access – embargoed or otherwise – reduces profits for publishers cannot be accepted uncritically. Numerous market conditions interact to generate effects in subscription rates. Growth of journals and papers in disciplines, the price – and pricing history - of a given journal and of competitive titles, the potential impact of required bundles, larger

library budget numbers and trends, and real revenue resulting from “long-tail” business all play an important role.

All of these market conditions regularly contribute to journal cancellations and must be accounted for so that effect of embargo period can be adequately isolated. The most successful embargos will be brief – lasting only as long as a critical evaluation of all market factors can justify – and in line with the established model that does not exceed 12 months.

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

Response to RFI

1. What steps can be taken to maximize the benefit of Government sponsored research in growing the economy?

The greatest benefit from government sponsored research would come from the widest public dissemination of that research product to all interested stakeholders at the lowest cost. The value of research is often not known until that research is combined with other research. The sum of multiple research products is more valuable than each one separately, and that extra value cannot be appreciated until all the different pieces are put together.

An excellent example of such a program that allows multiple types of research products to be accessed is arXiv.org. Extending the benefits of something like arXiv.org to all Government sponsored research, including book chapters would be beneficial.

2. What should be done to protect intellectual property interests?

Science only progresses by building on the work that has been done before. Proper citation is essential to properly credit scientists for their intellectual efforts. Existing patent law provides for protection of intellectual property interest in inventions. Patentability of inventions requires that the invention be new, useful and not obvious. Unfortunately there are patents of dubious novelty which seem obvious and which are stifling progress. Gene sequences identified in nature are not novel and should not be patentable.

It is necessary to balance the costs and benefits to various stakeholders. Copy right is granted to acknowledge the generation of unique materials and to allow wide dissemination while the originator retains ownership and can profit from it.

6, 7. To the extent possible, greater public dissemination will only produce greater public benefit.

8. The shorter the embargo time the better. To the extent that a for profit journal need to cover their costs in producing a specific article, once the cost of producing the article have been covered through subscriptions and single article purchase, the content should become open access. In other words, some journals do have an open access option, where the authors can make the article open access by paying a fee. Once an article has achieved revenues equal to the open access fee, the article should become open access.

An appropriate way to do this would be to remove copy right and impose an open access-type free unlimited use provided proper citation is used.

Thank you

David R. Whitlock

To whom it may concern,

I recently heard that there is currently a bill to stop open access to public, tax-payer funded research (H.R. 3699 Research Works Act: <http://thomas.loc.gov/cgi-bin/query/z?c112:H.R.3699>). As a recent graduate, and researcher, I think this bill should not pass for several reasons:

1. American taxpayers already funded this research, so they should get access to the results. Requiring taxpayers to pay for access to publicly-funded research articles is akin to asking someone to pay for groceries, then pay for being able to eat them. The latter would be unfair .
2. The inability to access cutting edge research means that the research will not be disseminated as easily. This can hinder further research in the U.S. Researchers who support open-access research (which, now seems like many researchers) will publish in journals that are abroad that have lower or no publication costs. This diminishes the research published within the U.S.
3. Doctors and patients need access to medical research to help cure diseases and improve the health of patients. Adding an access cost for patients and doctors creates a barrier to improving health care.

I urge you to not pass HR 369: Research Works Act.

Thank you,
Girish Tembe

Dear National Science and Technology Council's Task Force on Public Access to Scholarly Publications,
As a medical librarian at a public institution, here are my comments on your RFI for Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research.

- 1) Agencies should make all federally funded research publications mandatorily available to the public. The amount of information published makes it almost impossible to find the information that you need in order to conduct sound research. Because of commercial ownership on published articles, computation analysis to help “get to” the most appropriate information is almost impossible because programmers do not have a corpus of literature from which to text mine effectively, thus impeding scientific discovery. The cost of storage of these digital articles may be a barrier but is minimal in comparison to storing the great amounts of raw data that may be generated from activity such as gene sequence analysis or astronomy tools. In the least, a comparable database like PubMed Central that incorporates all of the subject disciplines with full-text articles would be extremely beneficial in more efficient and rapid scientific discovery and analysis. NIH has been a good role model for the biomedical sciences, but research in incredibly interdisciplinary and for true transformational research to occur, you cannot silo information into specific subject areas. Having a single source of federally funded research publications makes the most scientific sense.
- 2) I can appreciate the commercial need to create and sustain journals, however, the research that is being published has already been funded by tax payer dollars. An embargo period can continue to benefit private publishers without hindering the progress of good science. As a librarian, I am particularly angered by the incredible profit margins of these publishers (we have seen percent increases in a small society published journal that got acquired by a commercial publisher as high as 276%). Information is the single most empowering commodity in the world, and people have a right to it.
- 3) Pros – Data mining across literature of all subject disciplines will help to progress scientific discovery from small lab environments into commercial and public good. A single individual cannot retain all of the knowledge published that would necessarily lead them to new discovery. Text mining is an obvious immediate benefit of an open source, full-text, scholarly publications database. If content is distributed, then there would in the least need to be standards for metadata, archiving, back up and disaster recovery that the government would set. If the government pursues a distributed path, then I think it would be necessary to have “accepted” third-party providers so that the standards could be monitored and enforced.
- 4) Nature Genetics, a journal from the Nature Publishing Group, has standards for distributed storage of scientific data that could be looked at. The University of Michigan has several projects that demonstrate commercial publisher cooperation with accessibility of published materials such as their Deep Blue institutional repository and the Hathi Trust, a cooperative digital works repository.
- 5) Metadata standards such as those for Medline (PubMed) are a good starting point and requiring that all future requests for federal funding be tied into publications that meet these standards (such as the PMCID requirement by NIH) are a good way of ensuring compliance.
- 6) Tying compliance into the existing award structure for federal funds would be the most efficient method and probably least disruptive method for ensuring open standards. Publishers still have many options of providing the scholarly articles in a manner that can still be profitable but still do not impede the access to the information – more quickly (during the embargo period), via mobile applications (different formats that users may want and can afford to do so)
- 7) Ideally, all information that is federally funded should be made available since the research was funded by the public. However, the published, peer-reviewed article is a good starting point.

- 8) I do not have good evidence to support data for a specific embargo period. Existing citation metrics typically need a one year period in which to gather enough citation data to develop metrics such as journal impact factor, etc. This type of information may support a one year embargo period.

Jean Song
Research and Informatics Librarian
Taubman Health Sciences Library
University of Michigan

January 10, 2012

Chatham University Graduate Student Assembly

c/o Chatham University

Woodland Road

Pittsburgh, PA 15232

To Whom It May Concern:

We, the Graduate Student Assembly of Chatham University in Pittsburgh, PA, are writing to you in support of legislation for Request for Information (RFI) on Public Access. As an institution with undergraduate, graduate, and post-graduate studies in various realms of academia, it is pertinent that our students have access to the most current research available to them; especially when that research and the products thereof are publicly funded by their and their families' tax dollars.

Within our institution, nearly every department would benefit from the more efficient and more available information that this legislation will provide. In the sciences alone, Open Access will drive productivity, lead to an increase of follow-up research, provide innovations for new avenues of research, allow students to incorporate up-to-date findings in their current research, prevent overlaps and/or repetitiveness in research, as well as keep our student competitive in the expanding market of scientific research. In addition, it is necessary that this access to current information come with the legal rights to reuse the information with proper citations documented.

Restrictions that limit the use of this research information limit the value of this information to the citizens who paid for it to be completed in the first place. As such, enabling full reuse will mean that students can do more with less; research will not have to be duplicated in order to build on the results that have already been validly documented, extracting value from a single research-investment for years to come.

Thank you for your time and consideration on this matter.

Signed,

The Chatham University Graduate Student Assembly

University of Colorado at Boulder
School of Education
249 UCB
Boulder, CO 80309-0249

January 12, 2012

To: Office of Science and Technology Policy Executive Office of the President
725 17th Street Room 5228
Washington, DC 2050

From: Raymond C. Johnson, Doctoral Student in Mathematics Education School of Education, University of Colorado at Boulder

Re: Response to the White House RFI on OA publications

I am a researcher, concerned citizen, and a supporter of open, public access to publicly-funded research. I speak for myself and not on behalf of my colleagues or my institution, although I believe I express ideas and opinions shared by many researchers and educators.

In response to the White House Office of Science and Technology Policy request for information on "Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research," I urge you to preserve policies that require public access (such as from the National Institutes of Health) and expand similar policies to other federal funding agencies such as the National Science Foundation, a key source of funding for education research in mathematics, science, and technology. Currently it is with great jealousy I see the growth of open access publishing in areas such as health and medicine; as an education researcher I wish I had the ability to share the latest research with teachers and administrators, most of whom cannot afford the high fees charged by publishers of education research. Unfortunately, open access journals in education are relatively rare and undervalued. A change in policy, one that would require public access to federally-funded research, would quickly change the perceived valuation of open access publishing outlets and bring much-needed information to educators everywhere.

Prior to my becoming a researcher I was a high school mathematics teacher for six years in high poverty, rural Colorado school districts. I did not have the benefit of a nearby university or a district support staff with access to recent or prominent research. My main link to information was a powerful one: the internet. However, it seemed that my searches for research about teaching methods, curriculum, education policy implementation, etc., all eventually led me to paywalls put up by publishers to "protect" their work, requesting fees I could not afford to pay. Now, as a researcher, I realize that the authors of education research -- much of it funded with federal dollars -- are asked to give their copyrights to publishers in exchange for so-called "widest possible dissemination" of that research. Researchers neither receive nor expect any pay or rewards for giving away their work, other than some scholarly esteem and the hope their research somehow reaches and benefits students and educators. While that publishing model might have made sense twenty years ago, it does not any more. Any claim of "widest possible dissemination" that currently does not include searchable, full-text publication on the public internet is false, at best, and fraudulent, at worst.

In response to the eight questions in the RFI, I encourage you to consider the arguments and recommendations made by Harvard University in their response (<http://osc.hul.harvard.edu/stp-rfi->

[response-january-2012](#)). Their expertise in these matters far exceeds mine. However, I do wish to make the following amendments to their responses for questions (2) and (7):

(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?

Harvard's response refers to a need to divide and share rights between researchers and publishers. My recommendation beyond their statement is that any discussion of copyright include Creative Commons (<http://creativecommons.org/>), an organization dedicated to creating and defending content licenses that allow creators to reserve some, but not all, of their copyrights. Their expertise should be invaluable in any discussion about the sharing of intellectual property rights.

The Harvard response includes a recommendation of a Creative Commons license at the end of their response to question 1. I also urge you to consider the expertise of SPARC (<http://www.arl.org/sparc/>), the Scholarly Publishing & Academic Resources Coalition.

(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?

In Harvard's response, they say they "could support mandatory public access" for non-journal works, but consider these to be "secondary issues" and are "not prepared to list all the types of content to which a federal public-access policy ought to apply." I worry that this position is short-sighted and leaves too much room to abuse public access policies. Often the events that lead to research becoming a book chapter instead of a journal article are entirely matters of circumstance, and not a basis of quality or public importance. In fact, the entire distinction between article and chapter relates to a paper-based publishing economy, one that is increasingly irrelevant in a digital age. After all, if we were still limited to publishing on paper it is unlikely that this kind of public access policy discussion would even exist. If the spirit of these policies is to give the public access to research they have funded through federal tax dollars, there is no need to worry about "types of content" other than to say the research will consist of bytes and files traveling the internet. Furthermore, if the policy only requires "journal articles" to be published openly, what is to keep publishers from re-branding themselves as something other than a journal? By relabeling their products as books, magazines, or something entirely new, unwanted loopholes around public access are sure to emerge.

(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?

To maximize use of publications stemming from agency funded research requires their access from the public domain as rapidly as possible.

This provides innovators and knowledge users the greatest potential to translate research results into tangible economic benefits. Restricting access through, for example, commercial publication companies has the unfortunate side effect of limiting knowledge translation.

(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?

It is not clear what the actual intellectual property rights of publishers are. While they force copyright transference from authors, there is no apparent 'value added'. Current practices that are already in place to protect the rights of scientists, research institutions and federal agencies generally seem sufficient.

(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?

Pros: Library funds that are otherwise diverted to commercial publishing houses (which as noted above do not in themselves offer any additional value to the research), may instead be used to both maintain open access to published research as well as develop innovative new tools to access the research and would require a fraction of the costs that are currently being diverted into the publishing houses. It is not clear how the government can ensure long term stewardship if content is distributed across multiple private sources, some of which do not survive beyond a few years.

(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?

I am not aware of any

(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?

Full open text access together with supporting metadata is absolutely a minimum.

(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?

Take out the middle men - the publishing houses, which provide no value-added, restrict access and charge significant fees to US taxpayers through library subscriptions.

(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?

Yes.

(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?

This question does not appear to be based on a rationale argument - what purpose does an embargo serve except to stifle entrepreneurship and innovation?

Sincerely

John Parkinson Ph.D

--

Dr. John Parkinson

CIHR New Investigator

Senior Scientist - Molecular Structure and Function Hospital for Sick Children Associate Professor -

Depts. Biochemistry & Molecular Genetics University of Toronto Toronto Medical Discovery Tower (East)

OSTP Nov. 3, 2011 RFI on Public Access
Response from the
Association of Public and Land-grant Universities, APLU
R. Michael Tanner, Chief Academic Officer

On behalf of the Association of Public and Land-grant Universities (APLU), I write to reaffirm our support of providing public access to the results of research funded by the federal government and published in scholarly journals. This statement echoes many points the APLU sent in response to the OSTP's 2010 RFI on public access. APLU's endorsement of public access at that time was based on our polling of the Association's Board and of all the Provosts and Research Officers at our member universities. The role of our member universities in both the generation and the dissemination of new knowledge gives us a balanced perspective on the importance of publication and the desirability of ready public access to new research results. As we stated last year, timely and convenient public access to the fruits of federally-funded research benefits scholars and researchers, businesses, and our present and future students, and it enhances the vitality of intellectual inquiry generally. The intent of Article I, Section 8, of the U.S. Constitution was to "promote the Progress of Science and useful Arts," and how federal policy can best do that within the context of the Internet and search engines of today requires careful examination and weighing of the impacts of public access policies and the attributes of various models for dissemination, access, and preservation of the scholarly record.

Background of the Association-

The Association of Public and Land-grant Universities is a research and advocacy organization of public research universities, land-grant institutions, and state university systems. Our 217 members enroll more than 4.7 million students, award 60 percent of U.S. doctoral degrees and conduct nearly two-thirds of all federally-funded academic research, totaling more than \$34 billion annually.

We turn now to the specific questions of the RFI:

(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 APLU advocates broad and convenient public access at a reasonable cost, and at low or no cost as electronic media make that economically practicable. Paradoxically, the migration from print to electronic form has often reduced access. Electronic access to scholarly journals has been restricted largely to those who are members of the university community for which the electronic journals are licensed. Universities that once could lend copies of journals to the general public or permit them to have photocopies through inter-library loan, can no longer do so. Thus the continuing migration of the scholarly literature to electronic form reduces its lawful availability to the public. Small businesses and start-up companies need access to scholarly literature and the latest technological developments; for a budding enterprise, the cost and time required to negotiate licenses can be a big barrier to lawful access.

Frequently small business startups have as their principals or employees individuals with recent experience in university graduate programs and/or in research. Thus, they know the value that access scholarly record might hold for the success of their startups. Unfortunately, such businesses tend to be thinly capitalized and cannot afford expensive journal subscriptions. Great economic value can be created if a government-wide public access program that provides such access is put in place.

(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?

The patentable intellectual property resulting from federal research can be protected through provisions of U.S. patent law and international patent treaties. Nothing in the policies concerning peer-reviewed scholarly publications should conflict with that protection. Copyright protection on scholarly content has to be maintained for the purpose of assuring proper attribution and for capturing reasonable revenue flows to cover the costs of

publication production and distribution, recognizing that production and distribution have been greatly simplified by electronic tools and media. Publishers historically add to and enhance content through editorials, indexing, layout, copyediting, organization of material, etc., that must be respected and given due recognition. Historically a scholarly journal has bundled the scholarly content and the publisher's contributions, with copyright transferred by the author and held by the publisher, but achieving fairness to all contributing stakeholders calls for unbundling these elements conceptually.

(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 brief, a centralized approach can have advantages in terms of consistency of policy and practice while being correspondingly prone to disadvantages like rigidity and stultifying bureaucracy. Centralized approaches in general are vulnerable to failures when the "center" represents a concentration that can be subject to accident or even maliciously attacked at one physical location (e.g., the 1814 burning of the Library of Congress by the British). Also, a centralized approach may have a pervasive weakness in its uniform methods. Decentralized approaches can have greater resiliency and robustness and exhibit the positive attributes of diversity. At the same time, that diversity may pose challenges for consistency and interoperability.

With careful systems design and thorough execution, either approach can be workable. Clearly, wherever and however the material is stored, it must reside in multiple repositories in diverse geographic locations and otherwise protected against loss of data. For scholarly publications, a federal agency may reasonably keep custody of all published content for the sake of assuring long-term stewardship and beneficial redundancy, even if primary high-volume access to published content takes place by other means. Assuming reliable Internet connectivity, both centralized and decentralized models can be implemented without the content-searcher readily perceiving a difference.

(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?

We do not know of such models but, if publisher archives become part of a public access system, those archives must be compelled to adhere to the standards that characterize the maintenance of university archives. Such standards include but are not limited to 1) guarantees that material placed in such archives generally will not be removed or modified, 2) that access will be made available to all on nondiscriminatory terms, 3) that such archives will be actively linked to other public access archives such that unitary searches can be done, i.e. so that the location of the material in a privately held archive is not material to the conduct of searches and 4) that a mechanism will be put in place to ensure that all the material in the archive will be conveyed to a trusted successor organization should the private organization be unable to or choose not to maintain the archive.

(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?

(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?

Ease of compliance is crucial to achieving public access policy compliance. Faculty member authors can be motivated to make good faith efforts to comply, but the rate of compliance will be lower if a busy faculty member is expected to master intricate requirements and different posting protocols to publish research sponsored by different federal agencies. For all members of the stakeholder community, there is a significant advantage to simplicity of concept and consistency of approach. We suggest that, to the extent practicable, uniform requirements and procedures regarding deposit of papers be established across all funding agencies covered with, for example, the length of embargo period as parameter that may vary from field to field.

Consistent deposit protocols will reduce cost and complexity while increasing the rate of compliance.

Ease of access is similarly crucial for the user community. As information technology tools have leaped forward, the ability of search engines such as Google Scholar to digest, catalogue, and cache content extracts essential for searching is very important. A central electronic access point, such as PubMed Central, can offer specialized search capabilities for those who know to use them. For the public at large, immediacy of access is dependent on access to the published material, in whatever repository, by web crawlers, combined with legibility of the material with standard Internet browsers. To facilitate accuracy in scholarly references, it is far preferable to have the searchable and accessible text be the final published document, not a pre-print or the submitted manuscript.

The NIH public access model represents a balance between competing interests: researchers want timely access, and publishers want sufficient control and revenues from access to support a high quality production and distribution process. The NIH model has balanced these well, and it has proven very popular with our member universities.

Some propose that adoption of the NIH model by other federal agencies should allow for an embargo period that varies according the field. This will be taken up further under our response to question (8). The principle of simplicity and ease of compliance suggests that the embargo periods might have at most several tiers, but not unlimited variability. The embargo period should be consistent across similar types of publications within a given field and across similar fields whenever possible.

(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 principle that federally-funded research publications should be publicly accessible is independent of mode of publication, but pragmatically, there may be good reason to allow some differences in requirements. Books are generally considered a vehicle for publishing material that has a more lasting value and for which the publisher might anticipate sales to individuals over a number of years. The economics of the book market and the costs of production may be a rationale for a longer period of embargo, for example.

Even peer-refereed conference proceedings can vary significantly in the standards expected for quality and novelty of accepted contributions. While proceedings from prestigious conferences may have a standing comparable to or exceeding that of journal publications, other conference proceedings are topical offerings with much lower likelihood of future citations.

In the electronic era, the lines of demarcation between books, journals, and conference proceedings are blurring. A series of eBooks in a field could be very similar to a journal in impact. Conference proceedings may be published on the Web and have as large a readership as a journal. When there is little distinction among modes, the argument for a consistent public access policy is strong.

(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?

We are unaware of any journals whose financial viability has been significantly damaged by the NIH Public Access requirement and its 12-month embargo. The experience of the libraries of APLU members over the last two decades has been that of journal prices rising at well above the inflation rate, particularly in the sciences and engineering fields, a proliferation of journals, and publisher pricing to libraries for electronic access that involve bundling of groups of journals. Without doubt, some scholarly societies have relied on positive net revenues from journals to provide subsidies to other activities that have benefits for their membership. To the extent a short embargo demonstrably leads to the loss of subscriptions and associated revenue, those affected will argue that a longer embargo is necessary for the viability of their enterprise. To this point we have not been alerted to the extreme distress or disappearance of a journal critical to a field due to the current NIH policy.

It is important to recognize that information technology, and its ability to propagate information at close to zero incremental cost, is a disruptive technology in many “publication” businesses, and scholarly communication is not immune. The conventional distribution network of “printed” material is losing or lost ground rapidly to new Internet models in, for example, the

recorded music industry and the newspaper industry. Those industries have been forced to experiment with new business models. Losing consumers and subscribers, they have had to test new distribution ideas and seek new revenue sources. In the case of newspapers, they have looked to better targeted advertising and broader subscriber bases at lower costs to replace lost print advertising revenues. For those industries the trend has been painful, visibly damaging, but obviously unstoppable.

Included in the business analysis of the new order has to be the effectiveness and cost of enforcing copyright protection. There is no good rationale for a publishing framework that simply forces institutions that respect the law to carry the full economic burden of sustaining an outmoded business model. For scholarly communication, universities and university libraries have been painted into that corner. Increasingly scholars in various areas have embraced an “open access” model as a way to capitalize on information technology to improve the availability of their own work and make it more easily found by interested parties.

High-quality scholarly publication must continue to thrive. Whether through author fees for publication or subscriber fees for access, revenue is needed to sustain peer-review and triage, careful composition, editorial oversight, reliable distribution and access, and long-term preservation. The embargo period of the NIH policy allows the value of copyright to act during the initial period post-publication, when the demand for access is highest in science and engineering fields, to sustain the subscriber model.

The APLU has not done nor gathered results of studies that would provide empirical evidence for the appropriate length of the embargo period in different fields. We do believe that informative studies could be done and may already have been done in some fields. Publishers and libraries that have been providing primarily electronic access to journals probably have data or could collect data on the rate of access to articles, from the time they first become available and over a period of many months or years thereafter. Access activity that is strongly loaded into the first several months and drops off markedly thereafter is strong evidence that an embargo that extends, say, six months after the drop off should suffice. The immediacy of interest in the first months, during the embargo, will sustain subscriptions. If there is little drop off in access in a field, that would suggest that a longer period of embargo may be justified. Such studies would be valuable in establishing

suitable embargo tiers, as suggested above, for disciplinary fields that would be affected by an expanding public access policy.

Thank you for this opportunity to address your questions.

Thu 1/12/2012 5:23 PM

Gretchen Weibert

Free access to published results from NIH

As a 9 year survivor/warrior of IIC Ovarian Cancer I find it mandatory that I voice my concerns regarding the loss of free access to published results for NIH studies. Beyond the point that the NIH is a federally funded program, the ability for me, a "layperson" with a terminal illness to gain as much knowledge of my disease as i can helps me make decisions about my care and options for care. At minimal, the results should be available for reading online and downloading. That saves printing cost to the government, should that be concern. Research results should never be kept for only those who can pay for them. I understand proprietary concerns might be a concern, but when funding and support is paid, in full or in part by federal monies, then I think the results should be made available, something akin to Sunshine Laws.

G Weibert

Sent from my iPad

Here is my input to the RFI on Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research.

Igor Carron, Ph.D.

Executive Summary

In the past few years, we have seen a perfect storm mixing the growth of two phenomena, a data deluge stemming from access to cheap sensing and computational equipment and the growth of scholarly publications. At the same time, there has been a near constant supply of reviewers. Open access to government funded work is the only short- and long-term policy decision that quickly enables a larger pool of quality reviewing capability aside from imposing reproducible research standards. In the end, it enables a more robust scientific process.

Introduction

With the advent of cheap high throughput equipment, we are seeing the emergence of what one would call "the curse of dimensionality", .i.e. the ability to produce cheaply large amount of data and the somewhat still limited ability to makes sense of them. This data deluge is, in turn, the primary reason behind the growth of the number of scholarly journals and journal articles over the recent few years. Unfortunately, the pool of potential reviewers has remained about the same and has not caught up to the level needed to deal with these two growth factors. One can certainly wonder how this is having an impact on how Science is performed (i.e. judged). In particular, the growth of the number of journals has eventually yielded a reliance on a lower number of potential high quality reviewers per journals. More insidiously, the growth in data production and/or computational experiments has removed from most time-constrained reviewers the physical ability to take on real reviews.

Peer-Review

In light of this situation, the current response by non-profit and commercial publishing entities has been to exacerbate the problem by opening the gate for newer journals and conference venues instead of developing innovative processes to do the one function that is generally thought to be their value added to the process of scientific discovery: The management of the peer-review process. An item of considerable interest is the current lukewarm ability by publishers (commercial or non-profit) to deal pro-actively and fairly with retraction. In particular, there is currently no system in place for publications to address the fact that they may have referenced a recently retracted publication for instance.

Under a regime of government funded open access of publications, new or older players could change the way peer review is performed by enabling systems like a **post**-peer review capability. This is just an example but innovation has to enter this market in order for the different stakeholders to continue on producing high quality work, at the lowest price to the government.

Conclusions

The interest of the US Government to have open access of government funded work can be clearly delineated into the following reasons:

- Open access opens the ability of non-time constrained post peer-review processes by a larger pool of reviewers, thereby enabling a more robust scientific discovery process.
- Open access provides the ability for innovation in the marketplace by enabling new (commercial or non-profit) actors in the peer review process. The new players may provide the ability to create new opportunities that are currently seldom explored by the current landscape.
- Open access potentially reduces some large cost to the government in its ability to deal effectively with past flawed work and attendant retractions. Some of these retracted works may have had broad policy implications.
- Open access comforts the United States leadership in manners related to Science and Technology development.

Jan 12, 2012

Andrea Quintero.

University of California Davis

Davis, CA

Dear Mr Wackler

I will keep my comments brief and direct them to 4 of questions listed.

(1)

- The bill HR. 3699 as written contradicts the researched and justified recommendations of the Task Force on Public Access to Scholarly Publications. As such, this bill cannot be supported. Additionally, the NIH call to have all publicly-funded work available to the public must be enforced. And this responsibility should be on the shoulders of the private-sectors publishers, who knowingly accept, and profit from, this work for publication.
- Federally-funded work must survive a rigorous process in which must justify not only its scientific merit, but also how it benefits the public. If the results of this work are kept from the public, then how are we benefitting from the work?
- The proportion of citizens that can afford access to privately held publications, or have access to a university that can afford the cost, is relatively small. Innovation and discovery happen through the development of ideas mixing with creativity. Restricting access denies possible problem solvers from gaining the necessary information to push innovation and discovery forward

(2) The idea that private-sector publishers add unique value cannot be accepted as truth. Few publishers edit and review work submitted to them. Instead they enlist the expertise of researchers who decided if the work is appropriate for publication and suggest revisions if necessary. This is done without pay and generally anonymously. While the publishers should reserve the right to control the publications for a period of time, federally-funded work cannot be allowed to be kept from the source of the funding.

(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?

(4) Public Library of Science (PLOS)

Thank you for your attention,

Andrea I. Quintero

PhD Candidate

Neuroscience Graduate Group.

Cognitive Analysis & Brain Imaging Laboratory,

M.I.N.D. Institute.

University of California Davis

[Assigned ID #]

[Assigned Entry date]

Name/Email

Morteza Gharib, Vice [Provost](#)

Affiliation/Organization

California Institute of Technology

City, State

Pasadena, CA 91125

Caltech is a PhD university employing 922 principal investigators whose research funding comes largely from 6-10 different federal agencies. In addition, Caltech is committed to education and recognizes its profound obligation toward public dissemination of its research results ideally unfettered by the demands of commercial profit so that learning and discovery, two major pillars of the enterprise, will thrive. The global network provides the means to ensure maximum access for uptake of new knowledge via electronic distribution of publicly funded research results. Therefore, Caltech urges and supports action to require prompt public access to results of all government funded research.

In response to the request for information from your office released on November 3, 2011 on the topic of public access to peer-reviewed scholarly publications resulting from federally funded research we offer the following comments.

922 includes professorial faculty, research faculty and postdoctoral scholars.
The Dept. of Defense is counted as one agency.

Comment 1:

First, new markets frequently result from innovations as outcomes of research. Since the work of the economist, Dr. Edwin Mansfield in the mid to late 20th century, the synergistic relationship between research, industrial innovation, and commerce or social return has been well-established. Now the global network has created a communication revolution in which immediacy and reusability are significant variables in productivity or social return on research investment. Therefore more current economic research has delved into the different methods or business models for distributing peer-reviewed research papers, the primary vehicle for communicating reliable research results.

Dr. John Houghton of the Centre of Strategic Economic Studies at Victoria University has conducted a number of such studies for the U.K., as well as for Australia, Denmark, and even the U.S., showing that unfettered access to the results of publicly funded research have a significant positive economic impact for a country as a whole. See:

<http://www.cfses.com/projects/knowledge-access.htm> . In the U.K these results are taken seriously enough that government policy is actively shifting to ensure that the public (all individuals, all learning, all research, and all commercial entities regardless of size or means) can actively benefit on their own terms from government funded research. See: UK Department for Business Innovation and Skills report, <http://www.bis.gov.uk/assets/biscore/innovation/docs/i/11-1387-innovation-and-research-strategy-for-growth.pdf>

The U.S. should be concerned that small companies and entrepreneurs experience significant barriers to federally funded research results. Eight new companies are launched each year, based on Caltech intellectual property, yet the staffs of these small enterprises cannot readily access research papers from their work place as purpose and needs dictate. This circumstance results in inefficient use of time and unreasonable hurdles to staying abreast of developments since information seeking and use cannot be seamlessly integrated into the work.

Opening up access to the research papers will level the playing field to allow new approaches, new companies to provide value-add services. With accessible, reusable, digital access to publicly funded research papers, new businesses can compete by offering computational tools for data mining, subject or linguistic analysis, bibliometrics, indexing, alert services, and more. The more ways that information from research results can be utilized, the more productive will be society. The demand for unfettered access to research papers in order to create value-add services using network based technology is illustrated in the following 14 services (and there are more) that have sprung up in just the last few years: Academia.edu, Epernicus | Network, Google, Scholar Citations, iamResearcher, JournalFire, Laboratree, Mendeley, Microsoft Academic Search, Nature Network, PeerEvaluation, ResearcherID, ResearchGate, Researchr.

The federal government has a number of successful models in which a basic level of formatted information is made publicly available (census data, patents, Securities and Exchange Commission information), that and other third parties (e.g. Economagic, Derwent, Intellectual Property Network, Morningstar Document Research) reuse and augment according to diverse business models. In this way the government is not responsible for generating all possible uses and presentations. Instead the private sector performs those services and competes for clients.

Secondly, optimal scientific research productivity requires immediate and full public access to all government funded research results so that curious and creative minds, no matter where they are located, have opportunity to build on that work in whatever way is useful and productive. Caltech research applies first principles to problem areas of medicine, energy, and the environment, among many fields, - ultimately all critical areas of social need. Specific examples include: drug delivery systems, bio-inspired fluid flow systems, design of prostheses and cornea implants, heart pumps, signal processing, and communication in regard to intercellular signaling, mechanics of cell scattering leading to metastasis using a digital volume correlation, insect wood digesting enzymes, multiscale models for large-scale engineering on the scale of earthquakes, green IT, photovoltaics, multi-junction semiconductor composites for water purification, cheap catalysts for solar power, artificial photosynthesis, and more. These interdisciplinary research problems involve a broad range of federal agencies are involved. For the scientific research community to be optimally productive, access to all the research results from all the agencies is necessary.

Comment 2:

The current copyright framework allows licensing as needed to be flexible in meeting the needs of the creator, the funder, and those entities that provide services. This flexibility is exemplified in the Creative Commons Attribution (CC-BY) licenses that can be imbedded with XML tagging into documents. CC-BY is a copyright license that grants permission to the public to reproduce,

distribute, perform, display or adapt the licensed materials for any purpose so long as the user gives attribution to the author. See: <https://creativecommons.org/licenses/by/3.0/>

Thus, the rights travel with the object, thereby encouraging use and greater, faster impact because the rights are immediately clear. There is no delay for special permissions for uses that further learning and discovery as is the predominant intent of authors of research papers. The efficacy of the CC licenses vis-à-vis the copyright law has been upheld in the courts, and these licenses are enforceable under current law. See http://wiki.creativecommons.org/Case_Law/ It is entirely appropriate that publicly funded government agencies assert at the outset certain rights to research papers on behalf of the public. The NIH 2008 policy operates on that basis and has been successful. The one shortcoming that needs correcting is the restriction on further public use.

Copyright was intended in the U.S. Constitution to “To promote the Progress of Science and useful Arts...” It is time to “shift from a model that uses copyright to control reuse of content to one that uses copyright to encourage republication, preservation, and translation.” (Carroll MW , 2011 Why Full Open Access Matters. *PLoS Biol* 9(11): <http://www.plosbiology.org/e1001210>, [doi:10.1371/journal.pbio.1001210](https://doi.org/10.1371/journal.pbio.1001210)) . In short an IP model that maximizes the dissemination and uptake of a researcher’s output is in the best interest of the research authors, the originators of the content, and the public. In contrast to the author of a book, who may anticipate royalties on sales, those of research papers seek to make their work available to others. Their compensation comes in the form of recognition for their contribution to the field.

Comment 3:

There will need to be flexibility. We already see that a centralized repository such as PubMedCentral can work and yet there is also a role and purpose for institutional repositories. The main issue is that the administration of those reference repositories is committed to archiving, preservation and unfettered dissemination.

A centralized repository whose content can be downloaded may enable many analytic or discovery tools specific to a research field or project. WormBase at Caltech is a biological database that extracts information from scholarly publications and puts it into computable form. Some of the steps in this process are automatable using machine learning methods applied to article texts., for example, identifying articles worth reading by the professional curators. While curators can obtain through site license or purchase articles with key information, it is not practical to obtain all biomedical articles and then analyze their text to decide if it is worth human effort. A centralized repository of all scholarly articles that can be accessed electronically is crucial to allow efficient use of human effort.

No entity (private, for-profit, public, non-profit) should be hindered from harvesting and re-using scholarly output with attribution. Universities should be able to position their output for maximum distribution and measurable impact.

Comment 4:

Publishers are not cultural memory institutions over the long haul. They are businesses focused on current commerce. Universities and their libraries have performed the archival and

preservation role enabling discovery and access to the nearly infinite long tail of information for hundreds of years and will continue to do so. The shift to private, corporate control of the archives in electronic databases puts future continuity at risk.

Businesses do not naturally collaborate. They compete and are designed as proprietary silos. This model is suited to competitive services but it is not productive when unique results of publicly funded research are barricaded behind a publisher wall.

Partnerships between publishers and academic libraries and government are possible by teasing apart the social roles consistent with the funding models. Valuable publisher services must and can be remunerated on a price for service basis, but not at the cost of public access. Libraries are funded through their institution to provide the basic archival persistent infrastructure needed for learning and scholarship over the long haul while the government establishes the rules to ensure public benefit for publicly funded research.

Comment 5:

Standards are going to evolve over time and will need to be incorporated into repository activities as appropriate at the time. A basic specification of Dublin Core, OAI-PMH (Open Access Initiative Protocol for Metadata Harvesting), Data Cite Metadata Schema is needed at this time. In addition the OAI-ORE (Open Access Initiative Object Reuse and Exchange) must be included to facilitate computational tools and reuse. Licensing conditions are another concern. CC-BY licensing of content greatly facilitates interoperability and the creation of discovery services.

Nevertheless, publicly funded research results must first be made publicly accessible. Then clear national standards for interoperability can be formally designed and promulgated to ensure that research and business can optimally benefit,

Comment 6:

A basic consistent, mandatory policy (research results must be publicly accessible) across all federal agencies is needed so that all stakeholders have a single set of rules to follow for federally funded research.

The government can leverage infrastructure already developed within the NIH and build on tools already in development e.g. SWORD (Simple Web-Service Offering Repository Deposit).

Benefits will accrue in other spheres such as grant reporting, bibliography generation, return on investment measurements, and patent office review, reducing duplicative effort and documentation across the public sector. On a smaller scale, this is already happening at Caltech where the local digital repository dynamically delivers accurate descriptive metadata to researchers' web pages saving time in re-keying and editing. The government stands to gain from an order of magnitude improvement in output re-use.

Comment 7:

Before considering extending a public access requirement to other formats, first establish it firmly for peer-reviewed research articles. The rest will logically follow.

All federally funded projects need to include a statement regarding required public access to the certified results in whatever genre they are finalized for earliest distribution.

Comment 8:

The ultimate goal must be no embargo for optimal return on investment through research and business productivity. Internal studies at Caltech show that one third of researchers' reading or use activity involves the most current year's papers. Therefore, the greatest benefit to research productivity will be achieved by making papers immediately accessible.

A transition period starting with an embargo of six months for a defined period of time is appropriate. A twelve-month embargo is much too long and thwarts timely uptake of new information.

This proposed legislation would further inhibit the dissemination of information to the public. Appropriately, the government funds research and the one year delay in posting should be sufficient if not too liberal.
Please defeat this legislation.

Susan Vaughn
Associate Librarian for Collection Development
Brooklyn College Library

Graduate Student Association – gsa@unl.edu
University of Nebraska-Lincoln
Lincoln, NE

Comment 1

Increased access and analysis of peer reviewed publications means that research will be read by more people and used in new ways to create economic growth and improve higher education. Wider and more productive analytical methods mean a much higher return on investment for research publications. Agencies can grow and improve access and analysis markets by instituting rules that require publicly-funded research to be freely available and completely open to use.

More students, entrepreneurs, professors, and developers accessing research means that findings can be used in a wide variety of ways and can be put to use by a wide audience. Restricted access and use of publicly funded content means that public funds are not being used to the best of their abilities, and the return on investment is reduced. Full and open use means that businesses and individuals can build new products and services upon research. This is limited or impossible with the current structure that buries publicly funded research behind expensive pay walls. Open Access also means that research will be available to the general public, making research available to more eyes in a much broader scope than ever before. Open access also fosters interdisciplinary application and greatly increases the value of established research. The impact of open access on a paper's visibility and citation count is well-documented.¹

By allowing full use of research publications, readers can be much more productive with the information by using new techniques such as data mining and machine reading, and creating a new infrastructure for research. New pathways and connections can be made with open data and citation mapping. Under the current structures, information is locked into silos and users are not able to foster communication between research. Research is only as good as its reach and availability, and the current system is built to hinder access and use. Research can only be used by teachers and students if they have access and the more research that is available to students and teachers, the better and more up to date the education can be. Better education and available research means that American students will be better suited to compete internationally, especially in cutting edge fields like biotechnology and alternative energy, where new research is key to competition. Open data techniques will also enable private companies to capitalize even more on public resources.

Research publications can best be archived by making them immediately accessible and completely open to use in a centralized repository. Faster commercialization spurs economic growth, creating new jobs and advancing American businesses. Companies can also build upon public data and improve services analytical and finding structures, like Google Scholar and goPubMed. By allowing entrepreneurs, scholars, and students to access them without restrictions, the entire data base can be used for data mining or derivative works, and can make the sum greater than the parts.

¹ JISC Report - http://www.jisc.ac.uk/fundingopportunities/funding_calls/2011/02/benefits.aspx; Battelle Report - <http://www.battelle.org/publications/humangenomeproject.pdf>, Celera/HGS - <http://www.nber.org/papers/w16213.pdf>, Houghton Paper - <http://www.cfses.com/documents/wp23.pdf>

It is essential that the most recent research is available to the public. It would not be conducive to innovation and cutting edge research if students, entrepreneurs, and researchers were forced to depend on old research if newer and better research is available. New publicly funded research that is widely available best utilizes public resources and provides immediate benefits to universities and businesses. Open Access has proven to be the most productive dissemination method for research. Open Access increases citations, promotes a diversity of sources within research, increases new research pathways, and makes research immediately available for use in both application and further research.

Once students leave school, they are met with expensive barriers to keep them from research. This greatly hinders their performance in the workplace, as they are unable to keep up with the most recent research in their field. It also hinders entrepreneurs' businesses, leaving them at a disadvantage in the global economy. By making the most recent and advanced publicly funded research available, new businesses stand a much better chance to utilize their skills and compete for a share of the market.

Research can also be best utilized through storage in a centralized repository, similar to the current NIH model. The benefits of an NIH-style access policy are estimated to be approximately eight times larger than the costs, and can be instituted at a relatively small cost. The NIH spends about \$3.5 – 4.6 million annually to provide access to all public-funded research, which is about 1/100th of 1 percent of their overall budget. Because of this policy, research is widely accessed and used by a broad population. The NIH database is currently used by more than 500,000 users per day, and the majority of users are outside education, meaning that many of them likely would not have had access to the research in the pay wall model that blocks access to most research outside of the NIH. Full open access is ideal to making all of these ideas come to fruition. Restrictions on use also limit the possible value from research investment. It also means that less money needs to be spent on duplicate research, either through public funding or within the private sector. It's also important that students be taught the most up-to-date information possible to best prepare them for jobs and make them best prepared to compete in the global market.

Comment 2:

Publicly funded research can respect the intellectual rights of researchers and allow for the most complete utilization of research by implementing licenses like Creative Commons's CC-BY license.² The NIH currently allows articles to be used under "fair use," which protects authors, though it restricts some of the usefulness of the research. By allowing full use of this research, the public can get the most out of their investment. To further protect scholars' intellectual property; there could be an embargo period, where fair-use is applied, with the research moving to CC-BY or a similar open license. Again, this is not the best way to get productivity out of the research, but it does provide the author more rights over their work.

Comment 3:

The federal government should provide permanent stewardship of research because it ensures that research is permanently preserved, made accessible, and most efficiently usable. By pooling all research together in a centralized location, everything is easily available and searchable in one place, and it's possible to build databases that encourage communication between different research, rather than different collections of research stuck in a number of separate silos, where integration is difficult or

² Creative Commons' CC-BY allows licensees to copy, distribute, display, and perform the work and make derivative works on it only if they give the author or licensor the credit.

impossible. Federal stewardship is also very cost-effective, as stewardship for the NIH is only 1/100th of one percent of their budget.

At the very least, a federal archive should collect and mirror all publications that are published elsewhere to ensure stability and preservation of the research. However, it's essential that these publications are readily available through the federal repository to ensure that research is stable and constantly available to foster the use of derivative work and accessing tools.

Comment 6:

Uniform requirements and mandates are necessary for consistent creation of publicly-funded research. Because institutions often have researchers who hold grants from multiple agencies, all agencies should establish the same standards to smoothly implement research. Uniformity amongst agencies means lower costs for institutions and an increased rate of compliance. Policies should take advantage of existing protocols to facilitate automatic deposit of manuscripts, and encourage the development of additional tools. Additionally, policies should integrate articles with grants management systems to improve agency accountability and provide information to the public.

Policies to increase tools and other finding methods should work to increase bibliographies and principal investigator profiles to better raise the connectivity of research and raise the profile of those researchers whose works are used and cited the most. These methods would allow universities to better measure research output and impact ratings, and would create better pathways to locate better research and allow universities and libraries to use repositories as teaching tools.

Comment 7:

Educational materials such as book chapters, notes, texts, syllabi, and conference proceedings should also be made readily available to the public, but may require different policies than those directed at journal articles. These types of unpublished works, most notably peer-reviewed conference papers and proceedings, represent a large portion of research and teaching materials that are very relevant to other scholars, as well as the public at large. Feedback from these kinds of papers is integral to the research process, and a wider audience can significantly improve research, as well as keep others informed on current trends and burgeoning research.

Graduate Student Association

University of Nebraska-Lincoln



the virtual lifeline for proactive cancer patients

ICAN Boards and Councils

Chairman, Board of Trustees
Sherry Weinstein

President and CEO
Marcia K. Horn, Esq.

Vice Chairman, Medical Issues
Robert H. Tamis, M.D.

Chairman, Board Governance
Carey Gregory

Chairman, National Board of Advisors
Joan Weil

Chairman, Advisory Council
Cathy Dalzell

Federal Affairs and Strategic Development
Linda Jenckes

Outreach Issues
Linda Boatwright

Special Projects
Scott Kent Knauer

Honorary Trustees
Ira J. Gaines, Honorary Chairman
Cheryl J. Hintzen-Gaines, Honorary Chairman
Shirley and William Bird
Janet and Clive Cussler
Don W. DeMichele, Ph.D.
Ellen and William Goldstandt
Bunnie H. Marks
Carolyn Sechler, CPA
Marlene and Eugene Shapiro

Chairman, Scientific Advisory Council
G. Robert Pettit, Ph.D.

Chairman, Biomarkers Council
Scott M. Kahn, Ph.D.

Chairman, Physicians Advisory Council
Robert H. Tamis, M.D.

Chairman, Patients Council
Uri Harel

Special Advisors, Surgical Oncology Issues
Richard J. Gray, M.D.
Barbara A. Pockaj, M.D.

Special Advisors, Medical Oncology Issues
Mitesh Borad, M.D.
Jeffrey D. Isaacs, M.D.
Robert A. Nagourney, M.D.
Michael S. Roberts, M.D.

Chairman, Radiation Oncology Council
Nicholas E. Flores, M.D.

Special Advisor, Pathology Issues
Arthur Sitelman, M.D.

Intellectual Property Counsel
Kevin F. Murphy, Esq.
Farley I. Weiss, Esq.

Founding Trustee Emerita
Carol Ginsberg

Founding Chairman
Sidney M. Rosen, Esq.

Recipient

Michael J. Stebbins, Assistant Director of Biotechnology (Science Division), Task Force on Public Access to Scholarly Publications, Office of Science and Technology Policy

Assigned ID #

FR Doc. 2011-28623

Assigned Entry Date

Filed 11-3-11

Name/Email

Victoria Wang, vwang@pblbio.com

Affiliation/Organization

Legislative Director, Biomarkers Council
International Cancer Advocacy Network

City, State

Phoenix, Arizona

Background:

The International Cancer Advocacy Network (ICAN) was established in 1997 as a 501(c)(3) charitable organization that has to date, assisted nearly 8000 late stage cancer patients from all 50 contiguous states and 40 different countries in their battle against metastatic cancer. ICAN is very much a volunteer-driven organization; its 700+ volunteers provide world class expertise in many disciplines, including basic, translational and clinical cancer research. ICAN stresses a personalized approach to cancer care for each individual cancer patient. ICAN relies heavily on new and timely data published in peer reviewed publications for the benefit of our cancer patients. An important part of our care is guiding each patient through biomarker and functional testing with the goal of improving their survival. Thus, keeping abreast of the latest developments published in peer reviewed scientific journals is of the utmost importance to ICAN and our patients.

Addressing Question 6 of FR Doc. 2011-28623:

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?

Federal agencies that fund science maximize the benefit of public access policies to U.S. taxpayers **by eliminating all barriers for free public access to ICAN and similar 501(c)(3) charitable organizations, including access costs and all embargoes on free access.**

ICAN relies heavily on the literature for the most recent advances in cancer research and care. The costs of journal subscriptions are, simply put, overwhelming for ICAN, especially in this current economic climate. ICAN is delayed, often for long periods of time, from accessing state of the art research publications that would otherwise be of immediate help to our patients.

Waiving the cost barrier and embargo period for 501(c)(3) charitable organizations is, without question, a win-win situation for all interested parties, and it absolutely must be considered as a part of this legislation. It will have an immediate short term benefit of extending the survival of at least some of our patients. And, as discussed below, it will ultimately result in the development and discovery of new areas of research and innovative treatment protocols that will expand the scientific literature, thus providing an advantage for publishers.

Addressing Question 8 of FR Doc. 2011-28623:

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?

As a 501(c)(3) charitable organization that interacts on a daily basis with late stage cancer patients, ICAN absolutely requires timely access to breaking peer-reviewed journal articles. This is a heavy cost burden for both ICAN and our patients and this places an unnecessary and easily correctable financial barrier between our patients and their optimal care.

Considering the many therapeutic advances and accomplishments that were driven by creative and persistent patients and advocacy groups such as ICAN, it is paramount and crucial that 501(c)(3) charitable organizations be exempt from any embargo period placed on peer-reviewed publications attained from federally funded research. The new and novel areas of research and clinical trials resulting from such efforts will undoubtedly open up new chapters for publication in scientific journals, thereby adding to their bottom lines. In summary, creating cost and embargo exemptions for 501(c)(3) charitable organizations are in the best interests of all concerned parties.

Victoria Wang



Legislative Director, Biomarkers Council
International Cancer Advocacy Network

To: White House Office of Science and Technology Policy:
From: Digital Scholarship Lab, University of Richmond

Subject: Response to two OSTP RFIs, Public Access to Digital Data Resulting from Federally Funded Scientific Research and Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research.

Along with humanities organizations such as the Association for Computers and the Humanities, we would like to urge the Office of Science and Technology Policy to adopt a more expansive definition of 'research' to include the work of humanities scholars. Broad public access to research content produced with tax-payer dollars and the long-term survival of resulting digital data are as crucially important to the humanities as they are to the sciences.

Robert K. Nelson
Director

Scott Nesbit
Associate Director

Digital Scholarship Lab
University of Richmond

12 January 2012

Task Force on Scholarly Publications
National Science and Technology Council
Committee on Science
Office of Science and Technology Policy
Office of the President
Washington, DC

Sent via email to: publicaccess@ostp.gov

To Whom It May Concern:

These comments respond to the "Request for Information: Public Access to Peer Reviewed Scholarly Publications Resulting from Federally Funded Research," published in the *Federal Register* 76(214):68518-68520, on 4 November 2011.

Comment (1): One thing that agencies could do would be to require that the results of research they fund be prepared and submitted for peer-review publication and fund such efforts. Substantial amounts of archaeological research in the US are funded as part of environmental impact and historic preservation reviews required by NEPA, the National Historic Preservation Act, or the Archaeological Resources Protection Act as part of public project planning. The number of substantial archaeological investigations reported by federal agencies exceeds 50,000 annually (e.g., Department of the Interior 2009, 2010). It is rare for the results of the historical or scientific research from these investigations to be published in peer-reviewed journals or books. Requiring a peer-review publication from such studies, and making these publications widely accessible would increase the flow of information available for subsequent investigations on related topics or geographic areas.

Alternatively, agencies could require that the results of these kinds of investigations be subject to peer-review, and that any subsequent appropriate revisions be made, prior to accepting the final report(s) of the investigation. Such a procedure would not require publication in a traditional scholarly journal. Realistically, either of these requirements should be limited to projects of sizable scope in order for the review to be worthwhile.

Another alternative would be for agencies to require peer-reviews of all substantial reports created for environmental or historic preservation identification and evaluation studies or data-recovery and documentation studies. This would have the additional value of improving the

DIGITAL ANTIQUITY

School of Human Evolution and Social Change
an academic unit of the College of Liberal Arts and Sciences

Francis P. McManamon, Executive Director

PO Box 872402

Tempe, AZ 85287-2402

Direct: (480) 965-6510 Digital Antiquity: (480) 965-1369 Fax: (480) 965-7671

<http://digitalantiquity.org> <http://tdar.org> fpm@digitalantiquity.org

final reporting on such projects typically done not for strictly academic or scholarly functions, but as part of public project planning and construction projects.

One would expect that instituting either of these agency procedures requiring peer-review will broaden access to information that will make subsequent investigations more effective and efficient. Any new studies will have the advantage of better information from which they would be starting, information that is firmer and more widely based than if access to data and information from earlier studies is not accessible. Easier, more accurate, and quicker environmental reviews for public projects clearly would contribute to US economic growth and productivity.

Cited works:

Department of the Interior

2010 *The Goals and Accomplishments of the Federal Archeology Program: The Secretary of the Interior's Report to Congress on the Federal Archeology Program, 2004–2007*. Departmental Consulting Archeologist, Archeology Program, National Park Service, Washington, DC. (<http://www.nps.gov/archeology/SRC/reportPdfs/2004-07.pdf>).

2009 *The Goals and Accomplishments of the Federal Archeology Program: The Secretary of the Interior's Report to Congress on the Federal Archeology Program, 1998-2003*. Departmental Consulting Archeologist, Archeology Program, National Park Service, Washington, DC. (<http://www.nps.gov/archeology/SRC/reportPdfs/1998-03.pdf>).

Comment (3). A network of decentralized disciplinary-based digital repositories will be the most effective way in which to manage public access to federal research data and information. The variation in metadata organization and terminology among the wide variety of scientific disciplines involved in government research is too large to be effectively and efficiently accommodated by one or a few centralized repositories. However, the Federal government does have a role to play in establishing minimum metadata standards, regardless of discipline. Further, these disciplinary-based repositories must be interoperable, that is, linked through a central portal. In this way, the actual document (or other information resource) is stored in a decentralized repository but the descriptive information (metadata) about the item is accessible in a centralized repository like data.gov (<http://www.data.gov/>).

Comment (4). Existing publisher archives could be made better known and more widely used if metadata about the publishers' catalog listings, including summaries of the books, articles or book chapters they contain were exposed to searches by being accessible through digital repositories. In the field of archaeology, for example, [the Digital Archaeological Record \(tDAR\)](#) is open for publishers to create a metadata page for each of their archaeological publications. The metadata includes a description of the contents of the publication and standard archaeological metadata terms to assist with discover by individuals search the tDAR repository. Publishers may upload a portion of the publication the metadata page refers to (e.g., the front matter and perhaps an introductory chapter). Publishers also may include

information about how to order the publication, or a link to the publisher's web site for those who want to purchase it.

There are mutual benefits from this kind of commercial/not-for-profit partnership ([the Center for Digital Antiquity](#) which maintains tDAR is a not-for-profit organization being incubated at Arizona State University). The repository function that Digital Antiquity is set up to carry out gains additional digital resources that it can make available to its users. Publishers gain an inexpensive and easy way of advertizing their publications. The overall benefit is that available information is made more easily discoverable, accessible, and usable. In effect, open and not-for-profit repositories like tDAR are linking disparate information about a topic or an area, by including metadata from commercial publishing firms with the metadata and documents in open repositories. Users gain a "one-stop-shopping" experience that increases accessibility for users.

Comment (7). Besides scholarly journal articles with the peer-reviewed results of research funded by federal agencies, there are a number of other kinds of products from research that should be made accessible to the public. In my answer to Q.1, I noted that much of the research results from federally funded investigations are not peer-reviewed. I suggested that federal agencies should change this by instituting procedures requiring peer review, and funding it, at least for projects with sizable budgets and scopes.

There are a variety of research products that should be available to the public, within the limits of individual privacy protection and limiting the exposure of confidential information, copyrighted works, and individual intellectual property. Among these are technical and descriptive reports about the methods, techniques, and substantive information of the research, data sets (spreadsheets and databases with basic descriptive and analytical data), images, and scanned data of various sorts (e.g., GPS, GIS, object or landscape scans, etc.).

Other items for the TF to consider. As noted in several of the answers above, substantial results of research funded by federal agencies are not peer-reviewed. Agencies should institute procedures to provide funding for more peer-reviewing. Agencies also should make accessible more of the results that are not peer-reviewed so that these results can be used more widely.

Please feel free to contact me directly if you have any questions or seek additional information regarding our comments.

Sincerely,



Francis P. McManamon, Ph.D., RPA
Executive Director and Research Professor

Name/Email: C. Jeffrey Belliston

Affiliation/Organization: Private Citizen

City, State: Pleasant Grove, Utah

I have personally benefited from the public access mandate of the National Institutes of Health. My insurance company denied authorization for a physician-recommended treatment protocol for one of my daughters because their standards indicated that the protocol was approved for only one diagnosis. I work at a university and was able to access some articles through the databases that the university library licenses. Through these databases I found some useful articles. However, I also accessed additional useful articles published in journals that the university does not have paid access to via NIH's PubMed Central. I was able to cite research evidence indicating that the physician-recommended treatment protocol had a reasonable likelihood of success for my daughter with potentially far greater effectiveness and far fewer side effects than any other available protocol was likely to produce. This helped to persuade the insurance company to reverse their decision. This experience galvanized my feelings about public access. The vast majority of the American public does not have the kind of access I have as a university employee. They deserve access to what their taxes are paying for.

Comment 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?

Yes. I am not personally aware of anyone who has undertaken research and published their findings for the sole purpose that they enjoy research. Research is generally undertaken to advance knowledge and bring about positive change thereby. Among other things, positive change can come in the form of new products and services which "grow existing and new markets." The single most important steps any federal agency "could take ... related to the access and analysis of peer-reviewed publications that result from federally funded scientific research" is (1) to mandate that funded researchers grant the funding agency a non-exclusive, irrevocable license (before signing copyright over to a publisher) to make the final peer-reviewed manuscript available in a publicly open repository and (2) to act on that granted license and make the publication accessible.

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?

Archiving publications and making them freely accessible to anyone who is interested, rather than locking them up behind a pay wall as commercial publishers are wont to do, has the distinct possibility of releasing creativity, innovation, and entrepreneurship that would not otherwise be brought to bear. Such policies will make research available to institutions of higher education who would not otherwise have access. Further research could be done. Existing business enterprises--especially smaller enterprises that cannot afford commercial publisher prices for access--could use such research to enhance existing or make new products and services. If, as I frequently hear, small business is really the engine for job growth, such policies have the clear potential to spur such growth. Insofar as improving "the productivity of the scientific enterprise" is concerned, logic indicates that improved access (and I consider free access to be a definite improvement over paid access) would lead to less duplication in research endeavors.

Increasingly, I see references to meta-studies or meta-analyses. These are also likely to increase meaning that a single study has use even beyond what those who undertook it originally envisioned. This also means more "bang for the buck."

What are the relative costs and benefits of such policies?

I have not personally read the several studies that have been undertaken by the economist John Houghton examining such policies in various countries. I have read summaries of Houghton's studies. The summaries suggest that the most conservative estimates indicate that such policies bring 4 to 5 times more benefit than the costs involved.

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 ideal type of access is both the ability to read the publications and to reuse them. I am familiar with the Creative Commons and the variety of licenses they have created. The CC license enabling both reading and reuse is the CC-BY license. Both abilities should be made available to interested users as soon as possible and, ideally, immediately with no embargo period.

Comment 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?

I bristle somewhat that publishers are mentioned first in this listing of those who are potential intellectual property stakeholders. Publishers are among the last, if not THE last, to have any IP interest in a scholarly publication. Their intellectual contribution to any publication is minor at best and consists of copy-editing IF that can be considered an intellectual contribution rather than a stylistic or technical contribution. They become an intellectual property stakeholder when they force scholars to sign over their intellectual property rights.

The NIH public access policy is generally a good model to follow in this case. The requirement that a funded researcher grant a non-exclusive, irrevocable license to NIH to make the publication publicly available before the researcher signs any copyright transfer agreement with a publisher is clearly legal. If it were not the publishers would have sued to stop the NIH mandate.

The researchers interest is protected given that they retain their entire bundle of rights under copyright--at least unless or until they "voluntarily" sign them away to a publisher. Some may argue that the researcher's right is being abrogated because of the requirement that they grant the license to NIH is specious. The decision to seek NIH funding is a voluntary one. If a researcher objects to this requirement, s/he is under no obligation to solicit NIH funding.

In my view the federal agency is only a proxy for the taxpayer. The agency is the mechanism through which the taxpayer's interest is protected. Taxpayers pay for research and should be able to read and reuse what they have paid for without paying again! This leads to one of the drawbacks of the NIH policy. It does not go far enough. The NIH policy should allow not only for access; it should be modified to also permit reuse.

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?

Yes. There are policies that should NOT be adopted. No attempt should be made to require publishers to open up back files of scholarly content to which they legally acquired copyright

prior to any policy enactment. I would expect them to challenge any such attempt in court and would have to stand with them if such an attempt were made.

Comment 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?

I do not know all of the pros and cons of these approaches. One of the pros of a centralized approach is that NIH has already paved the way and it is reasonable to expect that other agencies could get up and going much more quickly because of NIH's experience (or that some agency could be tasked to be the central repository of all funded research from whatever agency). Another pro of a centralized approach is that it would provide consistency for researchers. Another is that if the federal government is likely to be more stable and enduring than any other enterprise whether commercial or non-profit, private or public.

Frequently, the federal government is neither nimble nor innovative. These qualities are more likely to be found in smaller entities. These qualities of a large federal bureaucracy are definite cons to a centralized approach.

The best approach might be a combination of approaches. If the decision is made not to establish either a single federal repository covering all agencies or repositories in each agency but to allow researchers to deposit in non-federal repositories meeting at least a set of minimum standards, then the federal government should somehow maintain a copy of publications from funded research in the event of a failure in/disappearance of non-federal repositories. Single copies are never a good idea to ensure long-term stewardship.

Comment 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?

Nothing immediately comes to mind as a model of such a partnership. There may well be and, due to the lack of nimbleness and innovation that too frequently characterizes the federal bureaucracy, may be an attractive alternative. Publishers should not have to participate if they choose not to but their participation should be welcomed IF (1) the price is right; (2) the intent of funder mandates to provide freely available access and, I would hope reuse, is not in any way compromised; and (3) a mirror site(s) is/are maintained to ensure the long-term stewardship as mentioned in Comment 3.

Comment 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?

I do not consider myself qualified to answer this question. I do know that standards help and that there are likely to be a number of both established and potential standards (some likely in

draft form already under consideration) covering a variety of areas of relevance. I also know that standards evolve and, typically, improve. All of this should be taken into consideration. However, adopting public access policies across federal funding agencies NOW is more important than waiting for some standard to be approved.

Comment 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?

To my mind, consistency is a critical factor in maximizing the benefit to taxpayers while simultaneously minimizing the burden and costs for other stakeholders. Preferably a single policy covering all current agencies (and any future agencies that might be created) would be established or, alternatively, identical policies for the different agencies would be established. Inconsistency breeds confusion, compliance headaches, and reduced benefits.

Another key piece to *maximizing* benefit is to build reuse rights, not just access rights, into the policies. This shortcoming of the existing NIH policy should be rectified.

Comment 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?

Potentially. From an internal logic standpoint (i.e. the taxpayers are funding it and therefore have every bit as much right to it) this makes sense. However, as I understand it, the bulk of the scholarly publications stemming from federal funding of scientific research is in the form of articles in scholarly journals. There is every reason to move forward expeditiously on this front especially given the existing NIH mandate and all that has been learned from the NIH experience. The policy could later be extended to cover additional publication types/outlets if warranted.

Comment 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 my opinion, there is no appropriate embargo period. Ideally the material is immediately accessible and reusable. An embargo period assumes that a commercial publisher's interests are of more importance than the interests of all other stakeholders in this process and that is, to my mind, a false assumption. An embargo period may, nevertheless, be necessary for political reasons. If that is the case, the embargo should be as short as possible and no more than 6 months. The empirical basis for that statement is that, aside from the NIH, embargoes allowed by other research funders with public access mandates are 6 months at a maximum. The NIH is an outlier and outlier should not be the basis for a policy decision.

Please identify any other items the Task Force might consider for Federal policies related to

public access to peer-reviewed scholarly publications resulting from federally supported research.

In an earlier career, I was a Foreign Service Officer with the Department of State. I know that the United States government expends a considerable sum on aid to foreign countries and that many Americans question the wisdom of doing so. Making federally funded research openly accessible to taxpaying Americans will inevitably open it up to researchers and the general public outside our borders as well. In my mind, that is an additional boon to the adoption of such policies. It is potentially the least expensive type of foreign aid we could engage in and has the possibility of generating significant and long-lasting good because it is more in line with the proverb that "If you give a man a fish you feed him for a day but if you teach a man to fish you feed him for a lifetime" than is much of our foreign aid.



American Educational
Research Association

Response to Request for Information (RFI): “Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research,” Office of Science and Technology Policy (OSTP)

76 Federal Register 214, pp. 68518-68520, November 4, 2011

**American Educational Research Association
Felice J. Levine, Executive Director (flevine@aera.net)**

January 12, 2012

About AERA

The American Educational Research Association (AERA) is the major national scientific association of 25,000 members dedicated to advancing knowledge about education, encouraging scholarly inquiry related to education, and promoting the use of research to serve the public good. Founded in 1916, AERA as a scientific and scholarly society has long been committed to knowledge dissemination, building cumulative knowledge, and promoting data access and data sharing. AERA publishes six highly ranked, peer-reviewed journals in the field and holds an annual meeting with approximately 14,000 participants, among other initiatives. In 2010, AERA introduced an online paper repository as a further vehicle to foster the sharing and dissemination of work prior to publication. AERA disseminates one of its highly ranked journals freely on its website.

AERA applauds the principles that lead OSTP to think through policy issues supporting the scientific enterprise and public access to knowledge. There are complex issues involved in assessing the responsibilities of the federal government and scholarly societies in such endeavors. The responses below seek to foster further examination of this issue, including the appropriate role of the federal government, from the vantage of sound research policy and viable business models of publishing.

As a non-profit research organization, AERA plays a key role in facilitating scholarly communication and knowledge dissemination. As with other associations in the social and behavioral sciences, AERA needs to maintain a peer review process of the highest quality (in an era of diminishing support for doing so); to provide access to publishing opportunities based on research quality, not individuals’ resource availability; and to serve science and society through affordable publishing and archiving. Because a short

embargo period that is inconsistent with a model of social science publishing would be problematic, AERA renews its previous recommendation advanced in a similar comment period in January 2010 to implement publisher-provided tollfree hyperlinks from federal agencies to the version of record immediately on publication. We speak from the vantage of a research society committed to affordable, sustainable publishing and maximizing opportunities for publishing research of the highest merit irrespective of the source of its funding. This step would efficiently and effectively achieve that goal.

Responses to RFI Questions

(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 publicly 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 federal government and its agencies could take steps to grow markets for access and analysis of research by endeavoring to support sustainable publishing efforts. Since the 1990s, publishers have invested heavily in online publishing platforms, tools, and functionality, and those platforms have both alleviated print archiving burdens on libraries and expanded opportunities for access to research. AERA has digitized the entire back content of its six highly ranked peer-reviewed journals and offers a wide variety of access options, including free access to *Educational Researcher* through the AERA website, a variety of subscription models—including discounted consortial packages—to address libraries' needs, pay-per-view options, and discounted and free online subscriptions to nonprofits in countries with low GDPs per capita. AERA also offers tollfree hyperlinks to its authors who wish to link to the version-of-record (VoR) of their AERA journal article, and these tollfree links may be implemented in institutional repositories or institutional webpages. Through these initiatives, AERA seeks to make peer-reviewed research available to the broadest possible audience while sustaining its publishing endeavors. Since moving to its current online platform, the number of institutions with access to AERA journals has increased dramatically: from 2,450 institutions in 2006 to 6,550 institutions in 2011, for a 167% increase.

In addition to the traditional models for access to research journals, AERA remains interested in alternative models such as open access, provided there are funds that will sustain the enterprise. The inclusion of author publishing fees in all government grants, including grants that fund social science research, would be one significant way that federal agencies could support accessible publishing models. It should be noted, however, that a more cost-effective cooperative endeavor between the federal government and publishers would be the acceptance of publishers' tollfree hyperlinks for implementation in grant agency databases. Such an approach also averts using

federal funds for such purposes or fostering a system where investigators without large grants or federal grants at all would have an undue burden.

(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?

Publishers such as AERA acquire from researchers license to publish peer-reviewed research in return for the publishing channels they offer and grant to researchers, in return, the license for certain noncommercial (e.g., educational) uses of the VoR. AERA also offers each researcher a tollfree hyperlink to the journal article VoR. AERA also provides Federal employees the ability to publish journal articles in AERA journals with the government retaining the right to use the VoR for government purposes.

In return for the considerable investments made in online publishing channels and platforms, publishers need the opportunity to recoup funds that help sustain the publishing endeavor and preserve their digital archives for access. Any federal mandate for full-text deposit less than 5 years after initial publication would endanger the social science publishing enterprise because social science journal articles are more expensive to produce than those of other disciplines and, thus, social science publishers require longer to recoup costs. Tollfree hyperlinks can be utilized immediately upon publication without danger to sustainability of the enterprise.

(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?

A centralized approach to managing public access has the advantage of yielding the most federal control but the disadvantages of (a) encumbering the federal government with long-term archiving expenses and (b) likely endangering publishers' enterprises that sustain their peer-review processes and their digital archiving solutions. A decentralized approach that features federal agencies partnering with publishers and other stakeholders could leverage the special strengths of those stakeholders and allow the federal government to focus on research oversight.

Within the decentralized approach, the federal government could ensure long-term stewardship of content if it includes deposit of a full-text VoR that remains dark to all users. This would act as an insurance policy if the published VoR becomes inaccessible. If tollfree hyperlinks were in use in the government database, the full-text VoR could be a backup in case the link ceases to function.

Most publishers' online platforms allow the use of searching and other interoperable functionality at no cost to online readers, and AERA's journals are on such a platform. If the federal government would wish to develop searchability and other interoperable functionality within any of its research databases, then a dark full-text VoR would enable searchability that could lead to the publisher's VoR.

(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?

Since 2007, AERA has offered authors tollfree hyperlinks to the VoR of their published articles. In its January 21, 2010, response to OSTP's previous request for information, AERA indicated that it would do the same for federal agencies funding research published in our journals. We advance that recommendation again here. Other nonprofit publishers are increasingly making use of these tollfree links. We applaud their expanded use and encourage the federal government to accept them in lieu of full-text articles that, in a free archive, would endanger the sustainability of social science publishing.

(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?

Most publishers such as AERA have already implemented arrangements for interoperable search and discovery across disciplines and archives. AERA's journals are hosted on Stanford University's HighWire Press platform, which offers state-of-the-art search and interoperable functionality. For example, on behalf of its publishers HighWire arranges full-text indexing of journal articles by Google and other web search engines. HighWire also offers tollfree reference linking, whereby those who subscribe to one journal whose article references another journal on the HighWire platform may access the second journal tollfree. HighWire facilitates reference linking beyond its platform via the use of the CrossRef consortium of publishers. In addition to offering a HighWire-wide search function, HighWire enables its publishers to design search

widgets that target searches from outside websites to certain HighWire journal content. Federal policies that protect the sustainability of publishers' enterprises will enable such search and discovery initiatives as these to continue to serve the research community.

AERA and most other publishers already make considerable metadata about journal articles available to the public, without access control. For example, anyone with Web access can search AERA and other HighWire journals without needing access control and can access abstracts without access control. Federal agencies often currently replicate this metadata; for example, the Department of Education's Education Resources Information Center (ERIC) replicates such metadata provided to it by AERA and other publishers.

There are numerous schemas that would serve any Federal agency providing research metadata to the public. In addition to ERIC's schema, CrossRef and DublinCore offer metadata schemas that agencies might utilize. We would encourage any agency adopting or developing a schema to ensure that it includes digital object identifiers (DOIs), in order to help online readers locate the VoR.

(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?

Federal agencies can maximize the benefit of public access policies by developing such policies that allow each set of stakeholders to focus on and sustain their areas of expertise. Those policies should allow researchers to focus on their work generating new knowledge and should allow publishers to evaluate and distribute that knowledge. Those policies might also perhaps be designed so as not to encumber taxpayers with new obligations, and the ideal way to do this is to partner with researchers and publishers to utilize each group of stakeholders' core strengths.

If ensuring full open access is the goal, then federal agencies should develop funding mechanisms that allow researchers to pay author fees to open access journals, in order to sustain the publishing enterprise. As discussed above, however, there may be better ways to maximize public access through cooperative efforts with publishers and to avert some of the downsides of a fee structure that could be large for authors.

(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?

It is difficult to comment on whether the scope of such policies should extend beyond peer-reviewed journal articles to other forms of scholarly products until they are further specified. In the cycle of social science research, conference presentation papers are

frequently initial reports of research results or findings that have not yet been fully developed by the researchers or vetted through the extensive peer review that is the hallmark of scholarly publication. AERA and a number of other research societies have introduced online paper repositories for peer reviewed papers to enhance knowledge dissemination at an earlier point in the cycle of knowledge dissemination. Our repository provides authors with the capacity to point later to publications and other final products.

Book chapters may have more extensive peer review than papers or works in conference proceedings, but are they are rarely reporting on single studies and more typically are creative products of larger scale that may only reflect in part specific studies or federally funded work. Also with such works as research volumes or handbooks, chapters may be invited by, conceived as, and parts of the intellectual creative product of others serving as volume editors. Such a situation may materially vary from journal articles developed and prepared independently by scientist author(s) to disseminate their findings and results. Also, the model of publishing books, including those published by scholarly societies like AERA, differs from journal publishing. For example, the formats in which e-books are currently published do not typically lend themselves to technological solutions such as tollfree links to chapters. Each form of scholarly publication should be examined in terms of the distinctiveness as well as similarities of its development and production, where federal funding fits, and whether or not there is an appropriate federal role.

(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 social science and humanities disciplines, the peer-review process is more time-intensive and far more selective than in the science, technical, and medical disciplines, and therefore much more costly per published article. In social science, the median age of cited social science journal articles (aka, citation half life) can be well over 10 years; this is certainly true for all AERA's journal articles. Because of low acceptance rates in social science journals, the business model of social science publishers, encumbered by high publishing costs per article, is predicated on this half life and keeping annual subscription costs low for the several years of citation usage. The 2009 report *The Future of Scholarly Journals Publishing among Social Science and Humanities Associations* (<http://www.nhalliance.org/bm~doc/hssreport.pdf>) provides evidence of this. A 5-year embargo timeframe could be one that allows financial sustainability in social science. But immediately implementing tollfree hyperlinks to the published VoR

would be far more effective in making research accessible without increasing federal costs or burden.

Overall Perspective

In conclusion, AERA is supportive of making peer reviewed publications, including work that is based on federal support, widely accessible. A commitment to fostering wide dissemination and cumulative knowledge is central to our mission and purpose. If there is evidence that peer-reviewed publications from federally funded research are not sufficiently accessible on a timely basis, we urge that a partnership be further forged and strengthened between the federal government and scholarly societies that pools our expertise and shared interest in science and the public good to address this issue. We think that the tollfree link is an accessible and efficient solution, and we continue to urge that it be embraced.

January 12, 2012

To: Office of Science and Technology Policy, Executive Office of the President
From: Ginger Strader, Director, Smithsonian Institution Scholarly Press
Re: Response to OSTP RFI on Public Access

In response to the Office of Science and Technology Policy request for information on “Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research,” I offer a compilation of opinions from individual staff members. These responses do not represent official responses or the formal position of the Smithsonian Institution.

The Smithsonian is not an executive branch agency, and many of its publications result from both federal and private funding and incorporate the work product of federal and non-federal employees. As a result of its unique classification, when considering the question of public access, the Smithsonian is mindful of its obligation to protect intellectual property rights of authors and other contributors when publications are not in the public domain.

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?

- Include funds in all research grants that will support reasonable authors’ fees for online publication in open access journals (or equivalents for different types of publications). If the grant is the only source of funding, grantees could be required to publish in open access venues.
- Continue to fund conversion and tagging of metadata and content so it may be harvested by existing, online databases and portals (e.g., Encyclopedia of Life).
- Researchers should take advantage of open repository options that may exist within their library systems (e.g., National Agriculture Library, Smithsonian Libraries) that offer archiving options for digitized publications. By exposing these repositories to standard harvesting tools and methods (e.g., OAI-PMH, Google Scholar), the publications will have higher visibility.

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?

- Encourage conspicuous identification of copyrighted elements.
- Do not require agencies to identify all content with general statements such as “government publication,” which may inadvertently mislead users to believe the entire publication is public domain and may lead to unpermitted reproduction and /or sale of intellectual property.
- Allow agencies to impose embargo periods on results prior to publication, during which grantees may protect their intellectual property.
- Restrict or limit assignment of rights to publishers that did not contribute to the creation of the IP, only to production and dissemination of the final product. Impose time limits to be applied to publishers’ exclusive rights to disseminate, so publicly funded research results can be re-issued in an open access journal after first publication.

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?

- Centralization of published scientific findings would allow an extension of simple search and download functionality. For example it would allow natural language processing and text mining software to inform research and guide discoveries that scientists cannot achieve by simple, linear reading of publications. This could be used to better inform hypothesis creation and to prevent duplication of effort.
- Encourage decentralized but open and standard metadata that can be aggregated into larger tools (e.g., the defunct National Science Digital Library).
- Unless a publisher distributes through an open access platform, restrict or limit assignment of rights to the publisher to ensure non-exclusivity, with a clause in publishing agreements for federally funded research that allows authors and/or funding agencies to reproduce and disseminate the same research independently through its own open access outlets.

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?

- In addition to Google Scholar, a good model that provides storage as well as access and social media activity is Mendeley (www.mendeley.com).
- Subscription models such as JStor.

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?

- Metadata: OAI-PMH, OAI-ORE, ONIX, Dublin Core, MODS, BibTeX, RIS, OPDS.
- Market forces seem to be taking care of this, led by open access journals, which are making content discoverable. In response, research communities are building systems for mining and processing that content. Some basic research on this area of information science is being funded. Federal agencies may want to follow suit.

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?

- Book chapters and essays in conference proceedings that are fully federally funded and, in isolation, are in the public domain should be included.
- Primary datasets should be included.

--

Ginger Strader
Director, Smithsonian Institution Scholarly Press
P.O. Box 37012, MRC 957
Washington, D.C. 20013-7012

Dear Science and Technology Policy Office,

Thank you for extending the deadline for comments on Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research. The Research Works Act has only very recently come to the notice of scientists, and it is because of this extraordinary proposal that it is now apparent to us that we need to reaffirm what we thought was settled: that OF COURSE scientific work funded by the public should be freely accessible to the public. I do not understand how this can even be a matter for discussion. The public pays: the public should benefit in every way possible.

The language in the RWA is highly misleading, attributing to publishers far more input into the scientific process than they really have. The truth is that scientists (often funded by public money) provide the underlying research, the writing and the figure preparation that result in a manuscript submitted for publication.

Other scientists then provide the editorial services and (contra publishers' claims, as can be easily verified) the peer review.

Publishers' contributions are limited essentially to typesetting, the provision of web hosting, and sometimes a very limited amount of compensation for senior editors only (usually not the handling editors who actually deal with authors' works). The notion that such a minor contribution should suffice to hand publishers, rather than the public, the right to determine how, where and under what regime the resulting works are disseminated, is ludicrous. It would be laughable if it were not so iniquitous.

Yours,

Dr. George E Homsy
Atair Aerospace, Inc.
454 Las Gallinas Ave, suite 175
San Rafael, CA 94903

I am writing in response to the Office of Science and Technology Policy's recent requests for information on public access to peer-reviewed scholarly publications and digital data resulting from federally-funded research (76 FR 68517 and 76 FR 68518) on behalf of the National Science and Technology Council.

I encourage the Council to issue strong recommendations to maximize public access to data and publications resulting from federally-funded research. Openness maximizes the benefits of research by increasing its scientific and economic impact while upholding scientific integrity.

To ensure that taxpayers derive the most benefit from the research they support, the Council should recommend that federal agencies require grantees to make their publications freely accessible to the public at no cost no later than six months after publication.

In addition, the Council should recommend that federal agencies require grantees to submit data management plans describing how they will manage, share, and provide public access to their data, if at all. The Council should also recommend that agencies establish expectations that grantees will provide public access to their data to the greatest extent possible, with narrow and specific exemptions (such as to protect human subjects and national security).

By issuing these recommendations and encouraging agencies to promptly implement them, the Council will fulfill its responsibility to advance federal science and ensure the best use of taxpayer dollars.

Sincerely,

Gavin R. Baker
Graduate student
School of Library and Information Studies Florida State University

[Assigned ID #]

[Assigned Entry date]

Name/Email

David W. Robinson, Ph.D.
Executive Vice Provost

Affiliation/Organization

Oregon Health & Science University

City, State

Portland, OR

Comment 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?

Complete, free, and re-usable access to the collection of scholarly output resulting from publicly funded research will dramatically spur market growth and scientific productivity. Several recent studies demonstrate a causal relationship between openness and an increase in the number and diversity of active researchers, an increase in the number of citations, an increase in new research lines, and an increase in upstream and downstream research activities.¹ With free and re-usable access, individuals and institutions—private and public—will drive innovation and invention. As such, open access policy must include liberal and explicit re-use rights in order to ensure the commercial legitimacy of resulting innovations, thereby encouraging, not stifling, economic investment. Finally, earlier access facilitates a quicker development cycle; new products and services are launched faster and more often. The ultimate results of free, re-usable, and timely access to this material will be diverse economic growth and an increased and earlier return on publicly funded research.

The resulting benefits in innovation and invention of supporting and managing an open access policy far outweigh the associated costs. In biomedical research, this is easily demonstrated. A host of recent studies support such a view and the U.S. can look to the

performance of the NIH Public Access Policy and the Human Genome Project as familiar and strong proofs of concept. The NIH reports that it costs between \$3.5 and \$4.6 million annually to provide access to its funded research results.² This figure represents less than 1/100 of 1 percent of the agency's overall budget.² Over 500,000 users access PubMed Central daily, demonstrating the profound demand for this information.² Initially, nearly \$4 billion was invested in the Human Genome Project. Since its inception, an entire industry has developed to support genomic research and R&D. The return on investment is dramatic; in 2010, the industry produced \$67 billion in U.S. economic output, \$20 billion in personal income for U.S. citizens, and 310,000 jobs.³ A powerful and specific example can be found at our own institution, Oregon Health & Science University. Our faculty member Dr. Brian Druker and his team developed the groundbreaking cancer drug Gleevec, an endeavor intrinsically linked to the research sharing and advances the Human Genome Project fostered. Gleevec's success has inspired a growing industry of second-generation gene-targeted cancer therapies. Houghton estimates that extending an NIH style open access policy to all other U.S. science funding agencies will conservatively result in a five-fold increase in ROI over a 30-year period with gains on the order of \$1.5 billion.⁴ Moreover, such an extension can leverage the existing infrastructures, investments, and successful management strategies of the NIH policy and PubMed Central to minimize additional costs. It should also be recognized that openness might reduce upstream expenditures, such as the time/cost of research, unnecessary duplication, and educational outcomes/attainment, lowering the price of research execution.⁴ Finally, open access to research increases accountability and enables more efficient funding and policy management. Agencies, budget drafters, and appropriators will have improved accounting on outcomes and enhanced information to assess value, identify promising research, and inform policy decisions.

Comment 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?

We strongly believe that publishers, scientists, federal agencies, and other stakeholders should be rewarded for the value they add to the research enterprise. Openness has the potential to increase and diversify the commercial and social-good opportunities founded on publically funded research and the associated rewards. Additionally, openness has the potential to increase and diversify the people and institutions participating in the exploration and execution of these opportunities. Working within

existing copyright framework and utilizing a stepped approach can ensure realistic stakeholder protection while enabling the fullest scientific, public, and commercial benefits.

A read-only access policy will not be sufficient. In order to unlock the scientific and commercial potential of publically funded research findings, individuals, institutions, and machines must be able to mine, analyze, and re-use the information. Appropriate licensing, such as the Creative Commons CC-BY 2.0, which allows users to share, re-use, adapt and make commercial use of the publication content, can facilitate this. To balance the interests of all stakeholders, full re-use rights could be activated after an appropriate embargo period.

Comment 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?

Access, technical operability, legal operability, and long-term preservation standards must guide the stewardship framework and its management. Third parties could maintain repositories that meet and demonstrate these conditions, presenting opportunities for partnerships and commercialization. Over the last twenty years, universities, academic libraries, and research institutions have built a network of institutional repositories, including PubMed Central. In the U.S, 235 repositories are currently cited in the Registry of Open Access Repositories. Standards to ensure these databases support human and machine based discovery, access, re-use, and innovation have been developed and continue to evolve. Hundreds of repositories and open access publishers utilize the Open Access Initiative's metadata harvesting protocols, for example. Additionally, a modest commercial sector has developed to support this work. This experience and infrastructure can and should be leveraged.

While access to publically funded research results can be supported through third-party partnerships, the federal government is the appropriate entity to provide ultimate stewardship. It should, at minimum, maintain an accessible mirrored version of all content, and public access policy must address standards and enforcement protocols for third party participation. A government maintained archive, its accessibility, and use is necessary to ensure research investment leverage and preservation. Moreover, as PubMed Central has demonstrated, this stewardship is cost-effective: PMC represents less than 1% of the overall NIH budget.

Comment 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?

Yes, there are numerous opportunities for public-private partnerships. The private sectors, and specifically publishers, bring to the table beneficial experience, funding, and technology. We support the creation of private-public partnerships as long as there are sufficient access, operability, and preservation standards and enforcement protocols. A broad view of public-private partnerships is ideal, one that not only recognizes opportunities related to publishers and other private entities as content repositories, but also as discovery experts, technology providers, content re-packagers, and business strategists.

It must be emphasized, however, that a healthy, successful access and preservation policy cannot be tied to a single site access point. Therefore, all associated public-private partnerships should be non-exclusive. As mentioned above, academic libraries, universities, and research institutions have extensive repository experience. This knowledge and infrastructure should also be mined for partnership opportunities with the same broad approach outlined above.

Comment 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?

Metadata, like the content it describes, has inherent value. In this view, metadata should be seen as facilitating specific actions, not merely as item description. It will be the foundation for discovery, powerful tools, and derivative products. As such, it is important that technical standards guide its definitions, expression, and communication in order to facilitate use, re-use, and analysis.

There are existing practices and standards that can inform these efforts. As mentioned above, the Open Archives Initiative's Protocol for Metadata Harvesting is in wide use

across the archives, repository, and open access publishing communities. The Dublin Core Metadata Initiative and its associated schema have done much to advance the creation and use of interoperable metadata standards for smarter discovery. Additionally, organizations like the National Information Standards Organization (NISO), DataCite, and the Library of Congress are working to ensure more intelligent, flexible discovery especially within the emerging context of the Semantic Web.

Comment 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?

Inter-institutional requirements and compliance standards for the deposit and delivery of peer-reviewed articles will keep implementation and long-term management costs low. Existing experience, like that of the NIH, can be utilized and improved upon as a cost model. Researchers and institutions rely on and must manage funding from multiple agencies. Standardization will generate better compliance, as stakeholders will be able to better navigate the necessary workflows. Standardization will also reduce the compliance burden on researchers and other content generators: it is absolutely essential that the compliance standards developed do not add to the considerable and ever-increasing regulatory burden that researchers already face. Whatever processes can be automatized should be made so.

Such consistency will also enable responsibility distribution across agencies, awardee organizations, publishers, and other stakeholders. As we have seen with the NIH Public Access Policy and PubMed Central, publishers will be attracted to low-cost, automatic and immediate deposit procedures. Awardee organizations will be better able to build management procedures around compliance. And, deposit and delivery standards will ease the participation of existing and new third-party contributors. This networking of responsibility will reduce costs and influence new market creation.

Finally, inter-institutional standards can serve as the foundation for new tools and services. For example, article deposit could be integrated into the grant management process; funding agencies would benefit from tools that revealed cross agency partnership opportunities; university's would profit from tools that highlight research output; and, researchers would gain from tools that created enhanced bibliographies and investigator profiles.

Comment 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?

Yes, other types of peer-reviewed materials resulting from publically funded research should be made readily accessible to the public. A successful and relevant public access policy must address all of the primary modes of communication for the funded disciplines. Access across these varied modes, will facilitate maximum impact and interdisciplinary discovery. However, the policies governing deposit compliancy should not create additional burdens for researchers and institutions. As mentioned above, policy must address practical and manageable compliance workflows.

Comment 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 the ideal world, there would be no delay. Free, immediate access will optimize scientific and commercial use. Faster access will facilitate more cutting edge science, derivative commercial services and market creation. For example, Houghton et al estimate that in contrast to a six-month embargo period, a zero embargo would increase incremental returns in R&D by \$120 million (NPV).⁴ Overall, studies investigating the citation advantage of open access articles, demonstrate at least a 25% lead.⁵

However, we acknowledge the position of those stakeholders, specifically publishers, who continue to rely on a subscription income. In these cases, limited embargo periods of no longer than 12 months have proven successful. The NIH relies on this timeframe, as do numerous international funders. At this time, we know of no studies or data demonstrating destructive consequences related to these polices.

To date the NIH open access policy has not altered Oregon Health & Science University's journal subscription buying patterns. It is not likely that a significantly reduced embargo period would change this trend, as our researchers need immediate access to this literature. Extending an NIH public access policy to all federal agencies and reducing the embargo period would significantly enhance our community's access to research results not covered by our Library's collection development scope and activities. This enhanced

access could bolster established interdisciplinary research and inspire new interdisciplinary opportunities.

1. Stern, S., Kolev, J., Dewatripont, M., Aghion, P., & Murray, F. (2009). Of Mice and Academics: Examining the Effect of Openness on Innovation. Research Paper or Report. Retrieved December 13, 2011, from <http://dash.harvard.edu/handle/1/4554220>
2. Battelle Technology Partnership Practice. (2011). Economic Impact of the Human Genome Project. Retrieved December 13, 2011, from http://www.battelle.org/spotlight/5-11-11_genome.aspx
3. ASL. (n.d.). Public Access to Federally-Funded Research. Retrieved December 13, 2011, from <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>
4. Houghton, J. (2010). Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs (SPARC). Retrieved December 13, 2011, from <http://www.arl.org/sparc/publications/papers/vuFRPAA/index.shtml>
5. Wagner, A.B. (2010). Open access citation advantage: An annotated bibliography. Retrieved December 22, 2011 from <http://www.istl.org/10-winter/article2.html>

Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Submitted by

John Willinsky, Khosla Family Professor of Education, Stanford University

Lauren Maggio, Director of Research and Instruction, Stanford University Medical Center, Stanford University

We work in schools of education and medicine, respectively, and are thus involved and deeply interested in the education of professionals, and the role that public access to relevant research can play in continuing professional development, improved practice, and, far more broadly, the general educational quality of democratic life. We are advocates, then, of greater access to knowledge, while being mindful of the need for both the quality controls of peer-reviewed scholarly publishing and the need for the financial investment that allows for such quality.

(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 publicly 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?

In terms of the benefits of an access policy to research, one consideration is the utilization of the research to improve the practice and knowledge base of professionals. We have recently conducted and published a [study](#) of physicians (N=90), in which a third expressed an interest in accessing primary research as part of their medical practice, with implications for improved quality and costs for health care. Furthermore, the physicians reported the weekly need to consult research was based on a need to inform their understanding regarding a specific patient, keep their practice up to date and to satisfy general curiosity. When interviewed on information use, one physician at a community clinic commented that her need for information is “constant” while another stated that he needs information daily. In our study, physicians were currently accessing information by making do with unreliable access through colleagues, the delayed assistance of librarians and the illegal hoarding and sharing of library passwords that provided them with such access.

(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?

We believe that it, above all, needs to be made perfectly clear in discussions of this matter that one of the principal intellectual property interests of scientists, federal agencies, and the public is the widest possible distribution of peer-reviewed publications, not only for the gratifying purposes of take-up, but also to ensure further evaluation and reassessment of published work. These interests appear to be currently, and perhaps needlessly, subordinated to the financial interests of publishers. While no one denies that the publishers provide a valuable service, for which they should be compensated, the question of whether this requires an exclusive monopoly over a public good. One indication that this might not have to necessarily be the case is the number of free, open access journals publishing today (upwards of 7,000 titles, according to the [Directory of Open Access Journals](#)), the majority of which do not charge either readers or authors, as they rely on institutional support and a greater extent of the academic labor that all journals receive at no cost, whether for content or peer review. At any rate, the very assertion of a public access *right*, on the part of federal funding agencies, makes it clear that the public has a stake and claim in this knowledge and that this public right should figure in the evolution of business models in the digital age.

(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?

We can see the value of the arguments on both sides of this question, but feel that what is unequivocally needed is federal support for arriving at centralized standards for metadata and other structural features, in consultation with research libraries, professional societies, and related federal agencies that are also in a position to advise on management of public access.

(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?

One of the most promising new models for public-private partnerships that has been discussed involves repositioning the research libraries as *partners* in scholarly publishing (rather than simply customers), given their expertise and capacities, and their willingness to invest in developing resources for public access. One step has been taken in this direction, judging by how the majority of such libraries now provide public-access repositories and hosting services for open access journals. This partnership or cooperative approach would be especially effective in working with professional societies and smaller publishers.

(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?

This would appear to be another opportunity for greater involvement of research libraries and librarians in setting metadata standards, as well as advising on interoperable search, as per response to #3. The model here might be the National Library of Medicine, with specialist librarians able to work with scholarly societies to create a similar level of standards, with the relevant federal agency support and endorsement.

(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?

A key element in keeping costs down is working toward, through federal mandates and the cooperation of publishers, a further integration of public-access provisions into the publishing process. This could involve, for example, automatic deposit on acceptance, along with all of the metadata, in a centralized repository, reducing the time and expense of separate deposit and multiple repositories.

(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?

Given the current standard (mandated by the NIH Public Access Policy) of providing public access to the “final draft” of the federally funded research, the published format should not be an issue. Although the NIH policy applies to *journal articles* alone, the focus on journals makes sense in medicine where the article is the standard. However, books and book chapters play a much bigger role in some fields, such as education, and they might reasonably fall within the mandate as well to honor the intention of making publicly funded research available.

(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?

One source of evidence that is available on the longer term effects of a zero embargo-time comes from arXiv.org, which since the early 1990s has been offering immediate pre- and post-print public access to virtually the entire literature in high-energy physics. The American Physical Society and Institute of Physics both reported to Alma Swan in 2005 that the arXiv.org repository did not appear to have resulted in any loss of subscriptions for these publishers. At the same time, the average subscriptions prices for physic journals remains second only to chemistry at \$3,252/a (2009) suggesting the sustainability of a healthy publishing industry in face of a zero embargo time in physics. The widespread use of arXiv.org among physicists can be attributed, in part, to a strong pre-print culture in physics dating back to the age of print. There is no reason why the rapid and open dissemination achieved by this field would not be advantageous to other fields. Again, a more collaborative, cooperative approach between research libraries and publishers could resolve and eliminate this embargo issue, as the libraries gain no advantage by having exclusive access to the literature, and thus may be willing to commit to supporting the

journals directly using existing allocations for subscriptions. This would guarantee the greater part of their revenue stream, while reducing subscription management costs on both sides. While this sort of change goes beyond the scope of public access policies, a move toward mandated public access to research is a strong first step in asserting that this knowledge was originally funded as, and is intended to stand as, a public good.

Hello,

I read the eight questions posed, very thoughtful. As an employee of a non-profit that specializes in energy and agriculture issues, I know access to research findings is critical. As a taxpayer, that access to publically financed research findings would be restricted is downright ridiculous.

USDA NIFA's SARE program seems to be do a good job of being clear about this matter, making sure results are available while still respecting the authors' other uses of their work.

Sincerely,

Al Kurki

Helena, Montana

TO: Office of Science and Technology Policy
FROM: Oya Rieger, Associate University Librarian, Cornell University Library, Ithaca, NY
EMAIL: rieger@cornell.edu
RE: Response to “Public Access to Peer-Reviewed Scholarly Publications”
DATE: January 12, 2012

Introduction

I am writing in support of governmental mandates that encourage scientists to share their research outputs, especially if the work is funded by the taxpayers. As the associate university librarian for digital scholarship and preservation services at Cornell University Library, I would like to offer examples from my own program area in support of the RFI, specifically related to the potential impact of open access for improving the scientific enterprise. I oversee the Library’s digitization, online repository, digital preservation, electronic publishing, and e-scholarship initiatives with a focus on needs assessment, requirements analysis, business modeling, and information policy development.

arXiv: An Open Access Success Story

The current publishing ecology has a diverse range of stakeholders including commercial and university publishers, scholarly societies, and libraries. Open access to social and scientific information is not mutually exclusive with commercial publishers. An excellent example of this dynamics is arXiv, which is the primary daily information source for hundreds of thousands of researchers in physics, and plays an increasingly prominent role in mathematics, computer science, and other related fields (Gingparg, 2011). With 700,000 e-prints, it provides an instant communication mechanism for scientists and complements the formal publishing process, which may take several months. Faster and unmediated access enables scientists, both in academic and entrepreneurial institutions, to incorporate new findings into their research faster. Since its launch in 1991, arXiv has achieved iconic status as an effective online distribution system and is often cited to illustrate digital repositories’ potential role in transforming scholarly communication. Such an impact is difficult to measure in financial terms due to its deep scholarly communication infrastructure roots.

One of the premises of arXiv has been making science more democratic by allowing for the rapid worldwide dissemination of scientific findings. The RFI is focusing on the U.S. economy an commerce; nevertheless, open access to scientific information has global implications without national boundaries. Figure 1 illustrates the international reach of arXiv. It is a global initiative, involving dedicated mirror sites in 17 countries and collaboration with U.S. and foreign professional societies and other international organizations. It has also provided a crucial life-line for isolated researchers in developing countries. Most scientists and researchers who post content on arXiv also submit it for

publication in traditional peer-reviewed journals. However, famously reclusive Russian mathematician Grigori Perelman's decision to post his proof of the 100-year-old Poincaré Conjecture solely in arXiv underscores the repository's increasing importance and its role in transforming scholarly communication.¹ We are in the process of parsing the 24% usage shown as “other” in Figure 1 but our early analysis indicates that there is strong use by commercial entities.

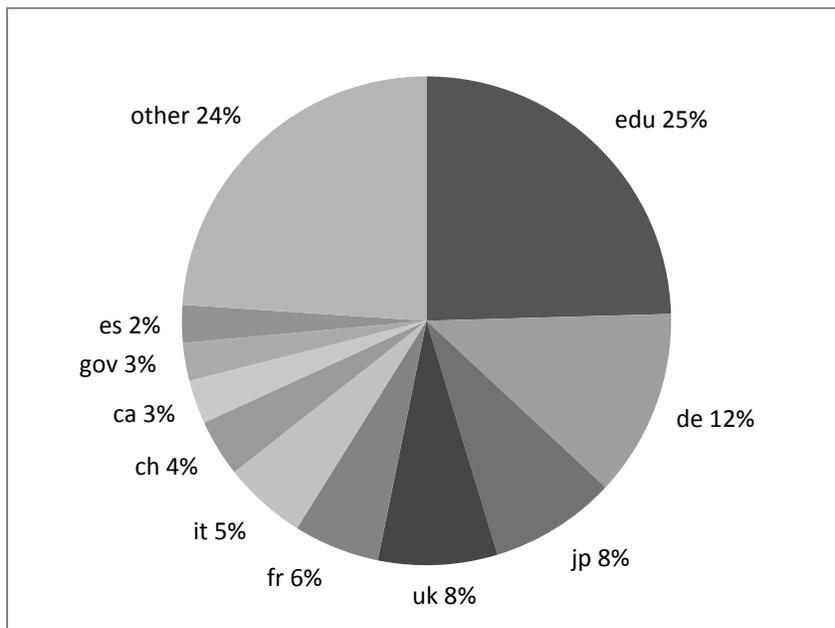


Figure 1: arXiv institutional downloads at main site by Internet domain of institutions (2010)

arXiv does not compete but co-exists with commercial publications. There is a unique role for each mode of dissemination. arXiv facilitates rapid and global dissemination of research results (*place to stake* intellectual precedence claims) whereas commercial journals continue to provide a venue for peer review and tenure requirements. In celebration of the arXiv's 20th anniversary, on September 23, 2011 Cornell University Library (CUL) hosted a meeting at Cornell with the representatives from several publishers and societies, including Elsevier, Wiley, and Springer. The goal of the forum was to discuss the feasibility and desirability of establishing a research and innovation collaboration in

¹ The Poincaré Conjecture is the only one of seven famous mathematical problems identified by the Clay Mathematics Institute that has been solved. For this work, Perelman was awarded the prestigious Fields Medal (which he declined) and in 2006 the journal Science named his proof of the Poincaré conjecture as its annual Breakthrough of the Year.

support of arXiv. The commonality among all the participants is their interest in understanding and meeting the needs of the scientific community.

arXiv is an example of a cost-effective open access delivery method. Since it moved to Cornell in 2001, the Cornell University Library has provided the bulk of arXiv's operating costs, which are projected to be approximately \$500,000 for 2011.² In January 2010, Cornell has established a voluntary institutional contribution model and invited pledges from the top 200 libraries and research laboratories accounting for more than 75 percent of annual institutional downloads (Rieger, 2011). Based on a budget of \$330,000 and 40 million paper downloads for 2010, each e-print costs merely 0.08 cents per download and the cost per submission is \$4.70.

Digital Preservation Through Open Access

Scholarship has been created and sustained through the interoperation of three key agents that have formed an infrastructure for sustainability:

- Scientists and their academic and professional networks
- Publishing organizations including scholarly societies
- Libraries and archives

This network has been changing due to the new modes of digital scholarship. Academic libraries are increasingly dependent on commercially-produced, born-digital content that is purchased or licensed. E-journals have replaced the majority of titles formerly produced in paper format. Cornell University Library spends more on e-materials than other forms of content. The finding of a recent Cornell and Columbia University Libraries study that assesses the role of LOCKSS and PORTICO in preserving each institution's e-journal collections was alarming (LOCKSS Team, 2010). Although LOCKSS and PORTICO are considered successful digital preservation initiatives, only 15-20% of the e-journal titles in the libraries' collections are currently preserved by these two initiatives.³ From users' perspective, there is an implicit assumption that today's electronic journal content will be refreshed and digitally manipulated as required to carry it forward indefinitely over time.

As we move to greater dependency on digital content, we must rethink how we go about managing our preservation responsibilities. An important benefit of open access mandates is supporting enduring access through redundancy of e-prints available from different

² The arXiv budget for 2011 is available at http://arxiv.org/help/support/2010_budget. It is based on an estimate and will be updated throughout the year to reflect the actual expenses.

³ Information about LOCKSS and PORTICO is available at http://www.jisc.ac.uk/publications/briefingpapers/2007/pub_ejournalspreservationbp.aspx

sources. For instance, if an article is published only in a born-digital commercial journal, the long-term accessibility of the work is solely based on the preservation provisions put in place by the publisher. Unlike journals with print and digital version, for born-digital articles, there is not a physical copy that will be archived and maintained by libraries and archives. Whereas an open access journal article is much more likely to be deposited in multiple repositories and therefore creating a security network through redundant digital copies maintained through different systems. So, unlike the proprietary publisher scenario, there is not a single point of failure and the risks are distributed.

The long-term viability of digital scholarly content managed solely by publishers' is uncertain. They are entrepreneurial entities with focus on return-on-investment. Due to the volatility in the publishing market, commercial publishers should not have the sole responsibility for providing long-term access to the output of the publicly funded research. When a publisher goes out of business, it is very likely that the content they manage will become inaccessible. To ensure that federally funded scholarly research outputs are permanently available online, there must be federal mandates to enable and authorize depositing articles to open access repositories. NIH Public Access Policy is exemplary as it requires that grantees submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.

To sum, although it is ideal to separate the issues of access to peer-reviewed research and digital preservation of the published intellectual and cultural record, these two areas are often intertwined with co-dependency. Preservation of the published versions of research outputs requires digital preservation infrastructures and services. However, due to the lack of a scalable and reliable preservation infrastructure for born-digital scholarly articles, the retention of prints and post-prints through institutional repositories or centralized subject-based repositories constitutes a critical archival strategy through redundancy and multiple copies.

Creating Sustainable Infrastructure for Open Access

Open access model allows transparency and accountability about how knowledge is created, verified, analyzed, and interpreted. Although I am supportive of encouraging free and open access wherever feasible, federal directives requiring uninhibited discovery of information is not an end in itself. Open access mandates from federal and state agencies are fundamental enablers; however, there needs to be a sustainable infrastructure based on the following principles:

- Scalable and cost-efficient databases to store, discover, access, and re-purpose information.
- Management policies and procedures in place to ensure the enduring usability, authenticity, discoverability, and accessibility of content over the very long-term (digital preservation).
- Interoperability arrangements that link a given repository to related systems, services, and communities.
- Leveraging existing technical and policy infrastructure to minimize unneeded duplication of efforts and support cost-effectiveness and long-term sustainability.
- Features that support supplementary information objects such as underlying data, auxiliary multimedia content, and research methodologies.
- Flexibility to accommodate different embargo periods in support of professional, academic, and entrepreneurial requirements of scientists and research institutions.
- Incentives, rewards, and recognition for scientists who share and archive the outputs of their research endeavors.
- Support for IPR, privacy, and confidentiality – especially if research entails human subjects or other sensitive and potentially misleading information.
- Functionality and arrangements that lower barriers for scientists to contribute content to multiple complementary repositories.
- Sensitivity to disciplinary cultures, practices, norms, and aspirations.
- Active participation in the development of community standards for deposition, use, and maintenance of scholarly information.
- Provisions for collaboration among communities and information types (e.g., articles, data, images) in order to encourage interdisciplinary scholarship.

In closing, I also urge the White House to adopt a more inclusive definition of research to include the work of humanities scholars. This will require that in addition to scientific organizations such as NSF and NIH, agencies such as NEH is also involved in developing open access policies. Interdisciplinary mandates will be instrumental in facilitating and

promoting collaboration among all kinds of scientists, including the humanities and cultural heritage community.

References

Ginsparg, P. (2011). arXiv at 20. *Nature*, no. 476, p. 145–147,
<http://www.nature.com/nature/journal/v476/n7359/full/476145a.html>

LOCKSS Assessment Team (2011). Final Report of the 2CUL LOCKSS Assessment Team, Cornell University Library and Columbia University Library,
<http://2cul.org/sites/default/files/2CULLOCKSSFinalReport.pdf>

Rieger, O.Y. (2011) Assessing the Value of Open Access Information Systems: Making a Case for Community-Based Sustainability Models. *Journal of Library Administration*, 51(5-6), 485-506.
http://bing.exp.sis.pitt.edu:8080/webdav/new_documents/rieger_open_access_value.pdf

Dear OSTP,

As both a librarian, a health consumer, and a taxpayer, I am strongly opposed to the Research Works Act. Taxpayers must not be made to pay more than once for the dissemination of research we have already supported. Please do not allow publishers to exploit US taxpayers. The NIH public access mandate has furthered the reach of health science.

Rather than repealing and/or limiting further dissemination of scientific information, it should be expanded.

Thank you.

Bette Anton

Bette Anton

Head, Fong Optometry & Health Sciences Library University of California

BEFORE THE OFFICE OF SCIENCE AND TECHNOLOGY POLICY

REQUEST FOR INFORMATION CONCERNING

PUBLIC ACCESS TO PEER-REVIEWED SCHOLARLY PUBLICATIONS

RESULTING FROM FEDERALLY FUNDED RESEARCH

January 12, 2012

My name is Ali Sternburg and I am a third-year law student at American University Washington College of Law. My views are influenced by my study of intellectual property law and policy and its history, my position on the Executive Board of the *American University Intellectual Property Brief*, an online, Creative Commons-licensed publication, and my role on the Steering Committee of the Right to Research Coalition. I write on my own behalf.

As a law student, I admit, I don't frequently read scientific articles. However, as a lawyer I may represent scientists, doctors, patent holders, entrepreneurs, and many others who do rely on scientific information—information that I would need to help them. As a student (both literally, for the next few months before I graduate this May, and figuratively, for the rest of my life as I continue to seek and share knowledge and information), I am concerned about the priorities of some policymakers who favor the private interests of certain publishers over the interests of the broad American public in research and knowledge; the problem is compounded by the fact that American taxpayers have funded this research. This information should be made publicly accessible, and I applaud the Office of Science and Technology Policy for seeking input from the public on this important issue. My generation—the future leaders of our country—must be properly educated and have the tools to innovate and create jobs, and the skills to be hired by those who have created jobs.

During law school, I have worked on two U.S. Supreme Court *amicus* briefs, in which we cited scientific articles and information. I worked on an *amicus* brief on the merits stage of *Sorrell v. IMS*, 131 S. Ct. 2653 (2011), a case which discussed prescription drugs and health care policy, and I am currently working on an *amicus* brief to grant a writ of *certiorari* in [a case where the cert petition has not yet been filed, so I cannot disclose the case name] which considers DNA molecules and human genes. Obtaining articles for these briefs showed me the challenges and costs of accessing specialized scientific information, in addition to the widely interdisciplinary nature of legal research.

In addition, I plan to work in support of the public interest. This means I am deeply invested in the public benefit that must be balanced with all private rights, especially in fields like intellectual property. This also means that I may not always be able to afford access to expensive paid resources. It is inconceivable for me and other members of the American public to not have access to the research that our tax dollars help fund as an investment for our future.

Also, the timing of this RFI allows me to briefly voice my strong opposition of H.R. 3699, the “Research Works Act,” introduced on December 16, 2011. This bill would prohibit Federal Agencies from conditioning their grant funding to require that all members of the public be guaranteed online access to the products of the research that their tax dollars fund; it essentially is aimed at reversing the highly successful National Institutes of Health (NIH) Public Access Policy.¹ Rather than impede access to these resources, as this bill would, the Government should actively ensure that students and the general public get the full benefit of our collective investment in science, a recognition that this RFI makes clear. The NIH and other Agencies must be allowed to ensure that taxpayers get timely, public access to the results of research funded with taxpayer dollars. The NIH policy should be expanded to other Agencies, rather than being reversed by sponsors of the Research Works Act.

(1) Are there steps that agencies can take to grow existing and new markets related to access & analysis of peer-reviewed publications? 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?

Comment 1:

Agencies certainly have the ability to enhance the market for access to and analysis of peer-reviewed publications. The most important step is to make the information freely and widely available, which encourages further research and collaboration. Making information available for free and without restrictions does not mean that it cannot be monetized and commercialized in the future; in fact, research has shown just the opposite.² The ability to access and reuse articles enables innovation, by individuals and companies, to build products and services, using content funded by the public to serve the public. Providing broad availability and allowing full utility of this information encourages innovation and development in diverse industries – from the biotech sector to pharmaceuticals to renewable energy to even the publishing industry.

This type of access is called Open Access. “Open Access (OA) is the free, immediate, unrestricted availability of high-quality, peer-reviewed scholarship over the Internet – combined with the rights to use this information to its fullest possible extent.”³ Open Access ensures that more students and more people in general – including particularly those who currently cannot afford access otherwise – not only stay informed of cutting-edge ideas, but also discover new uses and applications for research. Providing faster access allows ideas generated to be incorporated into development cycles more quickly,

¹ See also Michael B. Eisen, *Research Bought, Then Paid For*, N.Y. TIMES, January 10, 2012, available at <http://www.nytimes.com/2012/01/11/opinion/research-bought-then-paid-for.html>.

² See, e.g., Mike Masnick, *The Grand Unified Theory On The Economics Of Free*, TECHDIRT, May 3, 2007, available at <http://www.techdirt.com/articles/20070503/012939.shtml>.

³ The Right to Research Coalition, *The Solution: Open Access*, <http://www.righttoresearch.org/learn/solution/index.shtml>.

speeding the launch of new services and products into the marketplace, stimulating economic growth, and creating new job opportunities across broad sectors of the economy. The complete collection of articles resulting from publicly funded research must be made freely accessible, so that the public can fully use them (e.g., text mine, data mine, compute on them, create derivative works, etc.) without restrictions.

Policies for archiving publications and making them publicly accessible will directly grow the economy and improve the productivity of the scientific enterprise. Open Access to research articles is a critical driver of scientific innovation and productivity. Open Access lets people get to – and read more – information than they previously could. This is enhanced by new tools for incorporating more articles into research faster, including machines as a new category of readers and users, and leads to vast, previously unobtainable and unrealized ideas and connections. Opening access to the widest possible audience encourages contributions and citations by more minds, growing societal and institutional knowledge, and ultimately aiding this country.

Open Access allows research results to be quickly incorporated into the teaching and learning process – improving the quality of education quickly and cost-effectively. Professors can only teach what they have access to, with the most disparate impacts often felt in the regions that need intellectual advancement the most. Providing American students with the most complete, up-to-date education possible boosts U.S. economic competitiveness, especially in innovative, cutting-edge fields. Today's students will build the foundation of tomorrow's economy – Apple and Google were both started by entrepreneurs the age of today's current undergraduate and graduate students. If students don't have full access to critical publicly funded research, we're potentially missing out on innovative breakthroughs that could create jobs and be built into the next Apple and Google. Open Access helps students get projects off the ground and build businesses around their research. Losing access to the relevant literature is a significant barrier for students who might consider dropping out of school to start a business around their research. When students graduate, they lose access to the vast majority of research that is subscription-access only. This impedes students' ability to stay current in their field and hinders their ability to hit the ground running when they put their education to work. This cost is even greater in a weak economy such as the present, where students may spend a significant amount of time in their job search.

The relative costs of Open Access policies are minimal compared to the vast public benefit. The NIH Public Access Policy costs approximately \$4 million per year out of a \$30 billion budget, an investment of less than 1/1,000th of 1% that results in access to all NIH-funded research, which is used by more than 500,000 unique users per day through PubMed Central.⁴ According to a 2010 study, an expansion of the NIH public access policy to cover all federally funded research with a six-month embargo period would provide a 500% return on investment to the U.S. Government, generating benefits eight

⁴ Letter from Dr. Francis Collins, Director of the NIH, to Representative Joseph Pitts, December 2011, available at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

times greater than costs, resulting in a net present value gain worth approximately \$1.5 billion.⁵ Open Access is thus an excellent return on investment.

We need full Open Access (free, immediate, unrestricted availability of high-quality, peer-reviewed scholarship online, with the broadest possible information reuse policy), in order to create the environment that will improve students' educations, maximize scientific productivity, accelerate commercial innovation, and reinvigorate the U.S. economy. Restrictions that limit how we can access and use the scientific research we paid for limits the value and the return to American taxpayers. Broad reuse allows researchers to continue to find and add value from this public investment, now, and in the future, without having to duplicate research.

Students should be guaranteed Open Access to cutting-edge research upon which their education depends, and have the ability to advance scientific discovery and stimulate innovation in all scientific disciplines. Immediate, Open Access provides students with the most up-to-date education; anything less limits students'—and likely professors'—knowledge, stifling U.S. innovation and economic competitiveness.

(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 the same?

Comment 2:

Article I, Section 8, Clause 8 of the U.S. Constitution says “Congress shall have the power . . . To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors, the exclusive Right to their respective Writings and Discoveries.” Over time, copyright has expanded in scope, subject matter, and duration, generally at the interest of owners of existing content who are threatened by technology. Copyright in new works currently lasts for the life of the author plus 70 years. The public should not have to wait until copyright terms expire to have access to knowledge, especially given the rapid pace of development.

The fair use doctrine (17 U.S.C. §107) and other exceptions and limitations to copyright are available, but they rarely extend to copying entire works, even when it is for educational purposes. The fair use doctrine can be invoked as a defense to violating one of the copyright holder's exclusive rights under 17 U.S.C. §106—reproduction, distribution, making a derivative work, public performance, and public display. The right to make a derivative work extends to uses that build upon the work, transforming its context and adding value, such as making it searchable, machine-readable by new devices, translating it, downloading and analyzing data, or making other adaptations. This can be justified under fair use, but not always, especially when there is economic

⁵ Houghton, et al., *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*, July 2010, p. 7-8, available at <http://www.arl.org/sparc/bm~doc/vufirpaa.pdf>.

gain involved. The risks involved may stifle innovation and research, which harms everyone. Therefore, permissible actions with information should be broader than fair use. For instance, open licenses developed by Creative Commons permit users to do more with works than is allowed under copyright. In addition, information should not be locked down with Digital Rights Management and Technological Protection Measures that don't even allow for potential fair uses. Full use rights (e.g., distribution, reuse, text mining, data mining, computation, creation of derivative works, etc.) must be an integral part of a government-wide public access policy.

The publication I work on, the *American University Intellectual Property Brief*, publishes articles online under a CC-BY license, and we permit authors to retain their intellectual property rights and publish in other journals if they wish. To illustrate, below is an excerpt of an email from our Senior Articles Editor:

We currently publish the IP Brief under the Creative Commons Attribution 3.0 United States License. Anything that we jointly produce and publish will be able to be dispensed in print or online by anyone else as long as there is an attribution to you and to the American University Intellectual Property Brief. As far as you and the IP Brief go, you will always be free to publish your unedited work (the version you submitted to us). If you choose to update your article at some point in the future and you keep edits that we will work on over the semester, the [CC-BY license] will apply. But you could take that new article anywhere and publish it with anyone you would like. I hope that works for you. If you have concerns, we can always work out something else. We would be happy to do that.

(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?

Comment 3:

If the publications are decentralized among different sources, whether private or public, the Federal Government should have an accessible, mirrored repository that includes all articles and other content. This may be accomplished through an archive, which is put online and accessible to be used by all; all necessary rights must be given to the Federal Government for this purpose. Such an online resource should consider open-source programs and licenses, interoperability, accessibility for the disabled, translations, searchability, and other technological concerns, so that this an archival resource that will continue to be useful in the future.

(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?

Comment 6:

A successful policy will be easily implemented and consistent. The NIH policy can serve as a model, in which researchers consent at the time of grant acceptance to make their work freely accessible in PubMed Central. In addition, Agencies can require articles resulting from their funding to be made available under an open license, such as the Creative Commons CC-BY license.

(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?

Comment 7:

Ideally all peer-reviewed publications resulting from federally funded research should be made available to the taxpayers who funded them, which in turn allows them to be read by more people than just the few attendees privy to the meeting. Why keep it locked away? After all, don't researchers share research in order for it to be read and improved-upon by their peers? Another point to note is that this public access policy for other types of publications should be separate from the general public access policy for journal articles, due to inherent differences.

(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.

Comment 8:

The ideal embargo period is to not have an embargo at all. Students should have Open Access to the latest developments and research for their own educational benefit, for the benefit of their future employers, and ultimately for the benefit of this country. The U.S. Government should not sacrifice the education of its citizens in order to please the publishing industry and their lobbyists and the congressional campaigns they fund. An academic semester is generally 3 to 4 months long, and so the length of the embargo period (e.g., the difference between 0 months and 6 months and 12 months) can have a significant impact on what is taught and learned. Finally, there has been no evidence presented by any publisher that the NIH public access policy harms its business, which provides strong empirical proof that public access does not harm subscription-based publishers. This essentially means it does not harm anyone, while helping everyone.

Office of the President
Howard Hughes Medical Institute
4000 Jones Bridge Rd.
Chevy Chase, MD 20815-6789
T. 301.215.8646
F. 301.214.8863

January 12, 2012

To: Office of Science and Technology Policy
Executive Office of the President
725 17th Street Room 5228
Washington, DC 2050

From: Robert Tjian
President, Howard Hughes Medical Institute (HHMI)

Re: HHMI response to the White House RFI on OA publications

Via: Email, publicaccess@ostp.gov

We welcome the opportunity to contribute to the ongoing dialogue about how best to leverage the nation's investment in research and technology.

The Howard Hughes Medical Institute, the nation's largest non-profit funder of basic biomedical research, has long believed that society benefits when the fruits of discovery are made broadly available to the public through the published literature. That commitment has been manifested through the policies that govern publishing by Institute scientists, as well as our more recent partnership with the Wellcome Trust and the Max Planck Society to create a top tier open access journal in the biomedical and life sciences.

HHMI's policy, announced in 2007, requires our more than 400 scientists to publish in only those journals that make the contents freely available within six months of publication. It extends other policies that require HHMI scientists to share published research materials, databases, and software in a timely and useful fashion.

HHMI has designated PubMed Central (PMC), the digital archive of biomedical and life sciences literature maintained by the National Institutes of Health, as the repository for journals in the biological sciences. Articles published in journals that are outside the biological sciences are expected to be deposited in comparable repositories and made publicly available within six months.

The new open access journal – which will be called *eLife* and is slated to launch later in 2012 – takes this commitment one step further. The founders have agreed that the journal will utilize the Creative Commons Attribution 3.0 license (known as CC BY 3.0) so that the contents can be shared without restriction.

We fully appreciate the impact that the NIH policy has had on the sharing of scientific information and believe that it provides a useful starting point for other federal science and technology agencies seeking to enhance the public's access to scholarly information – including published articles and relevant data. PMC is a particularly valuable resource to the scientific community in this nation and around the world.

Both the research community and the wider society benefit when information is widely shared. That is the goal of publication in the first place – experiments are incomplete until the knowledge is shared – and currently available technology enables us to take that a step further by disseminating research results through publicly accessible repositories. There's also increasing evidence that sharing information also makes good economic sense because it fuels innovation and further discovery – outcomes that the American taxpayer has a right to expect.

Response to the Office of Science and Technology Policy public consultation on Public Access to Federally Funded Research

Submitted by the international Confederation of Open Access Repositories (COAR)

January 12, 2012

Introductory Comments

We would like to thank the White House Office of Science and Technology Policy for initiating this important consultation on public access to research outputs. The Confederation of Open Access Repositories (COAR) is a not-for-profit association of repository initiatives that was launched in October 2009. We represent 59 institutions in 23 countries from throughout Europe, Latin America, Asia, and North America. Our mission is to enhance greater visibility and application of research outputs through global networks of Open Access digital repositories. Our aim is to support the implementation of policies of governments, research funders and institutions. More information about COAR can be found on our website: <http://coar-repositories.org/>.

Current research dissemination practices do not adequately meet the needs of all stakeholders – especially the public who has funded much of this research through their taxes. Millions of policy makers, clinicians and practitioners, small businesses, students and educators, patients and their families, and others are without ready or affordable access. With the Internet comes the opportunity and the imperative to share these results widely so all citizens can access, use and build upon research results in new and innovative ways. (i)

In order to improve access and maximize investments in research, governments around the world are implementing policies that require the free availability of research results. The SHERPA-JULIET service in the UK, which monitors funding agency policies, now lists over 70 funding agencies with open access mandates from over 20 different countries. In Europe and elsewhere, open access is now acknowledge as an effective and low cost way to improve research impact and the efficiency of the scholarly communication system. (ii)

COAR's response to several of the questions contained in the Request for Information are as follows:

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?

There are two important steps that governments can take to gain further economic benefits from their investments in research and the peer-reviewed publications that result:

1. Implement policies that require researchers to deposit their articles into open access repositories.
2. Support the development of a national repository network for the collection and preservation of research outputs that is interoperable with the growing international network being developed.

Open access policies and infrastructure will enable the public, practitioners, industry, and others to make use of the valuable information contained within the peer-reviewed literature. Currently, this literature is available only to researchers who are affiliated with institutions that can afford to subscribe to these journals. Therefore there is tremendous unrealized potential for this content to be further used and exploited for the development of new products, practices and policies.

Similar to open data initiatives, open access to peer-reviewed publications will enable others to build effective value added services on top of the content. It is possible to envision the development of numerous value-added tools, such as discovery and indexing services, as well as data mining and text analysis technologies. These value added services will allow for new connections and discoveries, and lead to further scientific discovery, innovation and product development.

What are the relative costs and benefits of such policies?

There are costs associated with open access, such as staff and hardware costs for running repositories, however, these costs represent only a small portion of a nation's investment in research.

Economic analyses have shown that national approaches requiring open access

to publicly funded research papers open access system would result in significant cost savings, in comparison to the current subscription based system. A study conducted by Houghton et al. concluded, for example, that, "(s)haring research information via a more open access publishing model would bring millions of pounds worth of savings to the higher education sector as well as benefiting UK." (iii) In the three national studies of Denmark, the Netherlands and the UK, the costs and benefits of scholarly communication were compared based on three different publication models. All three concluded, "the greatest advantage would be offered by the Open Access model", via open access repositories." (iv) The study found that open access could lead to an "annual savings of around EUR 70 million in Denmark, EUR 133 million in The Netherlands and EUR 480 in the UK. Other analyses undertaken in Australia and the US have come to similar conclusions. (v)

In addition, much of our modern economy is already based on the free availability of information. Google, Facebook, Twitter are just a few examples of new services that have been developed because of the openness of information in the digital environment.

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?

There are advantages and disadvantages to centralized and decentralized approaches. The decentralized approach, such as a network of university repositories, ensures that the locus of deposit for articles is close to the working environment of the authors. A more centralized approach, such as PubMed Central, allows for the full corpus of literature in one field can be found in a single database. Ultimately, the best approach will likely depend on the history and traditions of a given discipline.

Many governments in Europe and elsewhere are adopting decentralized approaches by implementing networks of institutional repositories to make available the publicly funded research outputs. Countries in the European Union have benefited from two European Commission Seventh Framework Program (FP7) projects, DRIVER and DRIVER II (vi), which has funded the establishment and development of a European open access repository infrastructure. The projects provided funding at the national level to implement repositories, support for national help desks that provide expertise to repository developers, and also the development of a centralized search portal. The project ended in 2009, and the central portal, called DRIVER Search Portal, is now being maintained

collectively by national partners. It currently provides free access to over 5,790,000 research publications from 319 repositories in 43 countries. (vii)

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?

There are very important reasons why federal governments would want to maintain copies of their nation's research output. This collective content represents the official record of the world's knowledge and is a valuable publicly funded asset. While libraries have traditionally been the custodians of the scholarly literature, this is no longer the case in the digital environment. Yet, there are no other types of institutions currently with a mandate to ensure research papers are preserved and accessible to scholars and the public. There are numerous roles that private sources could play in ensuring the preservation of research outputs, however private industry is subject to the whims of the market and stockholders. Only stable institutions, such as universities, libraries and governments, that have a specific mandate to preserve, can be relied upon to ensure ongoing access over the long term.

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?

Ideally, all outputs from publicly funded research should be made openly available to the public. However, there is a need to develop the infrastructure to support open access in conjunction with implementing such policies.

In Europe, OpenAIREplus (2nd Generation of Open Access Infrastructure for Research in Europe) was launched in Pisa in early December. The 30 month project, also funded by the EC 7th Framework Programme, will work in tandem with OpenAIRE, extending the mission further to facilitate access to the entire Open Access scientific production of the European Research Area, providing cross-links from publications to data and funding schemes. This large-scale project brings together 41 pan-European partners, including three cross-disciplinary research communities. (viii)

Creating a robust, participatory service for the cross-linking of peer-reviewed scientific publications and associated datasets is the principal goal of OpenAIREplus. As scholarly communication touches upon many disciplines, the

project's horizontal outreach will facilitate collaboration across data infrastructures, providing information to scientists, non-scientists as well as to providers of value-added services. The project will establish an e-Infrastructure to harvest, enrich and store the metadata of Open Access scientific datasets. Innovative underlying technical structures will be deployed to support the management of and inter-linking between associated scientific data.

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?

Policies should require that articles be deposited immediately upon publication, and made accessible within a 6-months of publication. The optimal scenario is that papers are made available immediately upon publication. However, in general a 6-month delay is acceptable in order to allow publishers maintain a revenue stream for their journals. A delay of access beyond 6 month would decrease the value and impact of the public access policy.

Publishers will adapt their business models to accommodate any requirements imposed via these policies. and already are, adapt to the new open access requirements being imposed by funding agencies around the world. Many of the large publishers now offer an open access option for publication, and

Please identify any other items the Task Force might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.

Based on the previous experiences of other agencies around the world, we maintain that the following components are necessary to ensure compliance:

- Policies must be mandatory. The very low deposit rates of NIH funded researchers in response to the NIH voluntary policy demonstrated the need for a mandatory policy (ix). This was also exposed in a 2005 survey of UK researchers found that study which found that about 15% of authors are self-archiving voluntarily, but 95% indicated that they would self-archive if their institutions and/or funders mandated it.
- Policies must be monitored for compliance. Compliance with a public access policy should be attached to any future funding decisions. There are ways of monitoring this, through the use of grant numbers inserted into the metadata of the deposited papers. Grant numbers would then be searchable and granting agencies would hypothetically be able to glean

other valuable information related to funding decisions.

- Policies should be consistent across agencies. Researchers are often funded through multiple research agencies. In a global research context, it is increasingly problematic to have a wide variety of access policies with differing requirements of researchers. A consistent, nation-wide approach would cut down on confusion and greatly improve compliance levels. In addition, a uniform nation-wide approach to public access policies in the US would also be helpful for publishers in developing more consistent self-archiving policies.
- Complying with a public access policy should not be onerous for authors. Repositories can assist with deposit and much of the deposit procedures can be automated. For example, the SWORD protocol has developed a standard deposit mechanism that could be used for to simultaneous deposit into repository and publisher. (x) In addition, most repositories have the ability to embargo access for a given length of time.

References

- i. Text adapted from the Open Access Scholarly Information Sourcebook: http://www.openoasis.org/index.php?option=com_content&view=article&id=547&Itemid=265
- ii. <http://www.sherpa.ac.uk/juliet/index.php?sortby=country>
- iii. Houghton and Oppenheim *et al.* 2009. Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark: http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fdownloads%2fOA_What_are_the_economic_benefits_-_a_comparison_of_UK-NL-DK_FINAL_logos.pdf
- iv. Houghton and Oppenheim *et al.* 2009. Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark: http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fdownloads%2fOA_What_are_the_economic_benefits_-_a_comparison_of_UK-NL-DK_FINAL_logos.pdf
- v. Models: Exploring the costs and benefits. JISC, January 27, 2009.

- vi.** DRIVER I and DRIVER II. www.driver-repository.eu/
- vii.** <http://search.driver.research-infrastructures.eu/><http://www.openaire.eu/>
- viii.** <http://www.openaire.eu/>

Virginia Tech's University Libraries responds to the Office of Science and Technology Policy Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

<http://www.gpo.gov/fdsys/pkg/FR-2011-11-04/html/2011-28623.htm>

(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 publicly 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?

Academic libraries are clearly a stakeholder in necessary services such as archiving and preservation and, therefore, play a key role in enabling access to information and knowledge that can lead to economic growth and productivity. Libraries like Virginia Tech's play an important role in knowledge production; it is critical that our government agencies adopt policies that enable and support unrestricted public access, discovery and re-use of publications resulting from federally funded research.

Successful library services such as hosting open access electronic journals and commercial article databases provide concrete evidence that access to validated scientific research through peer-reviewed articles can occur simultaneously. There are benefits to both mechanisms and they should continue to coexist, making it possible to develop economic growth opportunities. Use and reuse of the nation's intellectual capital is built into our fair use guidelines and these should be allowed to evolve and match pace with evolving scholarship and communications technologies and applications.

(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?

While it is perhaps easier to have centralized access, there is potential for a single point of failure issues. Conversely, decentralized services enable a variety of capabilities but may increase coordination costs. University libraries are used to collaborating to combine the best of both, working with both the commercial and non-profit sectors to ensure continuous access and long-term preservation.

(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?

It is essential that we create new as well as maintain well-functioning discovery services that are interoperable across computer platforms. A key to this success will be metadata, especially when it contains information necessary for archiving and preservation as well as the potential development of new services. Persistent identifies for authors, publications, and links to data are also essential.

(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?

Maximizing the benefits of public access and investments in peer-reviewed literature should begin with mandates for depositing publications at every federal agency providing public funding for research. NIH set the example with PubMed Central. Since not every federal agency has such a system in place, the Library of Congress could be provided the funding to establish a centralized service for other agencies. Another possibility would be to centralize access through the LoC but distribute the funding support to the public university libraries since they already demonstrated these capabilities through their digital libraries and institutional repositories.

(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?

All peer-reviewed publications and gray literature resulting from federally funded research should be publicly accessible. We would also like to see support and encouragement for these publications to evolve beyond text and to be open to expression in new forms of media that use evolving technologies.

(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?

Embargo periods are borrowed from traditional publication mechanisms and should not be necessary when the public funded the research through taxpayer dollars. Embargo periods can disappear when the research becomes publicly available prior to the academically formatted and peer reviewed publication. The evidence already exists that public access models can coexist with scholarly as well as commercial publications. The prime example is ETD--electronic theses and dissertations that are publicly accessible and coexist with derivative and value-added commercial publications such as articles and books.

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

Document Citation: 76 FR 68518

Document Number: 2011-28623

From: Sage Bionetworks
1100 Fairview Ave. N.
Seattle WA 98109

Summary

Traditional business models for scholarly publication are a barrier to both public and scientist access to federally-funded research in the internet era. Agencies should require open access publication of federally-funded work as a condition of support.

Publishers provide a service that is fueled by researcher time, reviewer time, page charges and subscriptions; all of which are funded directly or indirectly by the federal government. All taxpayers should have access to the scholarly work they pay for.

It is of particular concern that Congress is considering H.R. 3699 that would establish anti-competitive protections when there are already many open-access alternatives available from both traditional publishers and innovators.

Specific Comments on RFI:

- (1) Funding agencies should assure that supported research is published through sustainable channels with public access. Current revenue problems at universities and research institutes have often resulted in the cancellation of many publication subscriptions based on cost rather than value demonstrating that traditional publications business models provide no archiving security.
- (2) Publishers have legitimate interests in operating a profitable service but the intellectual property rights should remain with the creators and the funders.
- (3) The federal government investment in central open archives such as Medline has paid huge dividends for medical research. It is, however, a large cost and there may be value in piloting a program to establish interoperability among multiple credentialed open access repositories.
- (4) The science of indexing and searching scholarly information is well advanced and support for existing initiatives as well as the exploration of new innovations would be a good investment.
- (5) While it is attractive to have a central authority establish clear standards and criteria for archiving and search, marketplace demand may well determine the specifications.
- (6) Subscription fees have unfortunately become prohibitive for many institutions and thus a move to front end payment and efficient electronic review and publication would be a large cost and capital savings for researchers, libraries and institutions.

(7) yes.

(8) There is no reason to have an embargo period in open access publishing models. Any delay slows down the progress of research

Context: The academic publication business is a big, rapidly changing marketplace. The very high prices that traditional publishers now must charge for physical or electronic access have severely restricted access for both the public and for an increasing number of researchers associated with institutions that can not afford the subscription fees. The result is that often neither the scientific colleagues for whom the publications are created or the taxpayers who paid for the research and its publication can access them.

The success of recent open access options that take full advantage of internet communication demonstrate that alternative and more cost-effective business models are legitimate and sometimes superior. Indeed, many traditional publishers have been leaders in open access models that use increased page charges and electronic distribution to offer full access to all interested researchers and the public.

There is no need for legislation such as House Bill 3699 to create anti-competitive rules for funding agencies in order to subsidize traditional business practices.

Submission prepared by;
Jonathan Izant PhD
Vice President, Sage Bionetworks

About Sage Bionetworks: Sage Bionetworks is a 501(c)(3) nonprofit biomedical research organization created to change how researchers approach the complexity of human biological information and the treatment of disease.

Sage Bionetworks' mission has five interdependent themes:

- Research on computational network models of disease
- Pilot projects trialing disruptive models of research cooperation
- Rules and rewards that promote data sharing and collective research
- Building the computational platform for a digital Commons
- Activating public engagement and access

We are driving a cultural change around the elimination of disease by activating patients, shifting scientists to share the data and models needed to build better models of disease. To do this, we are building an open Commons called 'Synapse' where data can be shared and a compute space where predictive disease models can be co-evolved so that industry and academia can jointly benefit from understanding biology.

<http://www.sagebase.org>

info@sagebase.org

Response to:

Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research [76 FR 68518]

Jan. 12, 2012

Submitted by:

Library and Information Technology
Bucknell University

Prepared by:

Andrew D. Asher, PhD
Council on Library and Information Resources (CLIR) Scholarly Communications Fellow
Bucknell University
Lewisburg, PA
andrew.asher@bucknell.edu

Background

As an institution, Bucknell University is a firm supporter of open access publishing and scholarship. In October 2011, the faculty of Bucknell University formally adopted an open access policy requiring faculty members to make all of their peer-reviewed journal articles open access and to place these articles in a repository that provides free public access without use restrictions. One of the principal arguments in support of this institutional open access policy is that research results are a public good, often funded directly or indirectly by public funds, and should therefore be made available to the public. Open access is also a form of social justice, allowing anyone to access research regardless of their own, or their institution's, ability to pay subscription fees. In this way, open access maximizes the value of publically funded research to its funders—the public—by enabling individuals, businesses, and educational institutions to more quickly and effectively utilize information to generate new and innovative ideas, products, and services, thus contributing to the overall development of knowledge, as well as the United States economy.

Comment 1(a)

[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?]

Library and Information Technology at Bucknell University believes that all federal agencies should adopt an open access policy requiring all articles resulting from publically funded research to be made immediately and freely accessible to the public. A policy of immediate public access would provide several economic benefits. By removing delays on access to information, an immediate access policy would enable companies and individuals to more quickly and efficiently build and launch products and services based on publically funded research. Furthermore, by generating new uses and applications for research, as well as enabling

the faster commercialization of this research, an immediate public access policy would support economic growth and job creation throughout the economy.

Comment 1(b)

[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?]

Open access to scholarly publications is a key driver of scientific productivity. By removing barriers to access, these policies promote a diversity of research interests and pathways, enable follow-on and corroborating investigation, and enable interdisciplinary and collaborative approaches to research questions. Open access policies also allow scientists to utilize new bibliographic and informatics tools to more effectively and quickly incorporate new information and data, as well as to use these tools to make additional discoveries. Finally, and importantly, open access policies encourage participation by the public in the scientific process, as well as contributions by “unforeseen participants,” which can often lead to new innovations.

Comment 1(c)

[What are the relative costs and benefits of such policies?]

As the National Institute of Health’s public access policy has shown, open access policies can be extremely cost effective. NIH spends only about 1/100th of 1 percent of its \$30 billion annual budget on its public access policies. In return, NIH receives increased return on its research investment (through the mechanisms discussed above) as well as improved accounting and oversight of the outcomes produced by its research funding. NIH has not only demonstrated an effective open access policy model that could be scaled up to apply to all federally funded research, but has also made infrastructure investments that could be leveraged by a broader policy initiative.

Comment 1(d)

[What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?]

Library and Information Technology at Bucknell University supports a policy of full open access, meaning free immediate access including re-use rights to all federally-funded research publications. These publications should be maintained in a fully digital, online environment, where they can be read, downloaded, searched, crawled, and indexed without restrictions. Restrictions on the use of these materials would place limits on the value that the public can derive from taxpayer-funded research as well as the return on taxpayer investment. Access alone without reuse rights is therefore insufficient to fully realize the value of these publications.

Comment 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?]

Open access policies should be implemented in ways that are fully compatible with current copyright laws. Mechanisms should be put in place that enable the full use of publications resulting from federally funded research (e.g. the re-use and distribution of articles, data mining, computation, indexing, etc.), while still maintaining the intellectual property rights of stakeholders and ensuring that these stakeholders receive credit for their work. Implementing licenses that are enforceable under current copyright law, such as the Creative Commons CC-BY license, would be one way to achieve this goal. In this manner, a federally mandated public access policy would help eliminate the effective enclosure of publically-funded research by (often for-profit) publishers who require the transfer of copyright from authors as a condition of publication.

Comment 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 federal government is uniquely able to mandate and ensure that articles resulting from publically funded research are made permanently accessible and useable. The federal government is therefore the appropriate entity to provide permanent stewardship of these articles, and any public access policy for these works must provide adequate rights to enable their archiving and distribution. A public access policy could include multiple repositories, provided that these repositories also support the same access and use policies and that they allow all interested parties to utilize them and the materials they contain.

Comment 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?]

Public/private partnerships should be encouraged, so long as they meet standards and conditions for accessibility, use rights, preservation and interoperability. Given their experience with designing, implementing, and maintaining digital archives, libraries and universities should be specifically encouraged to partner with federal agencies. Open source rather than proprietary software should be utilized for final archive sites for publically funded articles, and under no condition should any single site be the only point of access of these materials or a subset of these materials.

Comment 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?]

First and foremost, the metadata maintained with scholarly publications should enable the use, reuse, and analysis of these works. To this end, the metadata should be machine-readable and machine-interoperable. Existing metadata standards should be utilized, and metadata should provide context for the published articles, such as attribution of funding agencies, grant IDs, and the relationship between entities and articles. The metadata should also be flexible enough to support different specifications for publishing standards that support the analysis of texts as data objects as well as providing a bridge between publications and underlying data.

Comment 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?]

In order to minimize the compliance burden of a public access policy for federally funded research, care should be taken to create policies that ensure a consistency of requirements across all funding agencies which will both reduce the administrative complexity and cost of the policy and increase the rate of compliance by researchers. Public access policies can also maximize returns on taxpayer investment by using existing protocols (e.g. SWORD) to automate the deposit of articles in multiple repositories, integrating article management with grant management systems, and encouraging the development of enhanced productivity tools for authors, researchers, and universities.

Comment 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?]

Library and Information Technology at Bucknell University believes that all peer-reviewed publications resulting from federally funded research should be made readily accessible as soon as possible. However, the policies by which materials such as books, book chapters, and conference proceedings are made accessible may need to differ from those directed at journal articles due to the different conditions these materials are subject to upon publication. Policies should be developed that are cognizant of these differences.

Comment 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?]

Library and Information Technology at Bucknell University supports immediate open access as the ideal time frame for publicly funded articles to be made freely available. However, in deference to journal publishers that rely on subscription income, we find an embargo period of no more than 12 months to be an acceptable compromise.

Professor Victoria Stodden
Department of Statistics
Columbia University
New York, NY

<http://stodden.net>

(1a) 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?

Increased funding could be made available for grants for text-oriented scientific research. Even things like research metrics could be developed. We don't have open ways to evaluate citation and the importance of published paper, and we could also do this for scientific data and code reuse. But the key point is to make the scientific literature available to the public for reuse, by making the publications openly available in a centralized location on the web.

Just permitting the open access to the scientific literature cannot but increase the use and potentially commercialization applications that can be generated from the newly available knowledge. The National Institutes for Health, for example, is making its articles published since 2009, and one can only imagine the number of potentially commercialization ideas and discoveries in the engineering literature just to take one example.

(1b) 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?

It is clear that making more information about scientific discoveries available cannot reduce economic growth. But we can be sure of more than that. Embedded in the literature are many scientific discoveries that could be commercialized and developed by industry, that are either currently unavailable or available via patent licensing fees paid through the institution's Technology Transfer Office. With the onset of the Internet there is no need for such a convoluted way of making scientific knowledge available as the public good it is.

I believe the productivity of the scientific enterprise will increase, not only because of the increased ability to carry out metrics using a corpus of scientific publications, but because all the results will be open to wider scrutiny. With the concurrent movement toward reproducible computational science – making the data and code that created the published result conveniently available – there is a very power change in the ability of the broader public to understand the scientific discoveries being made.

(1c) What are the relative costs and benefits of such policies?

There are no costs, except publisher fees for some subscription journals, which could impact publisher profits. The publication model we have today grew from a pre-digital age when it was costly to print paper journals. Today, the writing, typesetting, and review – nearly all that is necessary for publication – is done by scientists on a volunteer basis. Furthermore, the support for such a publishing system is also archaic, heavily subsidized and influenced by federal funds.

We don't have a free market in academic publishing. Because of the integrated nature of federal funding for the research that is published, as well as the time for review, and for page fees and other publishing costs comes typically from federal funds. The argument is similar for scientific societies that publish subscription journals.

It is not clear what benefit they add for the millions and billions in subscription costs, outside of sorting discoveries by journal prestige.

Because the journal publishing is in such a federally controlled market it is appropriate to rectify market failures due to the advance of technology. Unlike in a free market where firms would have an incentive to adapt and reduce costs, the publishers have none. Because of the difficulty for scientists to publish outside the most prestigious journals, federal action is required to move to a more optimal place, from society's perspective: open access to scientific knowledge.

(1d) What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?

Without constraint, publicly availability of scientific publications, from their publication date. I believe the publishing industry must make the argument that sequestering scientific knowledge increases economic growth or scientific productivity – as a scientist I do not see how any argument exists other than protecting the publishing industry's bottom line, and that cannot be an interest of public policy, but rather creating the environment for open scientific knowledge.

(2a) 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?

We must remember scientific discoveries, a public good, are at the core of this discussion. How they are best transmitted to the public and built upon is with as much exposure and access as possible. To this end the federal agencies should enact a PubMed Central similar to the NIH's for published documents. Make these documents available under an attribution only license, for example CC-BY. Enforce deposit by making grant money contingent and have the deposit occur within 6 months of publication.

Establish federal guidelines recommending scientists keep copyright rather than sign it over the publishing house, who did almost zero in the production of the document and only hold the prestige of the journal as an inducement to publish.

(2b) 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?

Intellectual property rights should not be established with publishing houses. This is the root of the open access problems today. If publishing houses and academic societies could be trusted to act in the best interests of society and not in their own best interest, we would have open access already. Federal policies must rectify this accident left over from the pre-digital age.

(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?

The federal funding agencies have the most power to break the collection action lock preventing open access and should also be the central archivist for the published articles. Any decentralized archives are productive in tandem in that they provide backup and reduce the load on the main repository (provided the same versions of the papers are deposited in each). A central repository facilitates research on the corpus of text, encourages interdisciplinary research and breaks barriers to cross disciplinary communication (you don't have to know in advance which repository will house the information you are interested in) and will last. Links to papers will persist over time. This is the main reason why it is important for our scientific knowledge to be housed by a federal service, like the Library of Congress does for other material. Having private or even institutional repositories as the sole guardian of our scientific culture subjects it to potential loss when they go out of business or decide that such an effort is no longer aligned with their interests.

For example, researchdatasets.google.com was dissolved with about a month's notice after Google had established the site to warehouse scientific data there. Preserving our scientific culture is a federal government task.

(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?

I don't believe the publishers are interested in budging one inch toward open access and I believe the federal government will have to take steps to ensure these corpora are made open. It will be easier to do this for paper not yet published, but as the economic value (through access fees) declines over time publishers may be able to give up older papers, say older than a year, with no appreciable loss to profits.

(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?

I believe the most important first step is making the published scientific articles available in a repository. After that it will be easier to see what is missing in search and how to best rectify these problems. I would not try to establish standards before open access is established. It will create unnecessary barriers to establishing open access and the inevitable fact is that technologies will change and so will our search, tagging, location, classification, and retrieval abilities. The only thing that must be established is a DOI must be placed on each deposited paper for version control and citation reasons.

(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?

There is almost no burden to submitted a published article to an open access repository for scientists. There will be an expense in establishing such federal repositories which is unavoidable, but these textual corpora are small compared to, say, scientific data.

Publishers must also act in the public interest, since this is scientific publication and should not be considered a money making venture. Making profits by controlling scientific communication is bordering on unethical and certainly repugnant.

(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?

I believe the more of our scientific culture that is made publicly available the better, with exceptions for national security or confidentiality interests (which should be taken on a case by case basis).

(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?

Zero days is the appropriate embargo period. Scientific knowledge is a public good. This is why there are massive federal subsidies for scientific research in the first place.

I believe scientists will not suffer without an embargo period, since it is exceedingly difficult to scoop an authors next publication on his or her data. Evidence shows that the typical result is that research from a different field use the data in ways the original author did not anticipate and the original author gains citations with no loss of publications.

(9) Please identify any other items the Task Force might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.

With open access to journal articles, the task force may do well to coordinate with the data policies task force. The reason is because shared data must be versioned and linked to its published articles, which presumably is in the open access federal repository.

Thank you for creating this RFI and listening to our input.

Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research

Response from Anthony D. So, MD, MPA (anthony.so@duke.edu) and Quentin Ruiz-Esparza (qr2@duke.edu), Program on Global Health and Technology Access, Sanford School of Public Policy, Duke University, Durham, North Carolina

We welcome the opportunity to address several issues raised by the questions laid out by the Task Force on Public Access to Scholarly Publications. Specifically, we would emphasize that expanding public access to federally funded, peer-reviewed scholarly articles would help respond to the well-considered recommendations of the Institute of Medicine's report on *The U.S. Commitment to Global Health: Recommendations for the Public and Private Sectors*. In particular, Recommendation 3-3 is noteworthy:¹

The U.S. research community should promote global knowledge networks and the open exchange of information and tools that enable local problem solvers to conduct research to improve the health of their own populations.

- (A) Funders of global health research should require that all work supported by them will appear in public digital libraries, preferably at the time of publication and without constraints of copyright (through open access publishing), but no later than six months after publication in traditional subscription-based journals. Universities and other research institutions should foster compliance with such policies from funding agencies and supplement those policies with institution-based repositories of publications and databases.
- (B) The U.S. government, universities, and other research institutions should develop new methods—such as simplified web-based procedures for executing agreements like materials transfer and nondisclosure agreements—to expedite the sharing of information and research materials with researchers in low- and middle-income countries.
- (C) Scientists, clinicians, advocates, and other personnel involved in defined areas of global health should develop trustworthy websites that aggregate published literature, incorporate unpublished databases or clinical trial information, promote digital collaboration, and disseminate news and other information about common interests.
- (D) Universities and other research institutions that receive federal and philanthropic funding to conduct research should adopt patent policies and licensing practices that enable and encourage the development of technologies to create products for which traditional market forces are

¹ Institute of Medicine. *The U.S. Commitment to Global Health: Recommendations for the Public and Private Sectors*. Washington, DC: National Academies Press, 2009.

not sufficient, such as medicines, diagnostics, and therapeutics that primarily affect populations in low- and middle-income countries.

The U.S. National Institutes of Health is the leading global funder of neglected disease research. Nearly 40% of neglected disease funding in 2010 came from the NIH.² Of course, the value of NIH research for global health extends well beyond just the work funded on neglected diseases. The Report of the UN Secretary General prepared for the High-Level Meeting on Non-communicable Diseases this past September highlighted: “Death and disease from non-communicable diseases now outstrip communicable diseases in every region except Africa, where the rate of such diseases is quickly rising. By 2030, non-communicable diseases are projected to cause nearly five times as many deaths as communicable diseases worldwide, including in low- and middle-income countries.”³ So we would underscore the importance of publicly funded research for both U.S. and non-U.S. research institutions working on global health issues.

In keeping with the Institute of Medicine report recommendation, there would be no 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. As the IOM report recommends, “funders of global health research should require that all work supported by them will appear in public digital libraries, preferably at the time of publication and without constraints of copyright (through open access publishing), but no later than six months after publication in traditional subscription-based journals.” There would be no economic justification to have an embargo period on such publicly funded research in journals not specializing in coverage of neglected diseases, where the publication of occasional articles on these topics could be made available, without embargo, to the public without any risk to the subscriber base of the journal.

An embargo period of no more than six months would be consistent with requirements set by the European Research Council, the Wellcome Trust, and the Howard Hughes Medical Institute.^{4,5,6} In complying with Division G, Title II, Section

² PolicyCures. *Global Funding of Innovation for Neglected Diseases (G-FINDER)*. Sydney and London: PolicyCures, 2011. Available at:

<http://www.policycures.org/downloads/g-finder%20summary%202011.pdf>

³ *Prevention and control of non-communicable diseases: Report of the Secretary-General*. Sixty-sixth session, United Nations General Assembly, A/66/83, 19 May 2011. Available at:

http://www.un.org/ga/search/view_doc.asp?symbol=A/66/83&Lang=E

⁴ European Research Council. 2007. *ERC Scientific Guidelines for Open Access*.

Available at:

http://erc.europa.eu/sites/default/files/document/file/erc_scc_guidelines_open_access.pdf

218 of PL 110-161 (Consolidated Appropriations Act, 2008), the NIH’s Public Access Policy currently allows journal articles “to be made available no later than 12 months after the official date of publication.” Efforts to reduce further the delay to access to U.S. publicly funded research would be most welcomed.

The U.S. Copyright Act of 1976 prevents government employees from claiming copyright (or assigning it to journals) to publications they author, whether scholarly, peer-reviewed research or not. Yet public access—even to such journal articles written by government employees—might be improved through centralized approaches to managing public access. For example, we conducted a preliminary analysis of publications in PubMed.gov by three government agency heads—Dr. Francis Collins, Director of the U.S. National Institutes of Health, Dr. Margaret Hamburg, Director of the U.S. Food and Drug Administration, and Dr. Carolyn Clancy, Administrator of the Agency for Health Research and Quality—in the years 2010, 2011 and so far in 2012. Of the citations posted on PubMed.gov, we found that overall, full-text availability of journal publications by these three government agency heads only was accessible 42% of the time through the one-click away icon of “Free PMC Article” or “Free Article”.

We are not suggesting that these outstanding public servants bear responsibility for ensuring that their publications are one-click away on PubMed Central, but that PubMed Central be provided the resources it needs to do this with greater regularity as a centralized approach to managing both peer-reviewed scholarly publications that are publicly funded and also as a source for full-text publications authored by government employees.

	Year of Publication	No. of abstracts on PubMed	No. of abstracts on PubMed with article readily available*	% of abstracts with article readily available in PubMed
Francis Collins, NIH	2012	2	1	50%
	2011	17	6	35%
	2010	23	17	74%
	Totals:	42	24	57%
Margaret Hamburg, FDA	2012	0	0	N/A
	2011	2	1**	50%
	2010	4	4	100%

⁵ Wellcome Trust. 2007. Conditions under which a Grant is Awarded. Available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sf_central_grants_admin/documents/web_document/wtx026668.pdf

⁶ Howard Hughes Medical Foundation. 2007. Public Access to Publications. Available at: <http://www.hhmi.org/about/research/sc320.pdf>

	Totals:	6	5	83%
Carolyn Clancy, AHRQ	2012	2	0**	0%
	2011	13	1	8%
	2010	11	1	9%
	Totals:	26	2	8%
Overall totals:		74	31	42%

*By “readily available,” we refer to the PubMed.gov feature of flagging some journal articles with one-click away access, either as “Free PMC Article” or “Free Article”. Some of the articles are available on-line for free, but several clicks away. Others are not obviously accessible to non-subscribers to the journal.

**These articles are reportedly “in process” in PubMed.

The White House RFI also calls for “analyses that weigh public and private benefits and account for external market factors, such as competition, price changes, library budgets, and other factors.” In so doing, we would suggest that such analyses be viewed through the lens of several key policy considerations: 1) the context of how much public funding has gone into the research that led to the publication compared to the value added editing done by the journal (noting, of course, that peer review is usually done at no cost to the journal, apart from organizing such review); 2) the potential costs of delayed publication, including the scenario whereby life-saving treatment options might not become known to patients or health care providers in a timely way when publicly funded research might otherwise have made such options known; 3) the value added that might result from creating collections of publicly funded research--absent the transaction costs of seeking copyright permission from multiple journals--for republishing or providing links to public access versions, particularly for those in resource-limited health care settings (e.g., a specialized collection on the diagnosis and treatment of a neglected tropical disease); and 4) the alternative policy option that public funding, now supporting journal subscription costs, could be directed to supporting open access institutional repositories and open access journals.

We have argued elsewhere that:

This calculus of ‘pay now or pay more later’ might guide where the public ought to direct its investments to maximize the returns to the healthcare system. For example, in the value chain of scientific journal publication, paying the publication fees for open-access journals is one way of supporting a business model that encourages the sharing of knowledge. Going further, the U.S. government could develop a system of supporting open-access journals that publish peer-reviewed, publicly funded research. For those open-access journals that charge publication fees, it could build support into

the direct or indirect cost structure of grants. For those open-access journals that do not charge fees, it could provide direct or indirect subsidies. Either way, it could support journals that provide open access rather than impose subscription fees on patients, providers, and universities. This support could factor in transition costs, the citation impact factor of the journal in that field, the rejection rate, and the number of publicly funded research articles published by the journal.⁷

Finally, we would flag concerns raised over access to building blocks to knowledge more generally. Just as the private sector focuses on copyright, patents and trademarks as an incentive for investment, the public sector also should consider the strategic use of intellectual property rights in ensuring an enabling environment for innovation. These concerns have been reflected in the adoption of the Bermuda Rules, whereby leading funders of the Human Genome Project required research centers to deposit the sequencing of every 1000 base pairs on-line into the GenBank within 24 hours of completion. This purposefully prevented the patenting of our human genetic endowment through defensive publishing of prior art.⁸ Along similar lines, the NIH issued “Principles and Guidelines for Sharing of Biomedical Research Resources” in December 1999. This guidance counseled against exclusive licensing or even patenting if the government-funded research yielded “a broad, enabling invention that will be useful to many scientists, or multiple companies in developing multiple products, rather than a project or product-specific resource.”⁹ And most recently, the *New England Journal of Medicine* piece on “Copyright and Open Access at the Bedside” reminds us that protecting building blocks of knowledge for broad public use must extend to copyrighted tools, like the Mini-Mental State Examination.¹⁰ That a newer cognitive screening tool—the Sweet 16—could be removed from being available on an open access basis from the Internet because of a copyright dispute makes this case especially worrisome. This incident serves as a useful warning of the need for the U.S. government to take strong and strategic action to ensure fair returns from publicly funded investments and an enabling environment for innovation.

⁷ So AD, Stewart E. “Sharing Knowledge for Global Health,” in *The US Commitment to Global Health: Recommendations for the Public and Private Sectors*, Institute of Medicine Committee on the U.S. Commitment to Global Health. 2009, page 271.

⁸ Marshall E. “Bermuda Rules: community spirit, with teeth.” *Science* 2001; 291: 1192.

⁹ National Institutes of Health, U.S. Department of Health and Human Resources. “Principles and Guidelines for Recipients of NIH Research Grants and Contracts on Obtaining and Disseminating Biomedical Research Resources: Final Notice,” *Federal Register* 1999; 64(246): 72090-72096.

¹⁰ Newman JC, Feldman R. “Copyright and Open Access at the Bedside,” *New England Journal of Medicine* 2011; 365: 2447-2449.

[Assigned ID #]
[Assigned Entry date]
Nick Shockey; nick@arl.org
The Right to Research Coalition
Washington, DC

I am pleased to offer the following comments on behalf of the Right to Research Coalition. Founded by students in the summer of 2009, the Right to Research Coalition is an international alliance of undergraduate and graduate student organizations, representing nearly 7 million students, that promotes Open Access to scholarship. The Right to Research Coalition believes no student should be denied access to the published articles they need because they or their institution cannot afford access. The coalition works to educate the next generation of scholars and researchers about Open Access and to advocate for policies at the campus, national, and international levels that expand access to the results of research.

A full list of the Right to Research Coalition's members is available at the end of this document.

[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?

[Comment 1]

All peer-reviewed articles resulting from federally funded research should be open-access. That is, these articles should be made freely available immediately upon publication with full reuse rights, so users can text mine, data mine, compute on, and create derivative works – including further research – from the articles without commercial restriction.

Open access to federally funded research would greatly improve the resources available to students, at both the undergraduate and graduate levels, to achieve a complete, up-to-date education. Students' educations depend on access to the research literature. These articles are quite literally the building blocks of an education in any discipline; from its core, all the way out to the cutting edge.

Unfortunately, because of the often-high price of journal subscriptions – 15 academic disciplines have an average price per title in excess of \$1,000 per year¹ – students and the professors who teach them are often left without access to the research they need for a complete, up-to-date education. Not only do students routinely run into access barriers when researching for a paper, for a class, or simply to get a better understanding of a given issue, but students’ professors also run into those same barriers and cannot bring the most cutting-edge research into the classroom. Speaking to this point, Dr. Gary Ward, former PubMed Central Advisory Committee Chair and a researcher at the University of Vermont, has said, “In my role as educator, I often find myself teaching my graduate and medical students what I have access to rather than what they most need to know. Just as one example, in a recent lecture I was preparing for our medical students on how drugs can get across the barrier between the blood and the brain to treat neurological disease, I was only able to access about two thirds of the articles that I needed in order to make sure that I was providing these budding young doctors with everything they needed to know about the subject. I can tell you that’s extremely frustrating to me as an educator and it’s clearly not in the best interests of my students.”²

Because the federal government funds a large portion of all published research, a strong open access policy at the federal level would vastly expand the resources available for students to get a complete, up-to-date education. This improvement in education would translate into immediate and persistent economic benefits to the United States economy. As our economy continues to shift toward innovative, research-based sectors like biotechnology and alternative energies, the companies driving our economy will rely on a highly educated and trained workforce. By providing students with immediate access to cutting-edge research, federal agencies can help ensure students are ready to hit the ground running after graduation and put their education to use immediately, rather than having to play catch up.

Similarly, students’ library cards expire at graduation; however, with the current weak economy, it is increasingly common for students to take months or even years to find jobs in their chosen fields. With no institution to pay for journal subscriptions on their behalf, recent graduates lose all access to the subscription-based academic literature and have a limited ability to stay current in their discipline. A strong federal open access policy would open a wealth of cutting-edge research, enabling graduates to maintain an up-to-date understanding of their field and contribute more quickly once hired.

¹ Bosch, et al., *Periodicals Price Survey 2011: Under Pressure, Times Are Changing*. Library Journal. April 2011. Available at http://www.libraryjournal.com/lj/newslettersnewsletterbucketljxpress/890009-441/periodicals_price_survey_2011_under.html.csp.

² Dr. Ward’s full quote can be found at http://www.taxpayeraccess.org/issues/frpaa/frpaa_resources/press-conference-congressman-doyle-to-address-new-.shtml.

Difficulties in accessing the research literature disproportionately impact students at smaller and less wealthy institutions – especially community colleges – which cannot afford the multi-million dollar library budgets required to access large portions of the scientific and scholarly record.³ As our 21st century economy increasingly requires highly skilled workers, community colleges will become more and more essential in providing American businesses with the advanced workforce required for economic competitiveness. With strong open access policies, federal agencies could provide these institutions, which would otherwise have very limited access to cutting-edge research, the ability to incorporate the most up-to-date information into their students' educations. This would help level the playing field between students at less wealthy and wealthier institutions, and have a persistent positive effect on the skill level of the American workforce.

Beyond students, a federal open access policy would pay real dividends to the United States economy and the advancement of scientific research. A useful analogy can be found in the Human Genome Project (HGP), which sequenced the entire human genome and, critically, made the data immediately, openly available for anyone to use without commercial restriction. By any measure, the HGP was an incredible success in providing a return on taxpayer investment, with a \$5.6 billion federal investment yielding \$796 billion in economic output, over \$6 billion in federal, state, and local taxes, and over 3.8 million job-years of employment to date.⁴ Research has shown that the immediate, open availability of HGP data played a significant role in boosting this economic return. One study comparing the use of similar, but closed data from a parallel sequencing project run by the Celera Corporation found “robust evidence that the package of short-term IP used by Celera has been associated with reductions on the order of 30 percent in subsequent gene-level scientific research and product development outcomes.”⁵ There are strong reasons to believe a federal open access policy would lead to a similar increase in return on taxpayer investment in research.

Following the Human Genome Project's example, making articles resulting from federally funded research immediately and openly available would allow them to be utilized and built upon more quickly and by a larger, more diverse group of researchers and corporations. Immediate availability would shorten research cycles by providing researchers with faster access to breakthroughs, and would accelerate the advancement of science, decreasing the amount of time taken for businesses to

³ To get a sense of the size and variation in library journal subscription budgets, see the Association of Research Libraries' Statistics Report from 2008-2009, p. 40-46. Available at <http://www.arl.org/bm~doc/arlstat09.pdf>.

⁴ *Economic Impact of the Human Genome Project*. Battelle Technology Partnership Practice. May 2011. Available at <http://www.battelle.org/publications/humangenomeproject.pdf>.

⁵ Williams, Heidi., *Intellectual property rights and innovation: Evidence from the human genome*. National Bureau of Economic Research Working Paper Series. July 2010. p. 27. Available at http://www.nber.org/~heidiw/papers/5_12_10a_hlw.pdf.

translate theoretical breakthroughs into new products and services. Faster commercialization will, in turn, boost American economic growth and ultimately create new jobs across the economy as innovation can happen more quickly and with less restriction. Similarly, by making the full body of federally funded science openly available to all, federal agencies can greatly expand the number and diversity of those engaged in follow-on research. The expensive nature of journal subscriptions artificially and arbitrarily limits researchers' access to the journals they can afford rather than what they actually need. An open access policy would not only increase readership among an article's intended audience, but it would also lead to an increased likelihood the article would reach unintended readers in adjacent or seemingly unrelated disciplines. This increased diversity promotes additional paths of follow-on research across scientific domains, leading to breakthrough that would not have occurred without an article's availability to unintended readers.⁶

In addition to unintended readers, full open access allows machines as an entire new class of reader to use the literature to its fullest extent. With approximately 1,350,000 papers published annually,⁷ no single person can hope to read even a tiny fraction of all published articles. We will increasingly rely on computational text and data mining to get an overall picture of the state of a discipline and uncover trends, connections, and new research pathways that would otherwise remain hidden. These computational processes can identify relevant articles and enable scientists to work more efficiently, improving scientific productivity. These services also represent a new layer of potential commercialization on top of public databases, like PubMed Central, that is only possible with open licensing and full reuse rights. To be computed on to their fullest extent, articles must be available in a machine-readable format – XML, not proprietary PDFs – and come coupled with the reuse rights necessary to be crawled by computers and for businesses to sell services based on such computation.

One illustration of the value that can only be created from an open repository is the winner of the recent Binary Battle contest hosted by the Public Library of Science (PLOS) and Mendeley, a reference manager and social network for researchers. The winning application, OpenSNP, takes genomic data – either yours or other data that you upload – and "find[s] the latest relevant research and let[s] scientists discover

⁶ An analogous case of openness promoting the volume and diversity of follow-on research in the area of research materials can be found in Murray, et al., *Of Mice and Academics: Examining the Effect of Openness on Innovation*. National Bureau of Economic Research Working Papers. October 2008. Available at <http://www.hbs.edu/units/tom/seminars/2007/docs/Of%20Mice%20and%20Academics%20Ster n.pdf>.

⁷ Björk, et al., *Global annual volume of peer reviewed scholarly articles and the share available via different Open Access options*. Proceedings ELPUB2008 Conference on Electronic Publishing. June 2008. Available at <http://oacs.shh.fi/publications/elpub-2008.pdf>.

new genetic associations."⁸ This application is a great example of how text and data mining can uncover new connections in a way that is only possible when research is open.

It is important to point out that the aforementioned economic benefits of a federal open access policy only represent uses we can currently imagine. Opening this vast literature – not only to a larger audience of readers, but also for unrestricted use – will undoubtedly pay dividends in ways currently unimaginable.

The benefits of a federal open access policy would far exceed the costs. According to a study done last year by the Center for Strategic Economic Studies, an expansion of the NIH public access policy to cover all federally funded research with a six-month embargo period would provide a 500% return on investment to the United States government.⁹ Such a policy would also generate benefits eight times greater than costs, resulting in a net present value gain worth approximately \$1.5 billion.¹⁰ The impact could be even greater with a shorter embargo period or immediate open access. Furthermore, the NIH policy has a proven track record of cost-effectiveness over the past three years. The NIH spends approximately \$4 million per year to make the articles covered by its policy, approximately 90,000 annually, available through PubMed Central – a total of roughly 1/100th of 1% of the NIH's \$30 billion per year operating budget.¹¹

[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 of peer-reviewed scholarly publications resulting from federally funded scientific research?

[Comment 2]

Federal agencies should require articles resulting from federally funded research to be made available under a fully open license that allows the public to freely use, remix, revise, and redistribute the research without commercial restriction, such as the Creative Commons Attribution-Only ("CC BY") license.¹² Only by adopting fully open licensing will we maximize our collective investment in research and allow it to be used, reused, and built upon to its fullest possible extent. Such an approach

⁸ *Winners of the first Binary Battle Apps for Science Contest*. Mendeley Blog, November 2011. Available at <http://www.mendeley.com/blog/design-research-tools/winners-of-the-first-binary-battle-apps-for-science-contest>.

⁹ Houghton, et al., *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*. July 2010. p. 7-8. Available at <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>.

¹⁰ Ibid.

¹¹ Letter from Dr. Francis Collins, Director of the NIH, to Representative Joseph Pitts. December 2011. Available at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

¹² Creative Commons Attribution 3.0 Unported (CC BY 3.0) License summary available at <http://creativecommons.org/licenses/by/3.0>.

would adequately protect authors' interests by requiring citation – the primary mechanism by which researchers build reputation within their field – while allowing the widest possible distribution and use.

The government can implement a policy requiring open licensing fully within the current system of copyright. Using the same mechanism employed by the NIH policy, in which researchers consent at the time of grant acceptance to make their work freely accessible in PubMed Central, agencies can require articles resulting from their funding to be made available under an open license, such as CC BY. Open licenses, such as those offered by Creative Commons, operate within the current system of copyright and have been upheld as legally enforceable by the US Court of Appeals for the Federal Circuit.¹³ Furthermore, the CC BY license is already in use by a federal grant program, namely the Department of Labor's \$2 billion Trade Adjustment Assistance Community College and Career Training grant program (TAACCCT).¹⁴

While the NIH policy has been successful by all accounts, federal public access policies should now go beyond read-only access and include full reuse rights without commercial restriction. When taxpayers fund research, they deserve the full use of the results – to distribute, reuse, data or text mine, and build business on top of – rather than solely the permission to read resulting articles. As mentioned above, open licensing is crucial to maximizing the potential scientific and commercial benefit that can be realized from federally funded research. Opening this vast literature – not only to a larger audience of readers, but also for unrestricted use – will encourage the creation of innovative new tools, such as the OpenSNP application mentioned in comment 1, and pay dividends in ways we cannot presently imagine in the current closed system.

While an immediate open license maximizes the return on taxpayer investment in research, one compromise that could be considered to balance the interest of all stakeholders would be a stepped approach. Initially, articles would be under a period of embargoed access in which usage is restricted to only those uses allowed under copyright and fair use. Then, after the expiration of an embargo period of perhaps three to six months, the articles would be subject to an open license that would allow full reuse rights without commercial restriction, such as CC BY. This approach would allow publishers a sufficient period to recoup their investment, and would still give the public the full reuse rights they deserve for underwriting the research. Much of the additional economic benefit only gained when articles are made openly available would also be captured under this approach.

¹³ Case law supporting the legal enforceability of Creative Commons licenses can be found at http://wiki.creativecommons.org/Case_Law.

¹⁴ See TAACCCT's Notice of Solicitation for Grant Applications, p. 21: <http://www.doleta.gov/grants/pdf/SGA-DFA-PY-10-03.pdf>.

[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?

[Comment 3]

Federal agencies are the appropriate entity to maintain a centralized repository to ensure permanent, public access to publicly funded research. At a minimum, agencies should maintain a mirrored, publicly accessible copy of all articles resulting from federal funding and ensure they retain the rights necessary to do so, as the NIH has done through its public access policy since 2008.

Centralized repositories like PubMed Central (PMC) provide students, researchers, and others with a single point of access to a vast portion of the relevant research literature. This single interface provides students superior ease of use compared to collections of articles scattered across the websites of thousands of individual journals. This ease of use, in turn, enhances discoverability and scientific productivity. NIH's PubMed Central has convincingly demonstrated the excellent return on investment of such a repository. PubMed Central sees 500,000 unique users every day,¹⁵ three-quarters of whom are from outside of the academy.¹⁶

Federal custody is necessary to protect our investment in research by ensuring long-term stewardship over the course of decades or longer. Publishers' incentives and limitations necessitate such an approach. As with any business, publishers can and will fail, and without a properly maintained backup, large numbers of federally funded articles could be erased permanently when a publisher goes out of business. Publishers may also wish to prevent competitors from building products and services on top of their content by stipulating that any centralized repository be a "dark archive." However, public access to such a centralized repository is crucial to maintain archival veracity and maximize the return on our federal research investment.

Furthermore, establishing centralized repositories for other agencies (or groups of agencies) can be accomplished with relatively minor expense or effort. PubMed

¹⁵ Letter from Dr. Francis Collins, Director of the NIH, to Representative Joseph Pitts. December 2011. Available at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

¹⁶ Statement by David J. Lipman, MD, Director, National Center for Biotechnology Information, Public Access to Federally-Funded Research before the Committee on Oversight and Governmental Reform Subcommittee on Information Policy, Census and National Archives, United States House of Representatives. July 2010. Available at <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>.

Central's existing platform can be customized to meet the needs of other agencies at a fraction of the cost of starting from scratch. Alternatively, NIH's PubMed Central could be expanded to house all federally funded research in one central, cross-agency repository.

[Question 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?

[Comment 6]

For the reasons mentioned in previous comments, a policy requiring immediate, open access to articles through a centralized, PMC-like repository would maximize the benefit to the public and create the highest return on our federal investment in research.

To minimize the burden on all stakeholders, agencies should standardize the language, requirements, and procedures of their policies, being as consistent as possible. As institutions and researchers are often awarded grants by multiple federal agencies, such consistency will be essential to reduce complexity for grantees and increase policy compliance. Researchers should only need to learn one process, not be forced to navigate a web of different, conflicting requirements across federal agencies.

[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?

[Comment 7]

Any peer-reviewed publications resulting from federally funded research and created without the expectation of compensation should be made freely accessible to the public. Free access to these publications would provide significant value to students, researchers and others. For example, conference proceeding papers can provide additional or unique information not present in final publications, include preliminary results that allow insight into future publications, or contain comprehensive reviews of published research to date that can keep others informed of the current state of a given field. However, policies for making these other types of peer-review publications available may differ from those that apply to journal articles; thus, they should be considered separately.

[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?

[Comment 8]

The public should be granted open access to the results of federally funded research immediately upon publication.

American students, in particular, would benefit significantly from immediate, rather than embargoed, access. It is unacceptable to ask students to make do with outdated information. Instead, federal agencies should get cutting-edge research into students' hands immediately. Immediate access to these resources is especially crucial in rapidly evolving fields, such as biotechnology and alternative energy, that form a significant portion of the United States economy and represent some of its most innovative and high-growth sectors. By providing students with improved access to a cutting-edge education, public access policies that provide immediate access can boost American economic competitiveness by helping students hit the ground running after graduation and put their education to use immediately. Furthermore, courses only last three to four months. With an embargo period, a course may be taught many times before the newest research becomes available and thus can be integrated into the class.

If an embargo period is deemed necessary, it should be as short as possible, and the full opportunity cost of slowing the pace of research and delaying students' access to the most up-to-date research should be taken into account when considering the embargo's length. An embargo period should not exceed twelve months and would preferably be six months or less, as is the norm among research funders around the world with such policies.¹⁷ Similarly, hundreds of subscription-based journals voluntarily make their content freely available after embargo periods, typically of six to twelve months.¹⁸ This list includes publishers that have previously expressed concern over the potential negative impact of opening up access to their content. One such example is the Royal Society, the world's oldest scientific publisher, which earlier this year announced it would make its entire historical journal archive available online for free. Finally, the NIH's public access policy provides strong empirical proof that such measures do not harm subscription-based publishers. To date, no publisher has presented any evidence that the NIH policy has harmed its business. In fact, the largest commercial publisher, Elsevier, which owns a large

¹⁷ A complete list of funder access policies, including details and embargo periods, can be found at <http://roarmap.eprints.org/view/type/funder=5Fmandate.html>.

¹⁸ A complete list of subscription journals which allow embargoed access to their content can be found at <http://highwire.stanford.edu/lists/freart.dtl>.

number of journals affected by the NIH policy, has seen its profit margin and revenues increase every year since 2008 when the NIH policy took effect.¹⁹

Embargo periods have a cumulative impact, as they delay new research by their duration at each research cycle. For example, a paper under a twelve-month embargo will not be available to a large portion of researchers until a year after it is published, delaying follow-on research. If papers from that follow-on research are also subject to a twelve-month embargo, then the availability of those results is delayed a full two years. This delay will continue to accumulate with each cycle of research until it far exceeds the original embargo period.

¹⁹ Elsevier's most recent annual financial reports can be found at:
2010: http://reports.reedelsevier.com/documents/pdfs/reed_ar_2010.pdf; relevant figures: p. 134
2009: http://reports.reedelsevier.com/PDFFiles/ReedElsevier_AR09.pdf; relevant figures: p. 91

The Right to Research Coalition includes 48 member student organizations:

American:

- The American Medical Student Association
- The American University Washington College of Law Student Bar Association
- California Institute of Technology Graduate Student Council
- Columbia University Graduate Student Advisory Council
- Cornell University Graduate and Professional Student Assembly
- Dartmouth College Graduate Student Council
- Harvard Extension Pre-Health Society
- Library and Information Science Student Association, Simmons College
- Massachusetts Institute of Technology Graduate Student Council
- Massachusetts Institute of Technology Undergraduate Association
- National Association of Graduate-Professional Students
- Oberlin College Student Senate
- Oklahoma State University Graduate and Professional Student Government Association
- St. Olaf College Student Government Association
- Student Advocates for Graduate Education
- The Student Public Interest Research Groups
- Students for Free Culture
- Trinity University Association of Student Representatives
- Tufts Graduate Student Council
- Tufts University Friedman School of Nutrition Science and Policy Student Council
- The United States Student Association
- Universities Allied for Essential Medicines
- University of California, San Diego Graduate Student Association
- University of Minnesota Graduate and Professional Student Assembly
- University of Nebraska - Lincoln Graduate Student Association
- University of Tennessee - Knoxville Student Government Association

International:

- The Association of Medical Students in Bulgaria
- Athabasca University Graduate Students' Association
- The Canadian Federation of Students
- The Croatian Pharmacy and Medical Biochemistry Students' Association
- Direção Executiva Nacional dos Estudantes de Medicina (Brazil)
- The European Federation of Psychology Students' Associations
- The European Medical Students' Association
- The European Medical Students' Association - Turkey
- The European Pharmaceutical Students' Association
- The Indian Medical Student Association
- The International Association for Political Science Students
- The International Association of Students in Agricultural and Related Sciences
- The International Federation of Medical Students' Associations
- The International Federation of Medical Students' Associations - The Netherlands
- The Lebanese Medical Students' International Committee
- The Macedonian Medical Student's Association
- The Malta Medical Students' Association
- The Medical Students' Association of Kenya
- Medsin-UK
- National Graduate Caucus of the Canadian Federation of Students
- Udruga Studenata Dentalne Medicine (Croatia)
- University of Calgary Students' Academic Assembly

From: Meredith Jacob

To: publicaccess@ostp.gov

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

Michael Carroll; mcarroll@wcl.american.edu

Professor of Law and Executive Director; Program on Information Justice & Intellectual Property, American University, Washington College of Law

Meredith Jacob; mjacob@wcl.american.edu

Assistant Director; Program on Information Justice & Intellectual Property, American University, Washington College of Law

Washington, DC

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

(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 publicly 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?

Yes. All federal agencies that provide research support should require that the authors of any peer-reviewed publications that result from federally funded scientific research should deposit a copy of their post-peer-reviewed manuscript with a designated federal repository and that the author ensure that he or she has the legal authority to grant a copyright license to the federal agency to make the publications freely available over the Internet. Moreover, federal policy should require that the author(s) of such publications retain sufficient rights to grant the public a license to liberally reuse such publications; such license grant taking place no later than 24 months after the date of publication.

It is important to recognize how poor a fit copyright law is for scholarly publishing. Copyright is an author's right. The law grants these rights on the assumption that authors desire legal control over reuse in order to trade this control for remuneration. However, the authors of research articles reporting the results of scientific research do not depend upon the copyright incentive to produce these articles. As a consequence, public access to all scientific

research should not await the expiration of the automatic copyright term of life of the author plus 70 years before this research is made freely available to the public. Where public funds have been used to conduct the research and produce the articles reporting and interpreting the results, the argument for public access in the short term is even more compelling.

All agencies should require open access to peer-reviewed publications that result from federally funded research. At this point, federally operated central repositories are the best tool to serve this end. Federally operated repositories serve the interests of preservation and access that extend beyond the interests of commercial publishers. In addition to preservation, centralization and standardization in federally operated repositories provide the opportunity for novel uses, including the development of machine-aided research. Finally, the use of central, federally operated repositories provides the opportunity for historical data collection on document use, providing both a stronger understanding of the benefit of federal research funding, as well as data for analysis of research methodology.

The use of standardized, open formats, such as XML, improves usability of data, expanding both the speed at which we can make discoveries, but also increasing the ways that information can be accessed.

This standardized access enables machine-aided research and machine aided decision-making, increasing the productive capacities of the United States' innovation-based industries. Effective public access to the research literature also fuels innovation in the building of next generation machine aided discovery and analysis tools central to technology industry growth.

Toward this end, federal policies should address access and terms of reuse. Guidelines for both access and reuse should be as open and free as possible to widen the scope of potential innovation. There is huge value in building enough capacity and flexibility to enable tomorrow's unanticipated technological uses, not just those available today.

Science and technology education in high schools and community colleges also benefits from strong open access policies by providing the tools for students and teachers to engage in current science. The creation of strong science and technology students is crucial for continued job creation.

(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 is an author's right, vesting in the author at the time of creation and fixation in a tangible medium. Publishers do not employ investigators who conduct research nor do they finance scientific research nor do they compensate those who provide peer review services as an indirectly federally-subsidized professional service. With all due respect, the question is potentially misleading because publishers do not have intellectual property rights in scholarly publications that could be "undermined." Instead, these rights are authors' rights, and publishers only get intellectual property rights in scholarly publication through a transfer of rights under copyright by the federally-supported author of the copyrighted work.

Any policy requiring that the author(s) of federally-subsidized copyrighted work agree to grant the agency a copyright license ensuring that it may provide access and reuse rights to the public does not undermine any intellectual property rights of the publisher. Rather, this agreement between funding agency and author precedes any interest that the publisher may have in the copyrighted work. Publishers remain free to refuse to accept for publication federally funded articles with these conditions. If they choose to publish such articles, as biomedical publishers have done in response to the NIH Public Access Policy, they do so knowingly and with full consent to the terms and conditions to which the author has agreed in exchange for the federal support that enabled the article to be produced in the first place. As a result, federal policies that require that federally-funded researchers grant federal agencies a copyright license to scholarly articles emerging from such support are fully consistent with the basic policy that copyright is an author's right and can in no way "undermine" any rights owned by a publisher because such rights would have been acquired with full knowledge of the terms and conditions to which the author(s) agreed with the funding agency.

By requiring that the author manage copyright in a way that ensures legal public access and reuse rights through a federal repository, federal agencies are merely joining the large group of those who fund the creation of works of authorship and routinely demand terms and conditions on the distribution of those works of authorship. Where they are created by public funds, those terms and conditions should include deposit in an open access archive, including a grant of rights of reuse by the author or copyright holder a reasonable time after publication, such as 12-24 months.

With respect to the mechanics of how federal agencies receive their copyright license from federally-subsidized authors, it is very important for the White House to support NIH's view that the longstanding copyright license granted to the federal agencies under the terms of OMB Circular A-110 and similar regulations provides the legal basis for providing online access to scholarly articles. NIH chose to require the grant of a second copyright license at the time of submission, but as a legal matter this was entirely unnecessary. *See attached* Michael Carroll, Complying with the NIH Public Access Policy - Copyright Considerations and Options, available online at: www.arl.org/sparc/bm~doc/NIH_Copyright_v1.pdf

(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?

There are at least three distinct advantages to centralized federally operated repositories.

Search capabilities are increased in a central, standardized database. Incomplete interoperability of commercial databases with library search functions and incomplete holdings are a barrier to scientific innovation. Having a centralized search capability both increases the value of search results, and also enables the creation and continued development of powerful search tools. Centralized search also aids the linkage of scientific discoveries across disciplines when unanticipated results are found from other fields.

Standardization increases the ability to use the information for machine aided research and machine aided decision-making. This not only speeds the rate of discovery in established scientific disciplines but also provides a corpus of high-value information for the development of machine-aided discovery and analysis tools in developing disciplines and for non-research use.

Finally, a centralized repository can act as a powerful data-gathering tool for data about the research products being used. This data can improve Federal allocation of research funding, improve the mechanics and understanding of the usability of scholarly articles, and provide insight into the research methodologies of different user groups.

(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?

Publishers should adopt open standards for their archives. Interoperability with government archives should be a feature that users demand from publisher repositories for the most powerful search results and the highest value for subscription dollars spent. Central federal repositories can serve as educators and standard-setters for search interoperability.

(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?

The benefit to a central, federally operated repository would include the centralized authority for establishing metadata guidelines and ensuring consistent applications across scholarly fields.

(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?

Benefits are maximized and costs are minimized if Federal agencies have strong open access policies that are consistently, and vigorously, enforced. A blanket policy for all executive agencies would foster a single well-understood and streamlined process for all peer-reviewed articles that are funded in whole or in part by Federal agencies. If the deposit in a central Federal repository is a standard part of publication, it will become an efficient, even automated, part of the publication process with little administrative burden or cost. The broader and more consistent the open access scheme, the greater the benefit accrued to the public and the least administrative cost per article submitted. Instead, it is the current patchwork of policies that increases cost and administrative burden, while limiting the public benefit.

(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?

Yes, because copyright is an author's right, not one associated primarily or exclusively with the publisher. Any research result for which the author was paid in whole or in part through federal funds should be made publicly accessible. This includes, for example, book chapters on federally funded research.

(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?

For there to be an "embargo", there must first be a deposit of the embargoed article. Consequently, federal policy should require that federally-funded researchers deposit copies of their articles, as modified in response to peer review, immediately upon completion. If public access should be limited, this embargo period should last no later than 6 months after the date of publication. Additionally, rights for reuse should be provided on a standardized timeline, no later than 12-24 months from the date of publication.

COMPLYING WITH THE NATIONAL INSTITUTES OF HEALTH PUBLIC ACCESS POLICY:

Copyright considerations and options

A JOINT SPARC / SCIENCE COMMONS / ARL WHITE PAPER
By Michael W. Carroll

February 2008

SPARC[®]

 **science
commons**



ASSOCIATION OF RESEARCH LIBRARIES

Legal Disclaimer: This document is not legal advice, and this document does not create an attorney-client relationship. The author, SPARC, Science Commons, and ARL provide this information on an "as-is" basis. The author, SPARC, Science Commons, and ARL make no warranties regarding the information provided, and each disclaims liability for damages resulting from its use.

COMPLYING WITH THE NATIONAL INSTITUTES OF HEALTH ("NIH") PUBLIC ACCESS POLICY:
COPYRIGHT CONSIDERATIONS AND OPTIONS - A JOINT SPARC/SCIENCE COMMONS/ARL WHITE
PAPER

By Michael W. Carroll

February 2008

Copyright © 2008 by Michael W. Carroll. This White Paper is licensed to the public under a Creative Commons Attribution Non Commercial 3.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0>). Attribution under the license should be to Michael W. Carroll as author and to SPARC, Science Commons, and the Association of Research Libraries as co-sponsors.

Science Commons
c/o MIT CSAIL
Building 32-386D, Cambridge, MA 02139
<http://www.sciencecommons.org>

SPARC
21 Dupont Circle, Suite 800, Washington, DC 20036
<http://www.arl.org/sparc/>

Association of Research Libraries
21 Dupont Circle, Suite 800, Washington, DC 20036
<http://www.arl.org/sparc/>

Version 1.0, February 2008.

Available for free download at <http://www.arl.org/sparc>.

COMPLYING WITH THE NATIONAL INSTITUTES OF HEALTH
PUBLIC ACCESS POLICY: COPYRIGHT CONSIDERATIONS AND OPTIONS

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY.....	i
II. BACKGROUND	2
A. Brief History of the NIH Public Access Policy.....	2
B. A Note on Copyright Law	3
C. Copyright and the Public Access Policy	4
III. ANALYSIS OF GRANTEES’ COPYRIGHT-RELATED OBLIGATIONS UNDER THE PUBLIC ACCESS POLICY.....	8
A. The Copyright Compliance Risk	8
B. Timing.....	9
IV. COMPLIANCE OPTIONS FOR GRANTEES AND PRINCIPAL INVESTIGATORS	10
Option 1. Rely on individual authors to satisfy grantee’s obligation under the award.....	10
Option 2. Assist Authors with Copyright Management.	11
Option 3. The Grantee License.....	11
Option 4. The Grantee License - Plus	12
Option 5. Negotiate Directly with Publishers.....	13
Option 6. Pre-clear Journals.....	14
V. ADDITIONAL SUGGESTION AND CONCLUSION.....	15
About the Author	16
About SPARC	16
About Science Commons.....	16
About the Association of Research Libraries	16
APPENDIX A - SUGGESTED COVER LETTER FOR AUTHOR JOURNAL SUBMISSION	17
APPENDIX B - ANALYSIS OF SELECTED BIOMEDICAL PUBLISHERS’ COPYRIGHT POLICIES.....	18

I. EXECUTIVE SUMMARY

On January 11, 2008, the National Institutes of Health (“NIH”) adopted a revised Public Access Policy¹ for peer-reviewed journal articles reporting research supported in whole or in part by NIH funds. Under the revised policy, the grantee shall ensure that a copy of the author’s final manuscript, including any revisions made during the peer review process, be electronically submitted to the National Library of Medicine’s PubMed Central (“PMC”) archive and that the person submitting the manuscript will designate a time not later than 12 months after publication at which NIH may make the full text of the manuscript publicly accessible in PMC.

NIH adopted this policy to implement a new statutory requirement under which:

The Director of the National Institutes of Health shall require that all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication to be made publicly available no later than 12 months after the official date of publication: Provided, That the NIH shall implement the public access policy in a manner consistent with copyright law.²

This White Paper is written primarily for policymaking staff in universities and other institutional recipients of NIH support responsible for ensuring compliance with the Public Access Policy. The January 11, 2008, Public Access Policy imposes two new compliance mandates. First, the grantee must ensure proper manuscript submission. The version of the article to be submitted is the final version over which the author has control, which must include all revisions made after peer review. The statutory command directs that the manuscript be submitted to PMC “upon acceptance for publication.” That is, the author’s final manuscript should be submitted to PMC at the same time that it is sent to the publisher for final formatting and copy editing.

Proper submission is a two-stage process. The electronic manuscript must first be submitted through a process that requires input of additional information concerning the article, the author(s), and the nature of NIH support for the research reported. NIH then formats the manuscript into a uniform, XML-based format used for PMC versions of articles. In the second stage of the submission process, NIH sends a notice to the Principal Investigator requesting that the PMC-formatted version be reviewed and approved. Only after such approval has grantee’s manuscript submission obligation been satisfied.

Second, the grantee also has a distinct obligation to grant NIH copyright permission to make the manuscript publicly accessible through PMC not later than 12 months after the date of publication. This obligation is connected to manuscript submission because the author, or the person submitting the manuscript on the author’s behalf, must have the necessary rights under copyright at the time of submission to give NIH the copyright permission it requires. This White Paper explains and analyzes only the scope of the grantee’s³ copyright-related obligations under the revised Public Access Policy and suggests six options for compliance with that aspect of the grantee’s obligation.

¹ Revised Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research, NOT-OD-08-033, <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-033.html>.

² Consolidated Appropriations Act, 2008, Pub. L. No. 110-161 Div. G, Tit. II, § 218.

³ For the sake of convenience, recipients of NIH support are referred to as “grantees” hereafter regardless of the contractual form of NIH support.

Time is of the essence for NIH grantees. As a practical matter, the grantee should have a compliance process in place no later than April 7, 2008. More specifically, the new Public Access Policy applies to any article accepted for publication on or after April 7, 2008 if the article arose under (1) an NIH Grant or Cooperative Agreement active in Fiscal Year 2008, (2) direct funding from an NIH Contract signed after April 7, 2008, (3) direct funding from the NIH Intramural Program, or (4) from an NIH employee.

In addition, effective May 25, 2008, anyone submitting an application, proposal or progress report to the NIH must include the PMC reference number when citing articles arising from their NIH funded research. (This includes applications submitted to the NIH for the May 25, 2008 and subsequent due dates.)

Conceptually, the compliance challenge that the Public Access Policy poses for grantees is easily described. The grantee must depend to some extent upon the author(s) to take the necessary actions to ensure that the grantee is in compliance with the Public Access Policy because the electronic manuscripts and the copyrights in those manuscripts are initially under the control of the author(s). As a result, any compliance option will require an explicit understanding between the author(s) and the grantee about how the manuscript and the copyright in the manuscript are managed. It is useful to conceptually keep separate the grantee's manuscript submission obligation from its copyright permission obligation because the compliance personnel concerned with manuscript management may differ from those responsible for overseeing the author's copyright management.

With respect to copyright management, the grantee has the following six options:

- (1) rely on authors to manage copyright but also to request or to require that these authors take responsibility for amending publication agreements that call for transfer of too many rights to enable the author to grant NIH permission to make the manuscript publicly accessible ("the Public Access License");
- (2) take a more active role in assisting authors in negotiating the scope of any copyright transfer to a publisher by (a) providing advice to authors concerning their negotiations or (b) by acting as the author's agent in such negotiations;
- (3) enter into a side agreement with NIH-funded authors that grants a non-exclusive copyright license to the grantee sufficient to grant NIH the Public Access License;
- (4) enter into a side agreement with NIH-funded authors that grants a non-exclusive copyright license to the grantee sufficient to grant NIH the Public Access License and also grants a license to the grantee to make certain uses of the article, including posting a copy in the grantee's publicly accessible digital archive or repository and authorizing the article to be used in connection with teaching by university faculty;
- (5) negotiate a more systematic and comprehensive agreement with the biomedical publishers to ensure either that the publisher has a binding obligation to submit the manuscript and to grant NIH permission to make the manuscript publicly accessible or that the author retains sufficient rights to do so; or
- (6) instruct NIH-funded authors to submit manuscripts only to journals with binding deposit agreements with NIH or to journals whose copyright agreements permit authors to retain sufficient rights to authorize NIH to make manuscripts publicly accessible.

II. BACKGROUND

As is well known, PubMed Central is central to a suite of interconnected databases that is perhaps the most valuable research archive in biomedicine.⁴ Administered by the National Institutes of Health (NIH) through the National Library of Medicine (NLM), PMC is a free, Internet-accessible archive of full text articles from peer-reviewed scholarly biomedical journals. The Public Access Policy is designed to increase the value of this resource to the biomedical research community and to the general public.

A. Brief History of the NIH Public Access Policy

On July 14, 2004, the Appropriations Committee of the U.S. House of Representatives instructed NIH to develop a policy requiring free online access to articles arising from NIH-sponsored research no later than six months after the articles' publication in peer-reviewed journals.

NIH responded in September 2004 with a notice of a draft policy⁵ followed by a public comment period during which thousands of comments were received and reviewed by NIH.⁶ NIH released the final version of the Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research on February 3, 2005 with an effective date of May 2, 2005.⁷ The policy provided, in pertinent part:

Beginning May 2, 2005, NIH-funded investigators are requested to submit an electronic version of the author's final manuscript upon acceptance for publication, resulting from research supported, in whole or in part, with direct costs from NIH. The author's final manuscript is defined as the final version accepted for journal publication, and includes all modifications from the publishing peer review process.

...

At the time of submission, the author will specify the timing of the posting of his or her final manuscript for public accessibility through PMC. Posting for public accessibility through PMC is requested and strongly encouraged as soon as possible (and within twelve months of the publisher's official date of final publication).

The three key features of the final policy were: (1) it was voluntary instead of mandatory; (2) NIH decided as a matter of policy that it would require the person submitting a manuscript to grant NIH copyright permission to make the full text article publicly accessible; and (3) the duration of the optional embargo period for public access was extended from six to 12 months after publication.

⁴ See PMC Frequently Asked Questions, <http://www.pubmedcentral.nih.gov/about/faq.html>.

⁵ Enhanced Public Access to NIH Research Information, NOT-OD-04-064, Sept. 3, 2004, <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-04-064.html>.

⁶ NIH first released its draft policy on its web site on September 3, 2004, commencing a 60-day comment period. NIH then published the same text in the *Federal Register* on September 17, 2004, 69 Fed. Reg. 56074, also commencing a 60-day comment period. The comment periods were merged and the comment period closed on November 16, 2004.

⁷ Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research, NOT-OD-05-022, Feb. 3, 2005 at <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html>.

In early February 2006, the NIH sent a progress report to Congress (dated January 2006). Among other things, NIH reported that the rate of compliance with its request for public access was below 4%. Responding to the data demonstrating a low compliance rate under the voluntary policy, Congress enacted as part of the Consolidated Appropriations Act, 2008 a provision requiring NIH to make its voluntary policy mandatory.⁸ NIH released its revised policy on January 11, 2008 with an effective date of April 7, 2008.⁹

B. A Note on Copyright Law

The peer-reviewed journal articles subject to NIH's Public Access Policy are copyrighted works of authorship, and NIH has been directed by statute to make these articles publicly available in a manner consistent with copyright law. The remainder of this Section addresses the copyright issues that NIH must address with respect to its policy. The Public Access Policy requires grantees and investigators to take an active role in ensuring that public access to NIH funded research is consistent with copyright law. The following section identifies the copyright issues that grantees and Principal Investigators must address to comply with NIH's implementation of this statutory requirement.

1. Authorship and Transfers of Copyright Ownership

The author is automatically the initial owner of the copyright in an original work of authorship as soon as the work has been fixed in a tangible medium of expression. Originality requires independent creation by the author and a modicum of creativity. Facts and ideas are not copyrightable. Consequently, the results and underlying data reported in an article are facts that are not subject to copyright. Similarly, the insight or idea leading to an experiment is also not subject to copyright. In the case of journal articles, the copyright applies to the author's creative expression, such as the choice of text to describe materials and methods, an experiment or its result. Tables, figures, charts or other accompanying material are copyrightable only if some minimally creative decisions were required in their design.

Once the copyright vests in the author, s/he can authorize others to use the work in one of four ways: (1) assign the entire copyright; (2) grant an exclusive license; (3) grant a non-exclusive license; or (4) dedicate the copyright to the public domain. An author must sign a written document to effectively assign the copyright or grant an exclusive license. In contrast, a non-exclusive license or permission can be granted quite casually. A verbal okay or even conduct, such as posting a work on a publicly accessible web server, is deemed to be the grant of a non-exclusive license. The remainder of this White Paper uses the terms "permission" and "non-exclusive license" interchangeably.

In some quarters, confusion has arisen about whether the copyright in the first draft of an article (a.k.a. "pre-print") is distinct from the copyright in the final published version. It is important to note that the scope of the exclusive rights encompasses the exact text or any text that is "substantially similar." Although in some cases there may be a distinct copyright in the authors' revisions to the article, the substantial similarity standard usually means that the owner of the copyright in an article has the exclusive rights to control the dissemination of any version of the article. Thus, a transfer of copyright to a publisher does not leave the author with the rights to grant a license with respect to the author's final manuscript or any other earlier drafts of the article. It is, however, possible for the copyright owner to use licensing to allocate different rights with respect to different versions of an article. Some journal publishing agreements that transfer the copyright to the publisher do just this.

⁸ Consolidated Appropriations Act, 2008, Pub. L No. 110-161, Div. G, Tit. II, § 218.

⁹ Revised Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research, NOT-OD-08-033, <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-033.html>.

Since the author is initially in a position to manage the copyright in a biomedical journal article, it is important to identify the author or authors for purposes of complying with the Public Access Policy. Although it may seem obvious that those authors listed on a journal article jointly own the copyright in the article, this may not be so. First, there is a fundamental legal uncertainty about who the copyright law recognizes as the author of scholarly articles written by university faculty. Familiarity with this uncertainty is assumed, and this White Paper proceeds on the assumption that either by law or by university policy, faculty authors have the right to transfer exclusive rights under copyright to journal publishers.¹⁰ There is also a factual uncertainty that arises with every co-written article that affects the proper legal characterization of the authors' relationship and their respective rights to grant licenses or to transfer rights under copyright with respect to the article. This issue is discussed in Section III below.

2. Liability for Copyright Infringement

Distributing copies of full text journal articles over the Internet implicates copyright. The copyright owner has the exclusive rights to (1) reproduce the work in copies, (2) publicly distribute copies, (3) publicly perform the work, (4) publicly display the work, and (5) prepare derivative works. The courts have been imprecise in their analysis of how these exclusive rights apply on the Internet, but under the current interpretation of copyright law when a user downloads a copy of a work from an Internet server and views the work on the screen, the copyright owner's rights of reproduction, public distribution and public display have been exercised.

The liability analysis has two steps. First, one asks whether there is a valid copyright in the work and whether the use includes the exercise of one or more of the copyright owner's exclusive rights. Second, if the answers to those questions are yes, one asks whether such use is permitted under the fair use doctrine, under other limitations or exceptions to copyright or under a license or permission from the copyright owner.

In the case of public access to full text articles on PMC, NIH would be exercising the rights of reproduction or distribution and public display by sending copies of copyrighted manuscripts to members of the public who request them. Systematically distributing copies of another's copyrighted work over the Internet generally will not qualify as a fair use or under any of the other statutory limitations and exceptions to copyright. In this author's view, NIH would infringe copyright by systematically distributing copies of peer-reviewed journal articles from its PubMed Central archive without permission to do so.

C. Copyright and the Public Access Policy

Under the Public Access Policy, NIH receives two copyright licenses in connection with the journal articles written with NIH support. One is a license that NIH receives at the time the grant, cooperative agreement or contract comes into force. The second is permission granted during the process of submitting the manuscript to PMC. Under the terms of the Public Access Policy, NIH is relying only on the permission granted during manuscript submission as the basis for providing public access to full text articles. For this reason, it is essential that grantees and Principal Investigators ensure that copyright in NIH-funded articles is managed so that the author's final manuscript is submitted by one with authority to grant NIH permission to make it

¹⁰ University and other institutional copyright policies reflect this uncertainty about authorship under the law, with some asserting ownership by the university and a license back to the faculty and others purporting to recognize the faculty's customary or traditional rights under copyright. For the record, this author is of the view that most university policies that purport to treat faculty as the legal authors of their own work risk being ineffective if it is determined that scholarly articles are works made for hire. A transfer of exclusive rights requires a written instrument signed by the author. See 17 U.S.C. § 204. It is uncertain which, if any, general university policies satisfy this writing requirement.

publicly accessible within 12 months of publication. Although the first license is not directly relevant to compliance with the Public Access Policy, the mechanics of how that license is granted suggest a means for compliance with the Public Access Policy.

1. Federal Purpose License

Under applicable Health and Human Services regulations, those funded by the government may keep the copyright in works created with support from federal funds. However, NIH, as a part of its funding agreements, “reserves a royalty-free, nonexclusive and irrevocable right to reproduce, publish, or otherwise use the work for Federal purposes, and to authorize others to do so.”¹¹ This license is granted prior to the creation of the copyrighted work, and it comes into effect as soon as the work is created. Thus, subject to issues discussed in Section III below, any subsequent transfer of copyright by the author or grantee is subject to NIH’s Federal Purpose license.

One might ask whether this license suffices as the legal basis for the Public Access Policy. On its face, NIH’s license “to reproduce, publish or otherwise use” copyrighted manuscripts written with federal support would appear to cover public access to these works through NIH’s PMC server. During the comment period on the February 3, 2005 version of the policy, the American Physiological Society and the American Association of Immunologists filed a “legal analysis” that looked very much like a legal brief arguing that NIH had misinterpreted the scope of its own license and would infringe copyright if this license were relied upon as the basis for posting final manuscripts in PMC.¹² In its response to the comments accompanying the February 3, 2005 version of the policy, NIH explained:

Although the NIH, at this time, is not relying on the government purpose license, it is an available means for NIH to reproduce, publish or otherwise use copyrighted works resulting from NIH funding for Federal purposes, as well as to authorize others to do so. Arguments put forth and cases cited by the commenter as support for the premise that the government purpose license could not be used as a basis for PMC to post the manuscripts are not persuasive. None of the cases address circumstances where a government agency is acting to fulfill its own statutory purposes with regard to publications resulting from its own research funding. Creation of a publicly accessible, permanent archive of NIH-funded research publications is squarely within the statutory authorities of the NIH and the NLM and clearly constitutes a Federal purpose.¹³

In this author’s opinion, NIH is clearly correct about the scope of the Federal Purpose license, but it is also the case that based on the content, style and tenor of the APS/AAI analysis, NIH faced a non-trivial risk that it would have to litigate the issue had it chosen to rely on this license. Consequently, NIH chose as part of the February 3, 2005 version of the policy to require the person submitting the manuscript to set the embargo period and to specifically grant NIH permission to make the manuscript publicly accessible after that period.

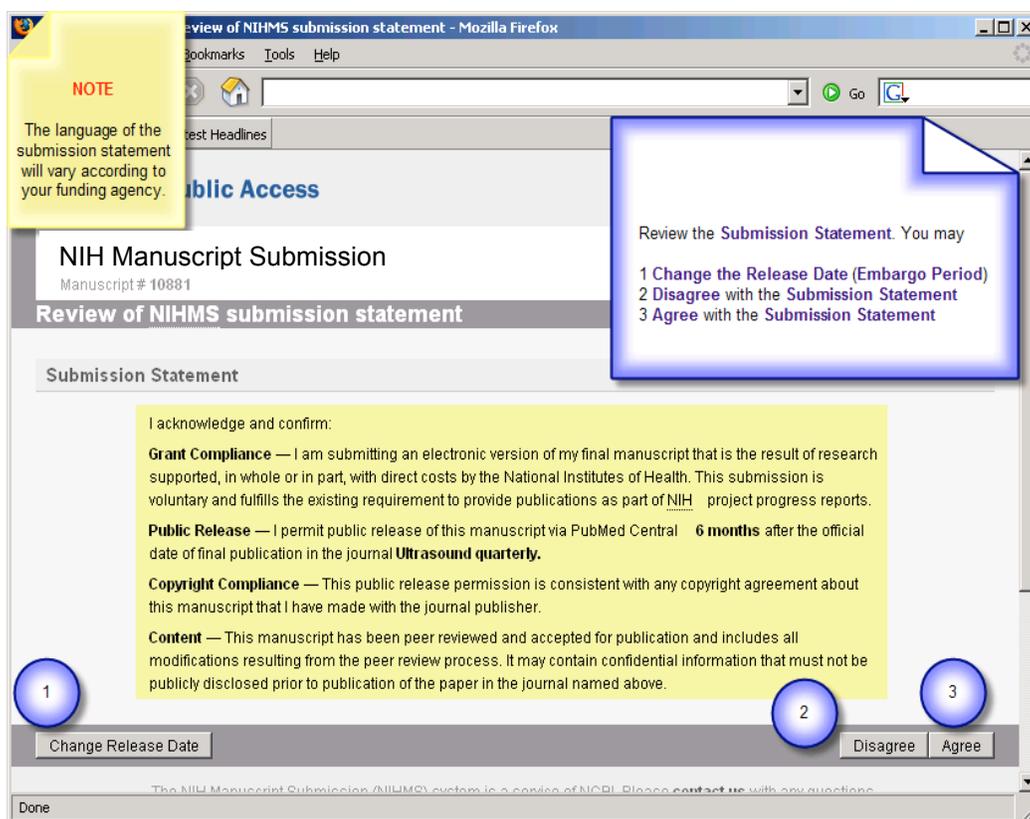
¹¹ 45 C.F.R. § 74.36 (2007).

¹² See Comment of the American Physiological Society (Nov. 16, 2004) (Attachment A), <http://www.dcprinciples.org/responses/aps.pdf>.

¹³ Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research, NOT-OD-05-022, Supplementary Information, § II.P.3 (Feb. 3, 2005), <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html>.

2. The Public Access License

Under the January 11, 2008 revision of the Public Access Policy, NIH has not changed the preexisting steps in the manuscript submission process. The only material change is that use of this process is now mandatory. The compliance issue discussed in this White Paper for grantees and Principal Investigators arises out of the required step in the submission process in which the person submitting the manuscript sets the time after publication (not to exceed 12 months) at which the manuscript should be made publicly accessible, and the submitter grants NIH copyright permission to make the manuscript publicly accessible. A screen shot of this step from NIH's submission guide is set forth below.¹⁴



Under copyright law grant of permission is the same as the grant of a non-exclusive license. For purposes of clarity, this White Paper refers to the “Public Release” permission granted to NIH in the above-referenced step in the manuscript submission process as the “Public Access License.” This is not a term that NIH uses in its explanation of the Public Access Policy, but it is useful to have a name for this permission or license because it is the legal basis for NIH’s compliance with the statutory command that the Public Access Policy be implemented “consistent with copyright law.”

In its response to comments accompanying the release of the February 3, 2005 policy, NIH directly explained why it had chosen to rely on the Public Access License as its policy choice for complying with copyright law:

Copyright: NIH received comments that the proposal infringes on copyright interests of Federal grantees. These commenters argued that copyright interests

¹⁴ The Guide can be found at <http://www.nihms.nih.gov/web-help/PI-NPMC/PI-NPMC-22.html>.

are well-established under Federal law, that NIH has no authority to alter them, and that the proposal is not consistent with controlling Department of Health and Human Services (HHS) regulations. They believe the proposal fails to recognize the need for copyright permission from authors and/or publishers. They argue that neither the principle of fair use, nor the Federal purpose license, can be used by NIH to implement the proposal. Finally, they argue that the PMC “open access” submission agreement constitutes a forced license and undermines copyright.

The Policy explicitly recognizes and upholds the principles of copyright. First, submission of final manuscripts is voluntary rather than mandatory; the voluntary submission to NIH by authors and institutions under the Policy constitutes permission to post the manuscripts on PMC and release to the public after the submitter's specified post-publication delay time. The fair use exemption to copyright infringement does not apply to the government's request for the manuscripts. It applies to the public use of the manuscripts as posted on PMC and provides a limitation on such use consistent with the terms of that exemption.

NIH does not need to seek permission from journals who may acquire copyrights from authors or institutions because any copyright transfer or assignment is currently subject to the government purpose license pursuant to 45 C.F.R. 74.36. Although the NIH is relying on permission, rather than the government purpose license, as the basis for its Policy, the government purpose license is fully available as a legal authority under which manuscripts could be reproduced, published, or otherwise used for Federal purposes. The comment that the proposal is not consistent with controlling HHS regulations granting copyright is not persuasive, since those same regulations grant the agency its government purpose license.

Finally, authors can indicate what copyright restrictions, if any, apply to their manuscripts when submitting them to PMC and can choose an appropriate PMC submission agreement that recognizes those rights.¹⁵

Although NIH stressed the voluntary nature of manuscript submission in its explanation, the change from voluntary to mandatory submission under the January 11, 2008 policy has no effect on NIH's compliance with copyright law. So long as the person submitting the manuscript has the authority to grant NIH the Public Access License, NIH's subsequent distribution of copies of manuscripts to the public will comply with copyright law.

¹⁵ Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research, NOT-OD-05-022, Supplementary Information, § II.P.2 (Feb. 3, 2005), <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html>. This author has one quibble with NIH's understanding of the copyright consequences of the reader's downloading a manuscript. NIH asserted that the reader must rely solely on fair use for the copy made during download. However, since the purpose of the Public Access License is to enable members of the public to display a copy of the manuscript on their computer screens and to download a copy to the reader's hard drive, the Public Access License granted during manuscript submission also covers the display copy or download copy made by the reader. Any further copies made by the reader would be subject to the normal restrictions imposed by copyright subject to the normal limitations and exceptions, including fair use.

III. ANALYSIS OF GRANTEES' COPYRIGHT-RELATED OBLIGATIONS UNDER THE PUBLIC ACCESS POLICY

Congress has mandated that NIH make peer-reviewed, copyrighted manuscripts publicly accessible through PubMed Central provided NIH does so consistent with copyright law. Copyright law requires that NIH have a license to publicly distribute copies of these manuscripts from one with the rights to grant it. NIH relies on the person submitting the manuscript to have the rights under copyright to grant NIH the Public Access License. Under the Public Access Policy, the *grantee* must ensure that the author's final manuscript is submitted to PMC by one with authority to grant NIH the Public Access License.

A. The Copyright Compliance Risk

Unless grantees intervene with their authors prior to April 2008, the risk of non-compliance with the Public Access Policy arising from these authors' business-as-usual copyright management practices is substantial. Since copyright in articles written with NIH support starts out in the hands of the author, the grantee must monitor how the author manages copyright in the manuscript or the grantee must take more direct action in order to ensure compliance. In particular, the focus of risk management should be on how investigators handle copyright agreements sent to them by journal publishers.

Copyright is an author's right. At the time the article is written, the author is in a position to grant NIH the Public Access License required by the January 11, 2008 revised policy. The risk arises with respect to publication agreements because under copyright law, for an author to transfer exclusive rights under copyright, he or she must do so in writing. Once an NIH-funded *author* signs an agreement that transfers all or most rights under copyright to the publisher, the *grantee* will be non-compliant with the terms and conditions of the grant award unless a publisher submits the manuscript. Even if the author or a member of grantee's staff subsequently submits a manuscript to PMC, the person submitting the manuscript will not have the legal authority to grant NIH the Public Access License and therefore such a submission will still be non-compliant.

No formal study of publication agreements was conducted in preparation of this White Paper, but this author has reviewed the copyright policies of selected large or prominent biomedical publishers as detailed in Appendix A. That review suggests that for a number of journals the author(s) must assign copyright or grant an exclusive license of sufficient breadth that neither the author nor the grantee retains sufficient rights to grant NIH the Public Access License after signing the publication agreement. In February 2008, only a small percentage of publishers had a binding agreement with NIH to submit manuscripts and to grant NIH the Public Access License.¹⁶

Thus, if the author(s) of an NIH-sponsored article subject to the Public Access Policy follow standard practice and sign some of the journal publishers' copyright agreements reviewed without altering the terms, the grantee will be non-compliant with the Public Access Policy unless the publisher, as a matter of grace, chooses to submit the manuscript to PMC. Some publishers have adopted a policy to voluntarily submit manuscripts to PMC. The largest of these is Elsevier.¹⁷ Grantees that rely on such policies for compliance run the risk that this policy may change or that such publishers will fail to execute and not deposit a manuscript.

¹⁶ See NIH, Journals that Submit Articles to PubMed Central, http://publicaccess.nih.gov/submit_process_journals.htm.

¹⁷ See Elsevier, Funding Body FAQ, http://www.elsevier.com/wps/find/authorsview.authors/author_sponsorship_information.

Many NIH-sponsored articles are co-written. The legal implications for compliance with the Public Access Policy are as follows. When many authors are listed on an article they may be (1) joint authors; (2) authors only of their expressive contributions (text, graphics, etc.); or (3) not authors at all for copyright purposes. Joint authors are those who contribute original expression to a work with an intent that their contributions will be merged into an inseparable whole. Joint authors share all rights in the copyright in an article as tenants in common. Each joint author may grant non-exclusive licenses, and the joint author may transfer his or her interest in the copyright to third parties. All joint authors must agree for a transfer of all rights under copyright to be effective. Under these principles, any of the co-authors individually has the legal power to ensure compliance with the Public Access Policy by granting NIH the Public Access License.

For the sake of completeness, there is a case-specific risk to flag. With respect to a small portion of co-written articles, the grantee may not have a relationship with the owners of all the rights under copyright and therefore may not be in a position to ensure compliance with the Public Access Policy. It is probably the case that most co-written articles are jointly authored, which means that the grantee's employees have the legal rights to ensure compliance with the Public Access Policy unless they sign away too many rights in a publication agreement.

In some cases, however, it may be that an author contributes only separately created material – such as a copyrightable figure or table – produced in relation to prior research. In these cases, there is some risk that each author has only the rights to grant a license with respect to his or her contribution to the article. It is still the case that those other authors and their institutions also are obliged to comply with the Public Access Policy, and it may also be the case that the other authors agreed to let such an author act as their agent. Consequently, although the risk has been flagged, in this author's view, this risk is insubstantial enough that it would not be worth investing resources to identify precisely what copyrightable expression a grantee's employee contributed to a co-authored article and under what terms.

A related risk is that not all persons listed as co-authors on an article are necessarily authors at all in the eyes of copyright law. Only those persons who contribute original expression to the text of the article or the associated materials are authors for copyright purposes, even though scientific norms require attribution for others who contributed to the underlying research. Where a grantee's employee is listed as a co-author but does not own any share of the copyright in an article, the grantee will have to rely on the employing institutions of those authors with rights under copyright to ensure that the manuscript is submitted to PMC and that NIH receives the Public Access License.

B. Timing

April 7, 2008 is the key date for compliance, but the precise contours of the obligation are defined by the date of funding and the date of acceptance for publication. From the risk management perspective, grantees should have a plan in place to address author(s)' copyright management practices with respect to NIH funded articles accepted for publication on or after April 7, 2008. However, not all of these articles are necessarily subject to the Public Access Policy. Such articles are subject to the policy if the article arose under (1) an NIH Grant or Cooperative Agreement active in Fiscal Year 2008, (2) direct funding from an NIH Contract signed after April 7, 2008, (3) direct funding from the NIH Intramural Program, or (4) from an NIH employee.

In addition, effective May 25, 2008, any person submitting an application, proposal or progress report to the NIH must include the PMC reference number when citing articles arising from their NIH funded research. (This includes applications submitted to the NIH for the May 25, 2008 and subsequent due dates.)

IV. COMPLIANCE OPTIONS FOR GRANTEES AND PRINCIPAL INVESTIGATORS

Grantees effectively have six options for complying with their copyright-related obligations under Public Access Policy.

Option 1. Rely on individual authors to satisfy grantee's obligation under the award.

From a grantee's perspective, relying on authors to ensure compliance with the Public Access Policy is attractive insofar as it avoids adding copyright-related overhead to the office responsible for grants compliance. But this option also poses a foreseeable risk of non-compliance with the attendant consequences for future funding. Many NIH funded authors are likely to submit their manuscripts to publishers whose respective copyright agreements would, if signed without alteration, leave the author with insufficient rights to grant NIH the Public Access License as required in the mandatory manuscript submission process.

As a result, this option would require substantial time and effort by authors in the short term to understand the scope of their obligation to retain sufficient rights by amending the publisher's offered agreement or to find a new publisher in some cases. It is foreseeable that some authors would not successfully amend the publisher's copyright agreement and would still sign such an agreement to get their work published. Even if the author submits the final manuscript to PMC and purports to grant NIH the Public Access License in the process, the author would lack the authority to do so. In such a situation, a publisher would be within its rights to demand that NIH disable public access to the manuscript, thus frustrating the purpose of the Public Access Policy.

To reduce the burden on authors, grantees interested in pursuing this option or related versions may be interested to know that there are a number of standardized author's addenda, including a joint SPARC and Science Commons addendum, each of which would leave the author with more than sufficient rights to grant NIH the Public Access License if accepted by a journal publisher. These addenda can be generated through the Scholar's Copyright Addendum Engine, which is freely available for local hosting.¹⁸

There are two types of risk associated with Option 1. First, many articles are co-authored, and the grantee's faculty member may not be the corresponding author who has been designated to negotiate copyright issues with the publisher by the co-authors. In such a situation, the grantee would be reliant upon a researcher in another institution to reserve sufficient rights under copyright to enable the grantee to comply with the Public Access Policy.¹⁹

Second, even when the corresponding author is employed by the grantee, the level of risk associated with Option 1 depends upon whether the author's interest in ensuring that the grantee remains in compliance is sufficiently strong to motivate the author to negotiate with a journal publisher or, in the worst case, to refuse an offer of publication if the publisher is unwilling to cooperate with the requirements of the Public Access Policy.

¹⁸ Scholar's Copyright Addendum Engine at <http://sciencecommons.org/projects/publishing/scae>.

¹⁹ The legal effect of the corresponding author's signature on a copyright transfer agreement is subject to some uncertainty. Under copyright law, joint authors may grant non-exclusive licenses to third parties without their co-author's permission, and a joint author may transfer his or her share in the copyright without a co-author's approval. However, all co-authors must agree for there to be a complete transfer of exclusive rights. If the corresponding author in fact has agreement from the co-authors to transfer rights to the publisher on behalf of all co-authors, the transfer will be legally effective. If, however, one or more co-authors has not agreed, then the publisher only receives a share of the exclusive rights, and the remaining co-author retains sufficient rights to comply with the Public Access Policy.

Option 2. Assist Authors with Copyright Management.

Instead of relying entirely on the author to satisfy the grantee's obligation, a grantee might engage more directly with the Principal Investigator and authors under his or her supervision to manage the copyrights in articles arising from NIH-sponsored research. Option 2 comes in two varieties and the administrative resources necessary to implement each vary accordingly.

(a) Author Education.

The first variation is to provide authors with dedicated resources to educate them about their rights under copyright and with some standardized forms, such as author's addenda, that could be used to amend a publisher's copyright agreement to ensure that the author has rights sufficient to comply with the Public Access Policy. These resources might also include tips for negotiating with publishers about copyright, lists of publishers or journals that have agreements with NIH, and lists of journals known to be cooperative with the Public Access Policy and those known to refuse to publish articles subject to the Public Access Policy. Depending on the number of faculty and associated researchers who receive NIH funding, this option is likely to require additional staff resources.

This variation provides the grantee with greater assurance than Option 1 that the author understands the nature of the copyright-related obligation under the Public Access Policy. Like Option 1, this variation relies on the author's incentive to keep the grantee in compliance to ensure that the author manages copyright appropriately.

(b) Author representation.

Alternatively, the grantee might ask or require NIH-funded authors to authorize the grantee to act as the author's agent in negotiating copyright issues with journals. Under this variation, the grantee would be responsible for reviewing and signing publication agreements on the author's behalf. The grantee's licensing agent would also be responsible for using a contractual addendum to alter those agreements that would otherwise render the author unable to grant NIH the Public Access License.

Implementing this option would likely require greater expenditure than Option 1 because the grantee would hire one or more licensing agents. The benefit of centralizing this function, however, is that such agents would be more familiar than authors with the range of publication agreements and would have experience with negotiating amendments to these as is necessary. Option 2 could be an interim step toward Option 5, under which the grantee would negotiate a standard copyright agreement with at least the major biomedical publishers to ensure that the author retains sufficient rights to grant the Public Access License.

Option 3. The Grantee License.

Options 1 and 2 focus on ensuring that the author retains sufficient rights to enable the grantee to comply with the Public Access Policy. As an alternative, the grantee may seek to directly acquire sufficient rights from the author to ensure its own compliance with the Public Access Policy.

The surest method for acquiring such a non-exclusive license is to require the Principal Investigator and any other researchers working on an NIH-sponsored project to grant to the grantee a non-exclusive copyright license at the time they commence work on the project. The potential authors of copyrightable journal articles make a legal commitment at the time they accept NIH support for their work. At the moment when the article is first drafted (and when subsequent revisions are made), the actual copyright license is granted automatically under the terms of the commitment to the grantee. The best way to implement Option 3 is to have each

researcher working on the NIH-sponsored project sign a standard form granting the license to the grantee at the time the researcher commences work on the project.

The scope of this license would include the right to grant the Public Access License to NIH, and thus the grantee ensures that it is in a legal position to comply with its contractual commitments to NIH. In concept, Option 3 is exactly like NIH's Federal Purpose license except that the licensee is the NIH grantee rather than NIH. As a legal matter, if the mechanics of the license are executed properly, this license would survive any subsequent action by the author that may seem in conflict with the grant of this license – such as the author's signing a publisher's copyright agreement that purports to transfer all rights under copyright to the publisher.

There is still a risk that, without proper education, the author may sign a form that requires the author to make a representation that is arguably false. But even if that occurs, the grantee is protected because the grantee still has the legal rights necessary to comply with the terms and conditions of its agreement with NIH.²⁰ To close the loop, because the grantee would still have the rights to grant NIH the Public Access License, NIH would be implementing the Public Access Policy consistent with copyright law even as to these articles. The publisher's only legal recourse would be against the author for representing that s/he could transfer all rights under copyright without any prior licenses when, in fact, a prior license had been granted. While it is unlikely, in this author's opinion, that a journal publisher would assert such a claim against an author, it would be advisable to invest institutional resources in making clear to authors the need to amend publication agreements that are inconsistent with the grantee's license.

Finally, even though the article may have authors from other institutions, each co-author (in the copyright sense of the word) is empowered to grant non-exclusive licenses to the copyrighted work. In general, a license from the grantee's own faculty member or employee in a co-authored journal article would be sufficient so long as all authors contributing copyrightable expression to the article and accompanying materials did so with an intention that these be merged into an inseparable whole.

Option 4. The Grantee License - Plus

Option 3 contemplates a strategy aimed only at compliance with the Public Access Policy. However, if the grantee chooses to negotiate the terms of copyright with the Principal Investigator and other researchers, the question arises why the grantee should not also secure permission to post a copy of the author's final manuscript, or perhaps even the final published version of the article, in the grantee's own digital repository.

No matter which option the grantee chooses, the unavoidable fact is that the Public Access Policy requires grantees to arrive at a more explicit understanding about copyright with the NIH-funded authors they employ than heretofore has been the case. Under the policy, the point of that conversation is to ensure public access to the NIH funded research. Since that conversation is now necessary, and since the grantee must assume a greater administrative role in copyright matters as a result, it seems sensible to think that the grantee might use this changed circumstance as a means for furthering its own institutional goals by also providing public access through the grantee's institution itself.

Some evidence suggests that some faculty would be receptive to granting the university a license as well. For years, faculty and librarians on campuses across the country have drawn

²⁰ See 17 U.S.C. § 205(e) (“A nonexclusive license, whether recorded or not, prevails over a conflicting transfer of copyright ownership if the license is evidenced by a written instrument signed by the owner of the rights licensed or such owner's duly authorized agent, and if (1) the license was taken before execution of the transfer; or (2) the license was taken in good faith before recordation of the transfer and without notice of it.”).

attention to the need for author education and better copyright management to improve scholarly communication. As mentioned in the discussion of Option 1 above, a number of standardized author addenda have been produced for use by authors to amend publication agreements so that authors retain the rights to make their work publicly accessible on the Internet, including through PMC.

Recent developments suggest that some faculty are interested in more explicit community commitments to manage copyright in a manner that facilitates public access to their work. For example, as of February 2008, the Faculty in the University of California system had under consideration a proposed policy by which faculty members would commit to routinely granting to the Regents of the University of California a non-exclusive license to place a copy of their scholarly work in a non-commercial, open access repository.²¹ In February 2008, the University of Oregon Faculty Senate passed a resolution calling on authors to retain rights to provide open access.²²

In addition, on February 12, 2008, the Faculty of Arts and Sciences at Harvard University became the first faculty in the United States to adopt a policy under which each author would grant to the university a license sufficient to permit posting of faculty-authored articles in the Harvard repository and to permit Harvard to permit reposting of such articles so long as access is available without a charge for profit. This license is waivable by the faculty member on an article-per-article basis.²³

As with Option 3, by entering into a separate agreement with the NIH-supported researcher-authors, the grantee can ensure its own compliance by taking a license prior to the signing of any publication agreements. Although Option 4 protects the grantee's interest, prudence dictates that the grantee should also educate authors to alert them to the risks of signing a publisher-drafted copyright form that calls for representations that the author cannot legally make. The measure suggested in Part V below also would mitigate this risk for the authors.

Option 5. Negotiate Directly with Publishers.

Options 1 and 2 rely on authors to ensure that the grantee is compliant. Options 3 and 4 ensure that the grantee is in a position to comply with the Public Access Policy but leave to the author the responsibility of alerting the journal of the prior license granted to the NIH grantee. These options pose the risk for unwary authors that they will sign forms that call for representations they cannot legally make. Options 5 and 6 focus on publishers' willingness to act as a partner to facilitate compliance with the Public Access Policy.

Under Option 5, grantees may seek to mitigate the risk of non-compliance by negotiating some more general form of copyright understanding with biomedical publishers. This solution might take the form of a binding agreement between the grantee (or a group of grantees) and the

²¹ University of California, Proposed Open Access Policy, <http://osc.universityofcalifornia.edu/openaccesspolicy/OpenAccess-Policy-DRAFT1-29-2007.pdf>; see also UC Open Access Policy Proposal, http://www.lib.berkeley.edu/scholarlycommunication/uc_open_access_policy.html (Feb. 2007).

²² Motion US 07/08 - 17 Initiative to protect the rights of faculty authors of scholarly publications (passed Feb. 13, 2008) at <http://www.uoregon.edu/~uosenate/dirsens078/US078-17.html>.

²³ Harvard is then in a different compliance position with respect to articles arising from NIH-funded research written by its Arts & Sciences faculty. On the one hand, the university is automatically in a position to comply with the Public Access Policy unless the faculty member has waived the Harvard license. On the other hand, if the faculty member waives the Harvard license, Harvard must then choose one of the options listed here to ensure compliance with the Public Access Policy. Of course, not all NIH-funded authors at Harvard are members of the Faculty of Arts & Sciences, so Harvard is in the same position as all other grantees with respect to these authors.

publisher(s) by which either (1) the author of any NIH-funded articles accepted by any of the publisher's journals retains sufficient rights to grant the Public Access License or (2) the publisher contractually commits to the grantee to deposit final manuscripts (or published versions) in PMC within 12 months of publication.

With respect to initial deposit of the manuscript and copyright licensing, some publishers already have binding agreements with NIH such that publication can be business-as-usual for the author and grantee with respect to articles they publish because NIH treats articles published in these journals as per se compliant.²⁴ There are other publishers who have made only voluntary commitments to post to PMC. Relying on these voluntary commitments is risky because NIH does not treat publication in these journals as per se compliant with Public Access Policy. In addition, it should be noted that Option 5 can fully address the copyright obligation under the Public Access Policy but not the deposit obligation. A publisher can make a binding agreement to deposit the final manuscript into PMC and to grant NIH the Public Access License, but the grantee must still ensure that the author receives the review copy of the manuscript in PMC's XML format and must approve that formatting for the deposit requirement to be met.

Chances are slim that Option 5 is a realistic possibility prior to April 7, 2008 because of the number of publishers who would have to agree to this solution. Consequently, Option 5 should be considered as a longer term solution that might follow short-term adoption of one of Options 1-4.

Option 6. Pre-clear Journals.

Finally, the grantee can manage its compliance risk perhaps most fully by limiting the field of journals to which the author(s) may submit the manuscript. The grantee would require researchers to agree as a condition of working on an NIH-sponsored project that any articles arising from the project would be submitted only to journals that have been pre-cleared by the grantee. These journals are likely to fall into one of four categories. First, the journal could be an open access journal that, for example, uses Creative Commons licenses. Any version of these public licenses automatically gives NIH the rights equivalent to those called for by the Public Access License. Second, journals may be pre-cleared because they have binding deposit agreements with NIH. Third, journals may be pre-cleared because they have binding deposit agreements with the grantee. Fourth, journals may be pre-cleared if their copyright agreements already give the author sufficient rights to comply with the Public Access Policy without need for amendment.

The risk management benefits to the grantee of Option 6 are self-evident, but researchers may resist this level of control over research dissemination. The feasibility of this option depends in part upon the quantity and quality of journals that would meet one of the four above-mentioned criteria for pre-clearance.

²⁴ See Journals That Submit Articles to PubMed Central, http://publicaccess.nih.gov/submit_process_journals.htm.

V. ADDITIONAL SUGGESTION AND CONCLUSION

The compliance options discussed in Section IV comprise the six possible ways in which a grantee could ensure that one with legal authority to do so submitted the author's final manuscript to PMC and granted NIH its Public Access License during the process of manuscript submission.

None of these options is foolproof. Options 1 and 2 carry the risk that the author may fail to retain sufficient rights to grant NIH its license. Options 3 and 4 ensure that the grantee is capable of complying but pose the risk that the author may make a misrepresentation to a publisher about the rights s/he has. Option 5 requires publisher agreement, and Option 6 would likely meet resistance from faculty and would require enforcement.

One additional measure to consider is to require that NIH-funded authors include with any article submission notice to the publisher that the article arises under an NIH-funded project and an agreement that, if accepted, the article will be published in a manner consistent with the Public Access Policy. This measure is not a substitute for the options discussed in Section IV, but it may serve to further mitigate risk in some cases.

Placing the publisher on early notice of NIH support for the research reported in an article provides a legal basis for arguing that the boilerplate terms of the publisher's copyright agreement must be read in light of the additional knowledge the publisher had upon receipt of the article. A variety of legal and equitable theories could be called upon to support the position that by accepting the article for consideration while knowing that the author's final manuscript must be submitted to PMC and be made publicly accessible within 12 months of the date of publication, the publisher cannot be heard to complain later when the article is made publicly accessible, even if the author happened to sign the publisher's form copyright agreement whose terms are arguable inconsistent with the grant of the Public Access License.

In at least some cases, this language alone may be legally ineffective to retain rights for the author. It is therefore recommended that this submission notice be used in conjunction with an addendum to the publisher's copyright agreement when necessary.

Please see Appendix A for suggested language that authors may use in a cover letter accompanying a manuscript submitted for publication in a peer-reviewed journal.

* * * * *

The National Institutes of Health Public Access Policy promises to improve knowledge dissemination in the biomedical sciences by making federally funded research publicly accessible to a range of audiences. While researchers at grantee institutions are among the likely beneficiaries of this policy, grantees must shoulder a new responsibility for ensuring that researchers properly manage their manuscripts and copyrights to comply with the Public Access Policy. Whichever option(s) for ensuring compliance with manuscript deposit and copyright permission to NIH seem most attractive, grantees and NIH-funded researchers will need to share an explicit understanding about the proper management of the approximately 80,000 manuscripts produced annually with NIH support and the 80,000 copyrights in those manuscripts.²⁵

²⁵ See NIH, Public Access Frequently Asked Questions, <http://publicaccess.nih.gov/FAQ.htm#f4>.

About the Author

Michael W. Carroll is a Professor at the Villanova University School of Law, and he serves on the Board of Directors of Creative Commons, Inc. His research and teaching interests are in the areas of intellectual property law and cyberlaw. As an extension of his scholarly and professional interests, Professor Carroll is an active advocate for open access to the scholarly literature. Prior to joining the Villanova faculty, Professor Carroll practiced law at Wilmer, Cutler & Pickering in Washington, D.C., specializing in intellectual property and e-commerce matters. He also served as a law clerk to Judge Judith W. Rogers, U.S. Court of Appeals for the D.C. Circuit and Judge Joyce Hens Green, U.S. District Court for the District of Columbia. Professor Carroll received his A.B., with general honors, from the University of Chicago and his J.D. magna cum laude from the Georgetown University Law Center.

About SPARC

SPARC, the Scholarly Publishing and Academic Resources Coalition, is an international alliance of academic and research libraries working to correct imbalances in the scholarly publishing system. Developed by the Association of Research Libraries, SPARC has become a catalyst for change. Its pragmatic focus is to stimulate the emergence of new scholarly communication models that expand the dissemination of scholarly research and reduce financial pressures on libraries. Action by SPARC in collaboration with stakeholders – including authors, publishers, and libraries – builds on the unprecedented opportunities created by the networked digital environment to advance the conduct of scholarship.

About Science Commons

Science Commons designs strategies and tools for faster, more efficient web-enabled scientific research. Science Commons identifies unnecessary barriers to research, crafts policy guidelines and legal agreements to lower those barriers, and develops technology to make research data and materials easier to find and use. The goal of Science Commons is to speed the translation of data into discovery and to unlock the value of research so more people can benefit from the work scientists are doing.

About the Association of Research Libraries

The Association of Research Libraries (ARL) is a nonprofit organization of 123 research libraries in North America. Its mission is to influence the changing environment of scholarly communication and the public policies that affect research libraries and the diverse communities they serve. ARL pursues this mission by advancing the goals of its member research libraries, providing leadership in public and information policy to the scholarly and higher education communities, fostering the exchange of ideas and expertise, and shaping a future environment that leverages its interests with those of allied organizations.

APPENDIX A

SUGGESTED COVER LETTER FOR AUTHOR JOURNAL SUBMISSION

Dear [Publisher or Editor name],

Enclosed is a manuscript to be considered for publication in _____ [Journal name]. The research reported in this manuscript has been funded through the National Institutes of Health and therefore its publication must comply with the NIH Public Access Policy (<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-033.html>).

In order to ensure compliance with the NIH policy I, as corresponding author on behalf of all the authors, am retaining the rights to:

- Provide a copy of the authors' final manuscript, including all modifications from the publishing and peer review process, to the NLM's PubMed Central (PMC) database at the time the manuscript is accepted for publication; and
- To authorize NIH to make such copy of the manuscript available in digital form for public access in PMC no later than ____ months (indicate 0 to 12 months) after publication.

[Universities may insert any additional terms pertaining to author and institutional rights for reproduction, distribution for academic activities, deposition in institutional archive, etc. Suggested language for this option is:

- To prepare derivative works from the manuscript;
- To authorize others to make any use of the manuscript provided that it is not sold for a profit and that the author receives credit as author and the journal in which the manuscript has been published is cited as the source of first publication; and
- To distribute copies of the manuscript in connection with teaching and research by the author and by the author's employer.]

By accepting this manuscript for review, [publisher name] accepts these terms and agrees that the terms of this agreement are paramount and supersede any provisions in any publication agreement for this article, already signed or to be signed at a later date, that may conflict.

(Signature of corresponding author on behalf of all authors)

APPENDIX B

ANALYSIS OF SELECTED BIOMEDICAL PUBLISHERS' COPYRIGHT POLICIES

American Academy for the Advancement of Science

Permits author to post final manuscript in PMC with six-month embargo.

<http://www.sciencemag.org/about/authors/prep/license.dtl>.

American Chemical Society

Does not permit author to post final manuscript in PMC.

<http://pubs.acs.org/copyright/forms/copyright.pdf>.

American Medical Association

Does not permit author to post final manuscript in PMC.

<http://jama.ama-assn.org/cgi/data/293/14/1788/DC1/1>

Elsevier

Does not permit author to post final manuscript in PMC).

http://www.elsevier.com/wps/find/supportfaq.cws_home/rightsasanauthor.

Elsevier has voluntarily committed to submit the author's accepted manuscript to PMC with a 12-month embargo.

http://www.elsevier.com/wps/find/authorsview.authors/author_sponsorship_information

Nature Publishing Group

Permits authors to post final manuscript in PMC with six-month embargo.

http://www.nature.com/authors/editorial_policies/license.html

Springer

General copyright agreement does not permit authors to post final manuscript in PMC.

<http://www.springer.com/?SGWID=3-102-45-69724-0>

However, authors may pay \$3,000 US publishing fee to select Springer Open Choice, under which author keeps copyright and grants the public the Open Choice License (which is the same as the Creative Commons Attribution Non Commercial 2.5 License). This scope of this public license necessarily gives NIH the rights it needs under the Public Access Policy.

<http://www.springer.com/open+choice?SGWID=0-40359-0-0-0>

Taylor & Francis

Unclear. The relevant language in its policy is:

[Author retains] the right to post your revised text version of the 'postprint' of the Article (i.e., the Article in the form accepted for publication in a Taylor & Francis journal following the process of peer review), after an embargo period commencing 12 months (STM) or 18 months (SSH) after first publication (either in print or online), as an electronic file on an Author's own website for personal or professional use, or on an Author's internal university, college, or corporate network or intranet, or within an Institutional or Subject Repository, but not for commercial sale or for any systematic external distribution by a third party (for example a listserv or database connected to a public access server)

<http://www.tandf.co.uk/journals/authorrighs.pdf>

Taylor & Francis also permits authors to opt for iOpenAccess, under which the author pays \$3,250 US and transfers copyright to Taylor & Francis, which then grants a public license in the published version of the article. (The license is the Creative Commons Attribution Non Commercial No Derivatives 3.0 License). This option gives NIH the rights necessary under the Public Access Policy.

Wiley-Blackwell

Although Wiley has acquired Blackwell, the company appears to have maintained separate copyright policies for each unit's journals. The Blackwell agreement appears to permit posting in PMC although the embargo period is not specified.

http://www.blackwellpublishing.com/bauthor/faqs_copyright.asp#1.3

For Wiley Interscience journals, the publisher has said with respect to the Public Access Policy that “[w]e will clarify our policy regarding the deposit of articles arising from research funded by the NIH when the NIH provides more details of the mandate that becomes effective on 7 April 2008.”

<http://www3.interscience.wiley.com/authorresources/journal-man-sub.html#afteracc>

A more complete list of publisher copyright policies has been collected by the SHERPA/RoMEO project in the United Kingdom. The list is a valuable resource that provides links to a range of biomedical and other scholarly publisher's copyright policies. The list also provides summary information, including SHERPA/RoMEO's opinion about the degree to which that author may make some version of an article available on the Internet and whether the publisher's policy is compliant with the NIH Public Access Policy. In this author's opinion, the SHERPA/RoMEO assertion that a publisher's copyright policy gives the author the rights necessary to grant NIH the Public Access License is not always legally accurate and should be used with caution. For the list of all scholarly publishers, see

<http://www.sherpa.ac.uk/romeo.php?all=yes>.



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.¹ 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

¹ A copy of the Scholarly Publishing Roundtable report may be found at 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.

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

Submitted by:

Jason Snyder

jasonscottsnnyder@gmail.com

Neuroscience and Mental Health Program

Hospital for Sick Children

Toronto, Ontario, Canada

(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?

With the appearance of new open access publishing models like that employed by PLoS ONE, it is clear that there is already major growth and change in the scientific publishing market. The benefits of open access to scientific information are innumerable and, while I do not have any specific recommendations as to how new markets can be developed, I think the US government could have a tremendous effect by promoting a culture of open access to science. Services such as Pubmed and Pubmed Central have revolutionized science by making it more effective and efficient, and they both are open-access-to-science services. They have enabled scientists to see the benefits of free access to scientific information and I think they have even promoted the view that access to scientific information is a moral right. As such, more and more scientists are paying to have their publications open access. I think the growth of new markets will occur naturally as this mindset is encouraged – new open access journals will appear, new types of data repositories will be created, new software and businesses that make use of freely-available data will appear, opportunities for new forms of education will arise. It will occur naturally, in this free market economy, as these incredible resources actually become accessible to the public. It just needs a nudge and a bit of guidance.

(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?

I myself have been a biomedical (neuroscience) researcher for over 10 years and have never come across any case of intellectual property aside from the seemingly artificial one created by journals that copyright the final published article. In no case, for my 8 peer-reviewed publications, have publishers ever added anything substantial to the quality of the manuscript itself, other than aesthetic reorganization. The vast majority of my colleagues produce scientific data meant solely to further our understanding of biology. They have no intellectual property interests. Maybe I'm missing something here?

(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?

I think databases curated by the National Library of Medicine, such as Pubmed and others, have proven fantastically useful and they are completely maintained by a Federal agency. Therefore, I can easily envision that federally-maintained approaches to accessing scholarly publications would also be largely effective. On the other hand, it is likely that different disciplines and projects will depend on different aspects of the scholarly literature and so it seems that allowing a decentralized approach would facilitate the rapid growth of different uses for scientific publications

and associated data. Perhaps a federal agency could maintain the “raw data” but allow independent, non-governmental agencies to access it and re-use. Another option would be for libraries (eg at universities) to be responsible for maintaining the literature, since funds would become available as massive journal subscription fees would no longer be an issue.

(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?

I can't think of many, maybe because the world of scientific publishing is so old fashioned (it's just articles, out there to be read). There are tools like Elsevier's Sciverse, which allows individuals to make applications that search articles for specific sentences, identify “hot” papers based on Twitter conversations etc but these tools often don't search the entire literature - I'm assuming because of fees and copyrights they can only search certain publishers. I think these sorts of collaborations will become more common as access to scientific literature grows (scientists are hungry for new ways to publish, people are hungry for new ways to use and explore data) I think the field is too young for there to be many examples at this point.

One example, perhaps, is the author-pays model used by PLoS ONE. Traditional scientific publications are rather old-fashioned. They don't share the underlying data, they're merely text on a page. But in the future if other services arose that, like PLoS ONE, charged authors for the ability to archive and disseminate relevant data, multimedia, and other novel scientific contributions there could be a wealth of new business growth.

(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?

I think Pubmed is a good starting point – it is very useful to be able to search based on a scientist's name or affiliation, chemicals used in the study, discipline etc etc. Obviously there will be more data available when the whole paper is freely available and so there will be potentially more metadata to create, but the general framework could be copied. How to decide on the exact items to include in the metadata is a tough question but will probably require experimentation. If different groups are able to create metadata as they see fit, for their own particular interests, then certainly some will discover which methods and metadata are most useful. I think of it like any type of service currently available online – similar services are offered by many different groups but, eventually, one group “gets it” and discovers the way to do it best.

(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?

If federal agencies were to play a role in the archiving, generation of metadata etc (as mentioned in the above questions) then, by providing the relevant services, third parties that re-use the scientific publications and data will be encouraged to maximize the benefit (to the public etc). I mean, just making this happen (open access to science) will massively benefit the public. But, just for example, off the top of my head, what if the metadata for each paper identified whether a study utilized things like toxic reagents, human surveys, or few material resources. Personally, I'm interested in “citizen science” – the idea that anyone can do an experiment and money doesn't have to be a limiting factor. If one was a high school teacher who wanted to educate their students on the scientific process while simultaneously generating real and relevant data that could actually be published they could search the literature and find studies that were safe (and didn't involve toxic chemicals) and affordable (e.g. by simply employing human surveys or experiments that didn't require costly reagents). This may not be the greatest example but my point is that by guiding the metadata one could encourage specific uses of the publication archive, e.g. ones that benefit the public.

I don't really understand who would be at risk of burden and costs. Publishers can charge authors once, up front, to publish their paper (a la PLoS ONE) and become very profitable. Libraries will save many thousands since subscription costs will be nil. Scientists' research will be more available to the public and to other researchers making it not a burden but a benefit (even at wealthy institutions like the University of Toronto I have found many promising articles that I would probably have cited but didn't because the library could not afford to subscribe to the journal. Free access makes it easier for scientists to get credit).

(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?

Absolutely! Of course! Why wouldn't they?! In fact, this is interesting. I have absolutely zero interest in writing book chapters. Why? Because they never get read (scientific books are often rare and expensive). Were they accessible it might be a different story...

(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?

There is no appropriate embargo period. There should be no embargos. If the science is taxpayer funded and if the science provides information that is useful to the health of the public then I think it should be given to the public asap. Why, as in the case of Pubmed Central, release the study up to an entire year after it is published? What if the information could have saved someone's life? What if a scientist working at a poorer institution or poorer country is working on something similar but they cannot read the paper for another year? And then they find out the last year's worth of work was totally misguided and a waste of time, money, energy, and emotional frustration? Again, paying the publisher up front, once, to publish your study means that no embargo is necessary.

Please identify any other items the Task Force might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.

In the future I can envision a new world, where scientific progress is not driven by competitiveness but by a pure desire to discover and share information for the betterment of the world. It is currently very far from this ideal but at the same time there is so much turmoil that it is clear that change is around the corner. Your interest in these issues is very encouraging and I thank you! I am sorry that my responses could not be more helpful in places.

-Jason

[Assigned ID#]

[Assigned Entry Date]

Marce Abare

Tim Anderson

Colin McCluney

Kristin Huntoon, PhD

Elizabeth Wiley, JD, MPH

American Medical Student Association

45610 Woodland Rd

Suite 300

Sterling, VA 20166

On behalf of the American Medical Student Association (AMSA), we are pleased to offer the following comments. AMSA is the oldest and largest independent association of physicians-in-training in the United States. Founded in 1950, AMSA is a student-governed, non-profit organization committed to representing the concerns of physicians-in-training. AMSA members rely on publicly funded research during their training as students and later as practicing physicians. Ensuring medical educators and trainees have unfettered access to the most cutting edge research will produce a health professions workforce better prepared to serve our nation's health care needs and compete in the global economy. On behalf of more than 32, 000 members across the country, AMSA supports open access to publicly funded research.

[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?

[Comment 1]

All publicly funded research should be made unconditionally and immediately available for public use and reuse. In contrast to current policy, "open access" refers to the free, immediate, unrestricted availability of high-quality, peer-reviewed scholarship over the internet, combined with the rights to use this information to its fullest possible extent as long as proper attribution to the original article is maintained. Such open access is critical to ensuring the quality of health professional education. Open access allows the most cutting edge research findings to be integrated into medical education to ultimately better prepare future physicians for evidence-based practice, extend standards of care across the board to all medical centers and have baseline therapies to improve upon.

Open access also has economic benefits. Empowering American health professions students with access to the latest research will improve students' ability to compete in the global marketplace and boost US biotechnology competitiveness. Investing in the knowledge and education of today's students will ensure that resources are truly invested in advancements and improvements

on known phenomena and efficient discussions of this knowledge to be exchanged across multidisciplinary groups.

Strong public access policies help level the playing field for students. Open access means American students across institutions will be better prepared to contribute when it's time to put their education to use in the private or public sector. For medical students, this entails providing patient care as a resident and, eventually, attending physician. Having unrestricted access to the most up-to-date clinical trials and evidence-based practices is essential to providing high quality care. For physicians and trainees practicing in rural and underserved communities in the United States and abroad, such access is not readily available as clinics and hospitals are not able to afford expensive journal subscriptions.

[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 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?

[Comment 2]

Open access policies are compatible with the existing intellectual property legal framework to protect interests of publishers, scientists, federal agencies and other stakeholders. While NIH policy currently provides for "fair use," broader access to this information is necessary to realize scientific and commercial benefits. Appropriate licenses - such as Creative Commons CC-BY licenses - would support access sooner than current term of copyright permits.

[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?

[Comment 3]

A majority of practicing clinicians and medical students utilize peer-reviewed evidence to guide clinical decision-making on a daily basis. Federal agencies are the most appropriate stewards of centralized repositories for the publicly funded scholarly literature that informs clinical practice. For example, Pubmed Central (PMC), among the most widely used online tools, is the free digital archive of biomedical sciences journal literature at the NIH National Library of Medicine and the designated repository for papers submitted in accordance with the NIH Public Access Policy. While critical to the ability of physicians and students to locate vetted evidence for patient care, this policy still delays access to research by a year and does not include other federal funding agencies such as the Centers for Disease Control or the non-medical research arms of the federal government.

Nonetheless, PMC sees roughly 500,000 unique users each day and is an invaluable resource for medical students, physicians and researchers. PMC is designed to store and cross-reference diverse data sources using a common format enabling efficient searches for full-text articles throughout the entire database to quickly locate pertinent information. PMC also allows integration of its database with other information resources and collaboration with international agencies that share similar goals—namely, free and immediate online access to digital biomedical literature to expand knowledge, innovation and evidence-based clinical care among practitioners and researchers. Having access to a non-biased, non-commercial repository of up-to-date scientific information enables medical students, physicians and researchers to be sure that they are practicing and innovating at the forefront of our collective knowledge.

[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?

[Comment 8]

Due to the rapid pace of medical discovery and therapeutic and technological advancements, it is critical that educational materials reflect the most up-to-date information possible. Physicians-in-training, like other professional students, should have immediate and unfettered access to peer-reviewed literature that results from taxpayer-funded research particularly in light of its implications for patient care and establishment of training and practice norms.

Coursework within problem-based learning and organ systems curricula typical of US medical training institutions is generally completed during one- to three-month-long sequences. An embargo period, therefore, necessarily excludes the latest information that could otherwise be available to students. In the context of hands-on training in teaching clinics and hospitals, the stakes are notably higher. From stroke treatment and care to the management of sepsis, acute coronary syndrome or other health emergencies, many common diseases and conditions are under active research. Appropriate management requires immediate access to the most cutting edge research findings.

The National Institutes of Health as well as other funding agencies worldwide currently use an embargo period from 0-12 months. To date no publisher has provided evidence of financial loss as a result of this policy. An embargo in which a stop-date between 0-12 months is determined by the author has been shown effective across multiple disciplines and hundreds of journals. When considering an embargo period, it is important to weigh the real costs associated with a delay in access. An embargo forces students and clinicians to rely on outdated information, impedes exchange of best practices and follow-on research, diminishes the ability of students to self-educate, and threatens patients' right to receive proper clinical care and recommendations based on latest medical evidence.

To Whom It May Concern,

As a humanities researcher, I am vitally interested in policies resulting from your discussions about public access to peer-reviewed scholarly publications. The policies you adopt will very likely have repercussions for all scholarly research, especially that which is supported by the National Endowment for the Humanities, and most especially that which is supported by the NEH's Office of Digital Humanities. Other federal agencies, too, support humanities research: the Department of Education, the Library of Congress, the Smithsonian, the Institute of Museum and Library Sciences, and the National Archives should all be included as you formulate answers to the questions you pose in your request for information.

You ask, "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?" Substituting "scholarly enterprise" for "scientific enterprise," I can certainly speak to the latter point: policies that ensure that federally funded publications are open will improve scholarly productivity in all fields. The fact is that scholars often communicate among themselves using the same communication tools that the public uses: e-mail lists, Google Groups, blogs, Facebook, Twitter, and so on. In order for research to be shared in these media, it must be shareable, and to be shareable, it must be open. If scholar A's institution subscribes to a particular journal and scholar B's does not, a link sent to scholar B by scholar A will not work. Even within a university, scholars will find that a link they send their students or graduate students often does not work if those students are off-campus. The "paywall" puts significant obstacles in the way of spreading information, which is the heart of scholarly productivity. Scholars who do not learn about relevant information may spend many months or even years in futile pursuits.

Those of us who work in the digital humanities are particularly aware that public access to our research makes our work more widely known by other scholars as well as by the public. The digital humanities researcher Melissa Terras has also written about the importance of public access in raising a publication's profile among other scholars and the public: in her piece "What Happens When You Tweet an Open Access Paper," she traces the increasing popularity of a peer-reviewed paper that she posted in an open repository: "Prior to me blogging and tweeting about the paper, it got downloaded twice (not by me). The day I tweeted and blogged it, it immediately got 140 downloads." The downloads only increased, and, ultimately, she wrote, "This post was mentioned in the Times Higher [Ed] last week, and the paper has now been downloaded 805 times in total." Note that open access to her paper, and her ability to link directly to the paper from social media such as her Twitter account and her blog, ultimately led to reporting on her work in a major newspaper.

In 2011 at the Modern Language Association annual meeting, I gave a paper with the tongue-in-cheek (but true) title "Your Twitter Followers and Facebook Friends Won't Read

Your Peer-reviewed Article if They Have to Pay for It, and Neither Will Strangers," in which I related the experience of discovering that several members of my social network, both scholars and non-scholars, were interested in reading my arcane work on Victorian poetic form if they could gain access to it freely. That (very short) paper is freely available at <http://amandafrench.net/blog/2011/01/07/twitter-facebook-article/> should you care to read it. In that paper, I cited a study by Jason Priem and Kaitlin Light Costello presented at the 2010 meeting of the American Society of Information Science and Technology titled "How and Why Scholars Cite on Twitter." As I wrote,

It was one of my most clicked-on links for the year, with 118 views—many of the links I tweet to news articles and so on get only thirty or so clicks. The authors studied a sample of 46,515 tweets from twenty-eight scholars — seven scientists, fourteen social scientists, and seven humanists — and reported that "In our sample of tweets containing hyperlinks, 6% were citations. Of these, 52% were first-order links and 48% were second-order." By this, they meant that 52% of the links went directly to peer-reviewed work, while 48% were links that went to non-peer-reviewed work about peer-reviewed work: blog posts and news articles, for instance.

One of the main reasons that scholars tweeted these "second-order" links was that they worked for everyone: "[S]cholars may prefer to link directly to the article when it is open access but will resort to second-order links to bypass paywall restrictions. Participants were attracted to open-access articles for Twitter citations; Ben said 'I would certainly be much more likely to link to things if they were more readily available.' "

That study, as well, is openly available at http://mail.asis.org/asist2010/proceedings/proceedings/ASIST_AM10/submissions/201_Final_Submission.pdf. As I hope is clear, I frequently make use of (and share) conference papers for my research, and therefore, I give a decided "yes" to your question, "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 format in which scholarly research is published should make no difference to its public availability.

It is true that the scholarly work I have mentioned so far has not been explicitly funded by the U.S. government. However, I reiterate that humanities researchers do indeed receive federal funding, and I am no exception. In 2009, a grant from the National Historic Publications and Records Commission, the funding arm of the National Archives, allowed me to work for a year on a project to update the curriculum of the Archives and Public History graduate program at NYU with department chair and principal investigator Dr. Peter Wosh. We were more than happy to distribute the results of this project publicly, online as well as through scholarly channels such as the annual meeting of the Society for American Archivists and the Mid-Atlantic Regional Archives Conference. This work would certainly come under the aegis of the Issa Research Works Act, an act which troubles me deeply. For another example, I have applied for a Kluge Fellowship at the Library of Congress for the year 2012, and although the funds for this fellowship come from a private foundation, it is possible to likely that any work I produced while doing research at the Library of Congress would also qualify as federally funded research.

Finally, I work at the Roy Rosenzweig Center for History and New Media (CHNM), a humanities research center which has benefited greatly from federal funds, and which as a body is committed to public access to scholarly publications. The Center itself, which has conducted more than \$20 million in grant-funded research, relies on a nearly \$3 million endowment achieved with the assistance of two Challenge Grants from the National Endowment for the Humanities. CHNM's work, like that of any scientific research center, relies on both federal and private funding: CHNM's work has been recognized with major grants from the National Endowment for the Humanities, the Department of Education, the Library of Congress, the Institute of Museum and Library Services, the National Historic Records and Publication Commission, and the Sloan, Mellon, Hewlett, Rockefeller, Gould, Delmas, and Kellogg foundations. Since 1994, CHNM has been a leader in improving students' understanding of history and the humanities through digital media, in building digital archives and mounting online exhibitions, and in developing software tools for scholarship. In 2010, CHNM's websites had almost 500 million hits and nearly 20 million unique users, and its software tools are used by more than a million scholars and students every day.

I hope to have convinced you that humanities researchers and the federal agencies that support them are interested parties in the development of policies related to public access to peer-reviewed scholarly publications. Please consider, too, the tremendous extent to which research done at institutions of higher education is made possible by the tax policies of the federal government: such research belongs to the public. Thank you for your work.

Amanda L. French, Ph.D.

--

Amanda L. French, Ph.D.

<http://amandafrench.net>

<http://twitter.com/amandafrench>

Skype: amandafrenchphd

AIM: habitrailgirl

Date: January 12, 2012

Name: Marilyn S Billings

Email: mbillings@library.umass.edu

Affiliation: University of Massachusetts Amherst, University Libraries

City, State: Amherst, MA

Summary:

Thank you for the opportunity to comment on "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research." These comments are submitted on behalf of the University of Massachusetts Amherst. UMass Amherst is the flagship campus of the University of Massachusetts system, sitting on nearly 1,450-acres in the scenic Pioneer Valley of Western Massachusetts, 90 miles from Boston and 175 miles from New York City. The campus provides a rich cultural environment in a rural setting close to major urban centers.

It enrolls over 21,000 undergraduate students and over 6,000 graduate students in 86 bachelor's degree programs, six associate's, 72 master's and 50 doctoral programs in eight schools and colleges. There are over 1000 full-time instructional faculty at this public land grant university that has the education of the public as one of its key missions. UMass Amherst is one of the nation's top public research universities, listed as a Carnegie Research Extensive University. The campus attracts over \$140 million in externally sponsored research each year, demonstrating its contribution to Massachusetts's position as a technological and economic leader. Funding supports the creation of new knowledge and its translation into the technical innovations and scholarly works that create opportunity for students, faculty and the public. The University Libraries is the largest public academic research library in Massachusetts. It has led conversations about open access to scholarly and creative works since 2001, actively engaged in promoting conversations through workshops, Faculty Senate debates, and providing solutions to open access through its digital repository ScholarWorks @ UMass Amherst.

The University of Massachusetts Amherst is a member of the Boston Library Consortium (BLC), the Association of Research Libraries (ARL), and is an affiliate member of the Coalition of Open Access Policy Institutions (COAPI) so many of these comments are reflections of comments provided by those entities that we want to reinforce as critical to this institution and its values. UMass Amherst fully subscribes to the principle that taxpayers are entitled to access the results of publicly-funded research, research funded by their tax dollars, immediately, and that taxpayers are entitled to fully reuse those results. The current NIH Public Access Policy, implemented in 2008, applies to the results of approximately one-third of all federally funded scientific research, and a significant amount of the research taking place at UMass Amherst. The NIH policy, while it is not without limitations, has been enormously successful in opening the results of NIH research to a broader audience - to the benefit of science and the general public. There is an urgent need for the federal government to adopt a comprehensive public access policy approach applicable to all major research funding agencies, one that would both extend and improve upon the current NIH policy. UMass Amherst agrees with the COAPI recommendation for the federal government to develop a policy framework that 1) is as uniform as possible for all agencies, 2) is mandatory for all researchers funded in whole or in part by those agencies, 3) results in rapid and open access to the results of peer-reviewed, government-funded research, and 4) allows flexible rights of reuse.

(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?

Comment 1:

Successful development of markets related to access and analysis of government-funded peer-reviewed publications depends in large part on the speed with which research information is made available and the terms under which it can be used. The combination of rapid public access and liberal reuse rights will drive software development that facilitates new types of information discovery and tools for research.

It will create the capacity for new information-based business models that draw on the innovations in information technology, such as the semantic web, which fosters sharing and reuse of information across applications and community boundaries. Full open access in this sense will also foster commercialization of products that increase access to and awareness of specialized research information.

All of these potential capacities will be reduced to the extent that access is delayed through embargoes or that reuse rights are limited unnecessarily.

Text mining, data mining, other forms of information computation, and the creation of derivative works are examples of new research and information dissemination capacities that can be enabled through appropriate reuse rights. For example, the University of Massachusetts Center for Intelligent Information Retrieval (CIIR) is one of the leading research groups working in the areas of information retrieval and information extraction. The CIIR studies and develops tools that provide effective and efficient access to large networks of heterogeneous, multimedia information. In addition to commercial applications, such tools could also be valuable to funding agencies by allowing them to monitor research developments in specific fields as part of the process of setting funding priorities.

A broader federal public access policy framework will also foster the continued development of open access journals (which now number more than 7,000 titles) and the transition of traditional publishing to open access business models ? again to the benefit of science, economic development, and public welfare. Commercial firms ? both new firms such as Hindawi and existing ones such as Springer ? are clearly realizing the economic benefits of open access through the creation of profitable new journals that follow open access business models. Nonprofit publishers are also experimenting with open access publishing and thereby extending the reach of the research they disseminate. The growth of publicly accessible research information will encourage scholarly publishers (both nonprofit and for-profit) to transition to open access in ways that meet both their scholarly missions and their economic interests. A broader federal public access policy framework will thus both add to and encourage the continued growth of openly accessible research information.

Numerous studies have demonstrated that openly accessible research information reaches wider audiences and produces more citations than research published under access restrictions. Recent studies are also showing that openly accessible research produces more diversity in follow-on research. It encourages contributions by participants who would have had no opportunity to contribute in an environment with access controls. It thus increases the potential for innovation and the interdisciplinary application of research through a larger pool of participants.

A government-wide public access policy or policies can be implemented by leveraging existing infrastructure in ways that minimize duplication of effort. The investments in software and other

resources that already support NIH's PubMed Central and similar repositories can be utilized by other agencies either individually or in a federated model. A comprehensive federal public access policy framework will have the added benefit of increasing the effectiveness of government research funding. One of the primary motivations of the NIH policy was improved documentation of the outcomes of sponsored research. A comprehensive federal policy will bring that benefit to all of the major scientific research funding agencies. It will also provide congressional appropriators and authorizers better information to assess the value of existing expenditures and better target strategic funding priorities. It will thus increase agency accountability and support informed, transparent, and evidence-based budget and policy decision-making in accordance with the Obama administration's emphasis on open government. In order to maximize the investments in cyber and information infrastructure, advance science, and promote innovation, free immediate access with full reuse rights to federally funded research literature would achieve the most benefits. There should be no restrictions placed on use of this literature or on who is able to use these federally funded information resources. This would be consistent with existing federal policy, the Paperwork Reduction Act and Circular A-130, concerning government information. If an embargo period is deemed necessary, it should be as short as possible.

(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 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?

Comment 2:

Key to the success of advancing research, and spurring innovation and commercialization, will be to provide unfettered access to federally funded research resources and permit the widest possible use within the law. If the goals of agency policies are to foster the development of science, encourage economic growth, and serve the public's interests in the broadest sense, then it will be important to construct the licensing framework for the policies according to principles that will facilitate those goals. Doing that requires no change in copyright law. It is only necessary to structure the licenses that authors grant to the agencies (as a condition of their funding) and the licenses that the agencies grant to the public in ways that facilitate both access to and maximum reuse of research information. A Creative Commons attribution license is an example of a license that would fulfill those purposes. Such a license would allow authors to receive full credit for their works while also creating great flexibility in terms of how their works can be used by others. Use of these licenses permits the user full use rights to mine data and text, and manipulate, reuse, and integrate data and information in publicly accessible digital repositories. Licenses that allow only for access to research information but not subsequent reuse or redistribution to colleagues are unnecessarily restrictive. Unlike the NIH policy, systematic downloading of articles should be allowed in order to facilitate flexibility in terms of reuse, for example, by programs that compute on the textual corpus. Since the licensing framework for the agency policies would be non-exclusive, authors would remain in a position to transfer appropriate rights to publishers. Like the NIH policy, agency policies should be mandatory, with authors required to deposit their final (post-peer-review) manuscripts in publicly accessible repositories. In view of that, publisher transfer of rights agreements for federally funded research articles could not be structured in ways that conflict with the licenses that researchers grant to the agencies. Publisher economic interests can be protected by brief embargo periods during which the

use of the research information would be governed either by fair use under copyright for journals in print form or ? in the case of electronic journals ? by the provisions of license agreements. Once the embargo is lifted, then full reuse rights should be associated with the research literature. Such an approach takes into account the needs and interests of all stakeholders. Regardless of where the publications reside, full reuse rights are essential elements of an effective policy. Metadata standards would include a full citation to the publisher copy of record. Such a policy framework would balance the needs and interests of research authors, agencies, publishers, and the general 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?

Comment 3:

The University of Massachusetts Amherst believes that a centralized or federated approach managed by the federal government is the most appropriate and effective strategy for ensuring interoperability as well as effective search mechanisms and analytic tools. Federally managed approaches are also the most feasible way to facilitate new research capabilities related to reuse (such as text and data mining, creation of derivative works, information discovery tools, and commercialization of products that increase access to and awareness of specialized research information). Even with carefully crafted regulatory requirements, it is clearly more difficult to establish and maintain such capabilities under a decentralized framework that includes partners outside the federal government.

The federal government has a long-term interest in making the results of its funded research permanently available and has a long history of ensuring that there is long-term preservation of and access to works via centralized deposit. For example, through a provision in the Copyright Act, printed copyrighted and public domain works are placed on deposit at the Library of Congress. Beginning in 2010, the Library extended this deposit requirement to include electronic-only serials.

The National Library of Medicine has been providing long-term preservation of and access to biomedical information for 175 years.

More recently, NIH implemented the NIH Public Access Policy, which is a natural continuation of this role. It is appropriate and necessary for the federal government to ensure that the long-term preservation of and access to these resources is undertaken and with appropriate use rights for the federal government and users alike. It is the only entity that has the capacity to make the full corpus of federally funded works publicly accessible, to establish and enforce standards of interoperability that ensure search access across repositories, and to establish and maintain an infrastructure that will allow new services and products to be built from publicly funded information.

As more and more institutions and organizations establish digital repositories, there will be many sites providing access to federally funded research literature, nationally and internationally. Any US policy must ensure that these repositories of federally funded research resources are interoperable and accessible with appropriate use rights both now and in the future, regardless of who is curating these resources. As we have learned, long-term preservation of and access to digital resources requires use; dark archives are not an option. To ensure that there is not deterioration of these digital resources and that there is a valid record going forward, continuous use is required.

Primary reliance on a federal government role does not preclude private or third parties from participating in a decentralized approach. However, any decentralized approach that involves entities outside the federal government, whether public or private, would need to provide all of the capacities

described above ? public access, interoperability, search functionality across repositories, adherence to standards, long-term archiving and preservation, openness and accountability, and the potential for creative reuse for research and commercial purposes. In addition, clearly delineated roles and responsibilities will be key. If the federal government found that a decentralized approach was feasible and decided to rely on it heavily, then government agencies should maintain mirrored and accessible versions of the decentralized repositories in order to protect the public's investment and ensure accountability. It will be critical to stipulate that if a provider for some reason is unable to meet its obligations of service?either short-term or long-term?a migration path should be in place to recover the resources. The federal government's stewardship over this valuable public good is critical.

(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?

Comment 4:

Academic research libraries have developed extensive experience and expertise in creating and managing digital archives designed for long-term preservation and access. Examples include arXiv (now managed by the Cornell University Libraries), the digital repositories of several research universities such as ScholarWorks @ UMass Amherst, and the HathiTrust, a major partnership of research libraries and research institutions that is designed to preserve digital books and broader cultural heritage. Given their expertise and focus on long-term preservation and access, research libraries could be important consultants in the development and implementation of federal, interagency and public/private partnerships in a public access policy. Some research universities could also partner with federal agencies to develop repositories for specific subject areas.

For example, UMass Amherst Libraries are engaged in the development of a subject repository for Nano manufacturing (InterNano) and received an NSF grant three years ago to develop a beta subject repository for materials in the ethical conduct of research in the sciences and social sciences (ESENCE). In addition, we note that some academic and research institutions have partnered with research funders to provide their permanent archives.

Publishers could be encouraged to participate in public-private partnerships by voluntarily providing the final published versions of articles after limited embargo periods that ensure their subscriptions and licensing revenues. However, given their focus on immediate income and the fact that they tend not to have long-term time horizons, commercial publishing firms in particular should not be relied upon solely for digital archiving. It should be obvious that long-term archiving and public access will be made much more difficult when corporate acquisitions, mergers, or business failures occur. For that reason, publishers should provide archiving and public access for the results of federally funded research only if the publishers' sites are mirrored by sites maintained by the federal government or by institutions that provide greater certainty of long-term preservation and access, such as research institutions. Publishers would also have to be able to comply with detailed rules for user interface, access formats, and interoperability.

(Question 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?

Comment 5:

The development of interoperable search, discovery, and analysis capacity across disciplines and archives depends on the creation of carefully crafted metadata standards that are implemented for all archives containing the results of federally funded research. It is critical that metadata be both machine-readable and machine-interoperable if agency policies are to realize their full potential. Metadata standards for archives should be designed to facilitate the functions of use, reuse, and analysis described above.

Federal agencies, through their public access policies, are best positioned to ensure the creation of metadata standards that will meet the functional goals of their policies. The research library community, including the Library of Congress and organizations such as OCLC, has developed a variety of metadata standards that have been endorsed by standards organizations (NISO, ISO, etc.). These can be drawn upon in developing a broad federal metadata specification.

The specification should support multiple metadata standards in order to develop metadata that is as rich as possible. Some of the primary goals of the specification (along with examples of related standards) would be to: 1) provide institutional information for published sources (grant IDs, funding organization, Institutional Identifier, etc.), 2) provide descriptive information for both the repository and published versions (Dublin Core, ORCID), 3) support searching through keywords as well as controlled vocabulary schema appropriate to disciplines, 4) incorporate abstracts, 5) facilitate full text searching and web crawling, 6) support metadata harvesting (OAI-PMH), 7) establish relationships through semantic web standards (RDF), 8) support usage tracking (COUNTER), 9) support description of related data (DataCite Metadata Schema), 10) support data exchange standards (JSON), and 11) document intellectual property rights.

It's especially important for metadata to support the capacity for machines to access and analyze both the publications themselves and the underlying data that support them in those instances where that data can be made openly accessible.

(Question 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?

Comment 6:

The benefits of public access policies to taxpayers will be realized to the extent that publicly funded research results are made openly accessible. The history of the development of the NIH Public Access Policy demonstrates conclusively that a broader federal public access policy (or policies) must be mandatory. The rate of compliance with the NIH policy increased dramatically following the end of the voluntary policy and the adoption of the current mandatory policy.

Average manuscript submissions have grown from approximately 1,000 per month prior to April 2008 (the date of adoption) to current levels that are well over 5,000 per month (for the most recent twelve-month period). See: <http://www.nihms.nih.gov/stats/> A broader federal policy must be consistent across all agencies in its requirements and mandates. Uniform requirements and procedures across all agencies will reduce burdens on researchers (who often hold grants from multiple agencies) and on the

institutions that support their compliance. Uniformity will reduce complexity and that in turn will reduce the time needed to educate researchers about policy requirements, to deposit articles, and to deal with deposit and compliance problems. Uniformity will also work to increase compliance rates. Publisher interests, for example those related to embargo periods and any deposit of final published versions of articles, are also best served by a uniform approach. Procedures should include standard criteria for what should be deposited as well as clear instructions for the deposit process. Existing grant management systems should also be integrated into the deposit process to facilitate agency and public accountability.

(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?

Comment 7:

Yes, definitely. There are other important types of scholarly communications beyond the peer-reviewed scholarly journal articles.

Monographs and book chapters, conference presentations, theses and dissertations, working papers, and datasets are also increasingly being made available via open access or public access policies.

Policies covering ETDs (electronic theses and dissertations) are also common, well developed, and generally supported by students as well as their faculty advisors. At the University of Massachusetts Amherst, these policies and procedures were developed in coordination with the Graduate School, the Faculty Senate Graduate Council, and the University Libraries. Open access ETDs have been captured in the ScholarWorks digital repository since 2008. It should be noted that, since there are different terms and conditions associated with each of these educational materials, it will be important to distinguish the various approaches to each type of scholarly output.

The related RFI concerning data policies indicates that data policies may be differentiated from peer-reviewed literature and other types of scholarly output as different terms and conditions may apply. Nevertheless, data is central to the scholarly and research enterprise and should be treated equally in terms of importance to the scholarly record and tenure and promotion.

(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?

Comment 8:

Advances in science and other scholarly disciplines build upon prior knowledge and the sharing of information. The scientific, economic, and public benefits of providing access to the return on our nation's investment in research diminish to the extent that access is delayed or denied. Immediate access at the time of publication is therefore ideal in terms of overall policy goals. It is time to accelerate such advances by significantly decreasing or eliminating embargoes to currently available, published research resources. Nationally and internationally, embargo periods of 12 months or less are the standard for journal publishing (<http://highwire.stanford.edu>). It is important to note that some publishers who have expressed concern in the past that public access would result in loss of subscription

revenue have changed both their views and their practices. If it is demonstrated through empirical evidence that embargoes are necessary, the University of Massachusetts agrees with members of COAPI that a uniform embargo period of six months or less should apply across all funding agencies. Such an approach has the benefits related to consistency across disciplines and would speed research access while also taking into account publisher interests. It is also important to note that the NIH Public Access Policy (with an embargo period of 12 months) is not representative of international biomedical funder policies. A six-month embargo is now standard (<http://roarmap.eprints.org/>). In addition, there is no evidence to support that academic and research libraries either have considered ? or would in the future consider ? public access to federally funded research to be an adequate substitute for journal subscriptions or licenses.

If a decision is made to adopt different embargo periods for individual disciplines or sub-disciplines, shorter embargo periods (less than six months, for example) should apply to rapidly changing fields and those where research results often lead directly to commercialization.

We would emphasize that the burden of proof for the need for embargoes should rest on those who believe they are necessary. The benefits of public access are clear. In the absence of empirical evidence clearly demonstrating the need for embargoes, immediate public access should be the norm, since it is the best way to foster innovation, competition, economic growth and scientific progress.

Please direct any questions about this response to:
Marilyn Billings
Scholarly Communication Librarian
University of Massachusetts Amherst

January 12, 2012

Office of Science and Technology Policy
Executive Office of the President
Washington, D.C. 20502

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

Dear Sir/Madam,

The American Association for Cancer Research (AACR) is pleased to respond to the Office of Science and Technology Policy Request for Information regarding “Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research.” AACR, a not-for-profit association with more than 32,000 members, is the oldest and largest non-governmental scientific organization in the world dedicated to advancing cancer research. The programs and services of AACR foster the exchange of knowledge among scientists involved in cancer research. AACR publishes 7 peer-reviewed scientific journals and a magazine for the general public; convenes topical scientific think tanks, conferences, workshops, and an annual meeting; offers fellowships and grants; raises public awareness of the progress and cause for hope in cancer research; and advocates for federal research funding.

As a scientific society publisher, AACR is dedicated to widely disseminating the results of research and supporting the scientific enterprise. AACR invests in the journals it publishes and the articles within them through various activities including peer review, copy editing, composition, electronic tagging, online journal hosting, printing, distribution, archiving, promoting the results to various audiences, holding editorial retreats, and applying new online features and functions.

AACR voluntarily makes *all* journal content freely available 12 months after publication through our online journal sites. AACR’s decision to make our content freely available after a 12-month embargo period was based on the particulars of our publications with the desire to sustain them and reinvest in the many activities the Association supports that contribute to the scientific endeavor. We join many other publishers in this regard—working together without government mandates to provide more access to scholarly content than ever before. Federal mandates that compete with the work of private-sector publishers jeopardize the sustainability of a robust peer-review publishing system which the vast majority of scientific researchers consider first-rate.

Publishers have an excellent record of providing long-term stewardship and broad public availability of the peer-reviewed scholarly publications that report on, analyze, and interpret federally funded scientific research. We believe that the best approach to achieving greater public availability to peer-reviewed content and to improve productivity is through public/private collaborations with all stakeholders. We appreciate the opportunity to respond to the Request for Information and provide comments in response to Questions 1, 3, 6, and 8. We look forward to working with the Office of Science and Technology Policy and other stakeholders to further consider public access.

(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?

Economic/Productivity Considerations

We believe that investment made by publishers has contributed to U.S. job creation and economic growth. Fifty thousand Americans are now working in the publishing industry. Access to peer-reviewed articles produced by private-sector publishers of research supported by federal funding should be considered with publishers based on the evidence of benefit while weighing the risk of destabilizing the publishing system upon which researchers and society depend for scientific integrity and, dissemination of information. This assessment should be determined by cooperation and collaboration, not by regulation.

AACR, like many American scientific society publishers, reinvests in the scientific enterprise and fosters its innovation and advancement. Long before the NIH Public Access mandate, AACR determined that its business model and mission would include free access to *all* content on our journal sites after a 12-month embargo, while making some other content immediately available. The creation of the costly PubMed Central database duplicated these efforts and spent federal funds that could have been better used on research itself. A more efficient method would be to leverage the valuable work already done by publishers by developing cooperative linking. Publishers could provide the federal agencies with the metadata and abstracts of federally funded peer-reviewed articles so that it could build an aggregated site linking together all content derived from the research.

Existing and New Markets

Lay Public

One market for federally funded research findings is to provide the general public with access to the information for which their tax dollars have paid. Much of the research that is funded by the NIH is pertinent to health, and people facing health issues are increasingly turning to online searches to find out more about prevention and treatments. Although many publishers make their peer-reviewed articles available to the lay public, the original research papers are often very technical and can be of limited use to much of the population. Many offices within federal agencies work to translate these findings into products that can be used by the average patient. Government examples of this type of compilation include the National Cancer Institute's Physician Data Query Program (PDQ) and the Agency for Healthcare Research and Quality's (AHRQ) Effective Healthcare Program (EHP), both of which prepare patient summaries and decision aids for various conditions and treatments. If the government would like to grow markets for lay consumption of research findings, then the focus should be on programs like PDQ and EHP rather than on simply providing access to original research articles having federal funding. However, the government can link its content to the final article published on the journal site so that it is available for the "expert" patient.

One way in which AACR and many other publishers have demonstrated our commitment to addressing patient and caregiver desire for research articles is by making them freely available upon request. Another example of a cooperative publisher initiative is patientINFORM (<http://www.patientinform.org/>)—a program that brings information from voluntary health organizations

together with scholarly articles for patients and caregivers. There is also a new initiative driven by the Association of American Publishers/ Professional/Scholarly Publishing Division; International Association of Scientific, Technical & Medical Publishers; and the Copyright Clearance Center. AACR is engaged in this pilot program, which aims to make it easier for patients and caregivers to obtain access to AACR articles as soon as they discover the material online. Finally, AACR continues to make any article immediately available to any patient or caregiver who requests it.

Research Community

Efficient research relies on the most complete and up-to-date understanding of a given research field, and journal articles are the gateway to that current understanding. While the information contained in a federally funded article is of importance, the advent of indexing and interoperability has given researchers easy access to information that is separated by one or two degrees from the article being accessed. This is done through active links in the bibliography to cited papers, and by the ability to see and access other articles that have cited the article since the time of its first publication, a feature that AACR makes available to readers. Dynamically updating a list of other articles that have referred to the article in question is a way for researchers to continually keep up to date on the state of the science. The value-added feature found on many publishers' sites points readers to the most recent and relevant articles, not those limited to a specific funding body. Rather than trying to create an accessible repository of documents that already exists on publishers' sites, the government should publish the metadata and abstracts of federally funded work and link users back to journal sites where the network of connected research is more fully presented.

New markets are available to the government, when the focus of the question moves beyond peer-reviewed publications. The many contributions of federally funded research that never gets published in journal articles are untapped. As the *British Medical Journal* stated, "...fewer than half of trials funded by NIH are published in a peer-reviewed biomedical journal indexed by Medline within 30 months of trial completion. Moreover, after a median of 51 months after trial completion, a third of trials remained unpublished." (BMJ 2012; 344 doi: 10.1136/bmj.d7292.) In addition, journals rarely publish research with negative findings, but access to this information can surely benefit the research community.

Agencies could grow new and related markets that originate from federally funded research by providing access to the final grant research report and the data that underpin that research. The final research reports of some agencies are already publicly available. Broadening these requirements and presenting report outcomes in a timely, consistent, and useful format with interactivity among agency sites would be of great value. For research that eventually gets published in peer-reviewed journals, linking to and from the research reports, data, and the final article on the publisher's site would assist scientists in analyzing and interpreting information. It would increase productivity, eliminate duplication of work and products, and free up resources that the government is currently spending on duplicating efforts of publishers, (e.g., PubMed Central). Publishers have already been working in partnership with groups such as CrossRef to develop standards for data and metadata to make research more readily searchable and discoverable. Collaboration on these and other efforts would benefit the scientific enterprise.

(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?

Although one might presume that a centralized approach would be the most efficient way to manage access to publications with federal funding, it is limiting in many ways. A centralized site created and managed by the government, such as PubMed Central, is costly, replicates the work that publishers have already done, and limits users access to content and a variety of functionality. A distributed, decentralized approach feeds innovation that is sparked by competition among publishers and other companies. The decentralized approach that includes information of all types, not just those derived from government-funded research, is of greater value to the user. Publishers and other companies have already successfully been promoting interoperability, advancing search, and developing analytic tools. Both not-for-profit and commercial publishers working in a competitive environment have moved quickly and decisively to introduce new technologies that meet researchers' demands for faster and more user-friendly delivery of scholarly information. Some examples of new technologies are mentioned below.

The Digital Object Identifier (DOI) is a unique identifier for each piece of content in a scholarly publication. The DOI, which has now been assigned to more than 50 million items, is a standard in the publishing industry with nearly 1,000 publisher participants. Work has been ongoing to standardize metadata for such identifiers for individuals, author contributions, and funding information. Federal agencies should work with publishers and other stakeholders who have expertise in developing and promulgating metadata to ensure standardization across disciplines and share best practices.

Publishers have collaborated with librarians and database providers to establish COUNTER (Counting Online Usage of NeTworked Electronic Resources), which has produced an international set of standards and protocols governing the recording and exchange of online usage data. These standards enable publishers to better understand the usage patterns of their digital content and for librarians to track the usage of their digital collections. A variety of Internet search engines, abstracting services, and other tools do an excellent job of ensuring the discoverability of research, and innovations and advancements of these and other tools continue to be developed.

A centralized registry of unique identifiers, Open Research and Contributor ID (ORCID), has been created to address the author name ambiguity problem in scholarly publishing. These identifiers can be linked to the researcher's output to enhance the scientific discovery process and to improve the efficiency of research funding and collaboration within the research community.

At a time of shrinking federal resources, use of funds to replicate work that is already being done by private-sector publishers is unwarranted. If the federal government is concerned that the long-term stewardship for peer-reviewed articles to which they fund the research is at risk, the Library of Congress and/or other agencies should be charged with creating an archive that can be used for access to these articles if they became unavailable.

(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?

Federal agencies invest in the research but it is the publishers who invest in the scholarly articles. As a result of the value-added activities AACR provides to its journal articles, no research article is published as it was originally submitted. These unique contributions strengthen the research literature and improve its accessibility—without direct taxpayer support. AACR invests in the submissions its journals receive through the work of our scientific editorial boards and staff who consider the submissions, identify peer reviewers, evaluate the peer reviewers' comments, and analyze requested changes or rebuttals regarding revised manuscripts. Work of this type is done on many more manuscripts than those that eventually get accepted and go through other added-value work such as copy editing, composition, electronic tagging, online journal hosting, printing, distribution, and promoting the results

to various audiences. Federal agencies and publishers should make voluntary agreements to make the peer-reviewed articles available on the publisher site within 12 months after publication, or whatever time is appropriate for the publisher to sustain its business. The publisher should provide the federal agency the metadata and abstract for it to link to the final article. This will avoid unnecessary duplication of products or unpaid access to content to which publishers have invested and hold copyright.

(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?

There is no uniform optimal embargo period across all fields or for all types of publications appearing at various frequencies. Content in different disciplines has diverse patterns of usage, citation, and life spans. A 12-month embargo, but not shorter, is acceptable to AACR, with the articles held on the AACR journal site. We have considered what is needed to sustain our publishing program and to reinvest in the many activities the Association supports that contribute to the scientific endeavor. The NIH Public Access mandate requires NIH-funded, peer-reviewed accepted manuscripts to be deposited in PubMed Central and made freely available 12 months after publication. The final versions of these articles, along with related content, such as letters, commentaries, and retractions or corrections, are available on the sites of the individual AACR journals. In the time following the NIH Public Access mandate, AACR has seen a loss of some usage to articles on our websites. AACR makes editorial and business decisions based on usage information, and we cannot get sufficient usage information from PubMed Central to inform these decisions. Because libraries and other institutions base journal purchasing decisions on usage, housing journal articles on PubMed Central not only duplicates our efforts but also interferes with them.

AACR stands behind its voluntary decision to make all content freely available on our site 12 months after publication. An embargo of an earlier release would threaten our ability to sustain our publishing program and contribute to the Association's many activities that advance cancer research and the scientific endeavor.

Sincerely,

Diane Scott-Lichter
Publisher



Graduate Student Association
253 South Silo
One Shields Avenue
Davis, California 95616

January 11, 2012

Office of Science and Technology Policy

Re: Request for information related to public access to peer-reviewed scholarly publications resulting from federally funded research

To Whom It May Concern:

Thank you for the opportunity to submit public comment on this important topic that affects research, the public, and American economic interests. We are submitting our comments on behalf of the Graduate Student Association (GSA) at the University of California, Davis (UC Davis). The GSA represents the more than 4,300 graduate students at UC Davis in a variety of fields. UC Davis relies heavily on federal support, receiving more than \$437 million in federal government grants for 2009-2010. The graduate students at UC Davis are an important component of these federally funded projects, often serving as research assistants or even co-principal investigators. As a result, open access to research from federally funded projects would have a significant impact on students at the graduate level. At the same time, UC graduate students rely considerably on peer-reviewed publications to assist in their coursework, their research and ultimately, their thesis or dissertation.

The UC Davis GSA strongly supports the policy of open access to research, pursuant to the model of the National Institutes of Health. We believe that it is important for transparency and intellectual development that those outside of academia be given access to the research paid for by the federal government and their tax dollars. In particular, as we face considerable cuts to higher education within the University of California system, money for libraries and journal publications is being cut. At smaller schools, subscriptions to top-level research journals, which may cost thousands of dollars per year, are often the first casualties of budget reductions. . Providing open access to research can ensure that American students at universities and colleges have access to state-of-the-art research funded by the federal government. At the same time, it gives American businesses and the public the opportunity to better learn about the research our schools have been performing and be informed about the most recent advancements in medicine, science, and technology. Open access allows entrepreneurs to focus on innovation, workers to access the latest developments in their field and cancer patients to better understand their own care. We have a number of suggestions and opinions related to open access to research that we will detail below as we answer the prompted questions listed in the federal register.

1) Economic Considerations

We believe that open access can provide significant opportunity to spur intellectual capacity and economic prosperity. Allowing others to freely access the developments in medicine, science, and technology can enable other researchers to test results and improve upon existing research. We



Graduate Student Association
253 South Silo
One Shields Avenue
Davis, California 95616

support full access to publications through open access policies, with embargo periods, to ensure that our scientists and researchers can advance our economy through these policies.

2) Intellectual Property Interests

The purpose of intellectual property (IP) protection is to ensure that whoever makes a discovery or develops an invention has a unique opportunity to profit from their labor. Open access does not change this in any way. Publication is a gold standard for establishing primacy over IP, it clearly establishes original authorship and broadcasts primacy over a particular discovery to others in the same field. Open access does not fundamentally change this practice. The decision on whether to submit an article for publication rests with the author and sponsors and the decision on whether to accept an article for publication rests with journal editors. Open access only changes the channels through which the published article is disseminated to other members of the community

In fact, open access could possibly improve the ability of IP regulators to resolve contested claims. A common consideration in IP cases is whether a discovery substantially reproduces prior art or whether it is, in fact, novel. By increasing researchers' access to the latest scholarship, it may be easier to demonstrate that a discovery does not offer an advance over prior art. Additionally, by increasing access to research and transparency within government-funded research, Open Access may encourage researchers to more efficiently choose which projects to pursue.

5) Data

We recognize that publishing raw data is often not in the public or an individual researcher's interests. In fact, in many cases, this could jeopardize national security and competitiveness. As such, we support that data be released at least as much as what is within a peer-reviewed journal, and to the extent that Institutional Review Board and other ethics requirements permit.

6) Public Access

Open Access policies directly maximizes the benefit of science to U.S. taxpayers by making research available to a broad audience at no cost. These policies impose virtually no cost on scientists or researchers, since the overwhelming majority of time and expense involved in publishing stems from conducting the research. There are multiple existing online databases for scholarly journal articles that would almost certainly be willing to index articles published under open access provisions at little or no cost to the government.

In order to maximize the public's access to this research we would suggest that the government utilize a user-friendly website, with tutorials and help files written in plain English, as the default access point. We would suggest that all open access research be published in one database, as opposed to multiple domain-specific ones, in order to optimize presentation of cross-disciplinary articles. Finally, we would strongly suggest that all Open Access articles be indexed through the Google Scholar web service, as well as any specialized web sites.

7) Other Publications



Graduate Student Association
253 South Silo
One Shields Avenue
Davis, California 95616

The UC Davis GSA supports all non-confidential, federally funded research being made available through Open Access policies. We feel that government resources should be made available to all taxpayers and that the social benefit of such resources should be maximized. Other publication venues, such as book chapters and conference proceedings may have different routes to publication, but they are not inherently different from journal articles in that they represent a communication of scholarship to a broad community.

8) Embargo Periods

The UC Davis GSA feels that a six month to one year grace period is more than sufficient to ensure that private, for-profit distribution services can maintain their business model, while still allowing the public to reap a great benefit from this research. Scholars and researchers who are in highly competitive fields gain a distinct advantage from having immediate access to state-of-the-art research, and so they will continue to subscribe to journals or database services. This will allow the companies whose business is to provide such services access to the overwhelming majority of their revenue stream. The one year embargo period also allows for critical debate and editorial corrections to be made before it is released for wide publication.

Thank you kindly for the opportunity to submit comment on this issue. Please contact us if you have any questions or require additional information.

Sincerely,

Handwritten signature of Meredith Niles in black ink.

Meredith Niles

Deputy External Chair
Graduate Student Association
University of California, Davis
mtniles@ucdavis.edu
443-536-8390

Handwritten signature of Colin Murphy in black ink.

Colin Murphy

External Chair
Graduate Student Association
University of California, Davis
cwmurphy@ucdavis.edu

Name/Email: Abigail Goben,

Affiliation/Organization: I am writing this as a private citizen, though I am employed by the library at a publicly funded university which receives federal grants. Views herein expressed are my own as my employer has submitted their own response already.

City, State: Chicago, IL

Comment 1:

In response to comment one, I would suggest expanding the use of library and information scientists, in collaboration with federally funded scientific research, to facilitate access. Access will grow new markets and opportunities but it will help you have publications archived. Access should be publicly available, not behind any sort of paywall, and the articles need to have appropriate metadata applied to them. As yet, this is not something that computers can do for us. They're working on it, but you still need a person to oversee it. Benefits of this is people who can know the areas wherein they work.

Show citation box

Comment 2:

A standardized way to cite data and the encouragement of the federal government that data set creation and publication is valuable to the academic community must be created. Right now, scientists are so afraid of sharing their publications and their data because they have to hold onto things in order to get tenure and get further grants. There needs to be fiscal rewards for sharing and cross institutional collaboration. And a standardized way of citation would help to protect the IP of the scientists.

Federally funded research should be accessible to those who fund it, namely the public. If an embargo is developed, such as is presently in place with the NIH Open Access mandate, then it should be very short--no more than a year.

There should not be ANY policies that are put into place that place the rights of the public to federally funded research behind the rights of corporate entities (e.g. publishers).

Show citation box

Comment 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?

Pro of centralized: people know how to get there, you're not having to search way too many websites
Pro of centralized: Cross collaboration across types of science. This is very big in science and medicine right now. If researchers can see one place
Pro of decentralized: different places can take up different pieces and specialized. This should be done in conjunction and by funding local universities, specifically uni libraries and librarians, to develop these specialized digital libraries. It shouldn't be long term private sources--there are too many public opportunities to do this right.

Show citation box

Comment 4

NIH Open Access Mandate does nicely, though I don't think the publishers would agree with me. It certainly offers a lot of benefits to students and researchers at smaller institutions who can't pay the gateway access and numerous professionals who aren't on staff at a R1 institution.

Comment 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?

Provide a centralized location for them to deposit. Make it as easy and seamless as possible--these are people who already don't have time to do this, don't add mandates that are a) unfunded b) toothless and c) burdensome. Simple is ALWAYS going to be better
Tie future grants to compliance
Hire librarians to write the metadata for you rather than hiring scientists, we tend to be a more cost effective option

Show citation box

Comment 6

Make it simple
Make sure that it is funded as part of EVERY SINGLE GRANT. If a researcher is going to need metadata, then they need to BUDGET FOR IT NOW and the funders need to allow that budgeting or be willing to pay for it themselves. This wouldn't be a burden for libraries if it were funded. If we had the money to hire people to do this, we would jump on the opportunity. certainly I would. We are far too stretched financially though to take this on at present. Also a centralized server would be nice, unrealistic but nice.

Comment 7:

I don't see why not. E-books are where new things are being published. Let's do this right the first time. If there is concern about sniping by other researchers, put an embargo on this--2 years would probably do it.

Show citation box

Show citation box

Comment 8

One of the things I do as part of my job as an academic librarian is teach. Any mandate, ANY mandate, that embargos material more than 5 years will make that information instantly useless to my students. They refuse to go back further than that on a first search, often I have trouble getting them to look past 18 months. I doubt that they are a rarity and as they are our future scientists I imagine this will be pretty standard across the board unless someone is looking specifically for historical record.

Show citation box

Please identify any other items the Task Force might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.

--

Abigail Gobin, MLS

<http://HedgehogLibrarian.com>

We are witnessing improvement in access to scholarly publications through a variety of high-quality mostly not-for-profit initiatives related to metadata standards, archiving, linking, low-cost article renting, plagiarism mitigation, and information dissemination to the resource-compromised. These information-dissemination and preservation systems are catalysts of innovation acceleration in the sciences and the arts, and they will thrive without interference from the government. Free public access to reports of federally funded research is a reasonable expectation. However, the investments of publishers, whether operating under subscription-based or open-access models, should not be compromised by public-access mandates regarding the peer-reviewed literature because society gains scientifically, culturally, and economically as systems for providing added value are sustained. Government can help by focusing on communications infrastructure expansion on a scale that it may be uniquely positioned to provide.

Ted Bakamjian
Director, Publications
Society of Exploration Geophysicists

Michael Eisen
mbeisen@berkeley.edu
University of California, Berkeley
Howard Hughes Medical Institute
Public Library of Science

About me: I am an associate professor of Molecular and Cell Biology at the University of California, Berkeley and an Investigator of the Howard Hughes Medical Institute. I am also a co-founder and member of the board of directors of the Public Library of Science, a San Francisco based non-profit publisher of open access scientific journals.

I am responding to your request for information on how to best provide public access to publications describing the results of scientific and medical research funded by the US government. My views on this topic are begin with a simple principle:

If the taxpayers paid for it, the taxpayers own it.

In particular, when the taxpayers pay to generate information, by default that information should be in the public domain. The only exception should be if it is unambiguously in the public interest to restricting access in some way. For example:

- taxpayers pay to generate classified military intelligence, but do not have access to it in order to protect public safety
- taxpayers pay to collect information filed with income tax returns, but that information is not publicly available to protect individual's privacy

However, the scientific and medical research literature is different. Its value to the public is maximized when it is as widely available as possible. Research is a cumulative endeavor whose progress depends upon the flow of ideas, methods, data and discoveries. And there are countless people outside of the research community who benefit directly from access to the latest scientific and medical discoveries, including physicians and their patients, teachers and their students, and any member of the public interested in seeing where their tax dollars go.

Today, virtually all scientific and medical journals disseminate their results electronically (the number of journals that still print is dwindling, as is the number who read printed copies of these journals). The only reasons that the published results of all government funded scientific research is not freely available online at the moment they are published are 1) that most publishers of biomedical research journals cling to an economic model developed in the 17th century that depends upon granting access only to paying readers, and 2) that the government has been unwilling to defend the taxpayers fundamental right to access this material.

Prior to the internet, when information was primarily transmitted in printed form, the public good was advanced by an effective collaboration between the public research community, who generated information at taxpayer expense, and publishers, who curated, printed and distributed journals around the world. To facilitate the journals' role in the process, it became standard practice for researchers to assign their copyright in the work to the publishers. Because every copy of a printed journal costs money to print and distribute, it made sense for publishers to charge for every copy they delivered. Of course not everybody could have access to every journal whenever they wanted. But this was a limit of the technology available at the time.

But these limitations evaporated in the 1990s when the internet replaced printed journals as the primary mode of distribution for most scientific journals. There is now no technological obstacle to providing anyone with an internet connection access to the latest scientific and medical discoveries. And the fundamental economics changed as well. The costs of electronic publishing almost entirely come in production of the final version of the paper – there is essentially no marginal cost when a new reader accesses an article. Thus, while it once made economic sense to charge readers, it now no longer does. Unfortunately, publishers have largely failed to seize the opportunity provided by this technological revolution. Most publishers of research journals still charge individual readers and institutions to access their content, and deny access to all others.

This denial of access for most Americans to the results of biomedical research they paid for is completely unnecessary. In the past decade entrepreneurs in the US and elsewhere have seized on the remarkable advances in electronic communication to create a new business model that funds the value added by journals (overseeing peer review, editing and formatting of electronic documents) with requiring the massive subtracted value of denying millions of Americans access to something they paid for.

This new business model – known as open access publishing – treats the activities carried out by scientific publishers as a service provided to the research community and its funders. Open access journals are paid for the service they provide, and the peer-reviewed and edited papers they produce are made freely available from the moment of publication. Although initially dismissed as naïve, open access journals are now thriving. The Public Library of Science (PLOS) – a San Francisco based non-profit publishers that I co-founded along with Patrick Brown and Harold Varmus – now publishes the largest journal in the world, with over 14,000 papers published in the last year. And the organization is thriving financially. BioMed Central, a UK based for-profit open access publisher was acquired by a large commercial publisher (Springer) after it achieved profitability.

The success of open access provides an obvious means for the government to achieve universal taxpayer access to the research it funds. I would like to propose that the government make it a condition of the receipt of federal funds for scientific

research that any papers derived from their work be made immediately freely available through the National Library of Medicine. Researchers would have a wide choice of open access publishers to choose for publishing their papers, and existing non open access journals would have the opportunity to shift to the open access model or risk the loss of papers from federally funded scientists.

[I would like to emphasize at this point that I have no financial interest in the success of PLoS, which is a non-profit, or any other open access publisher].

Some publishers are sure to argue that this would constitute an intrusion of the government into the free market. But this is not the case. Publishers would, in fact, not be party to such a policy. In requiring that its grantees make the work available to the public, the government would be continuing a long tradition of imposing conditions on the receipt of federal funds. This would undeniably alter the marketplace for the provision of publishing services to the research community. But markets change all the time, and the free market will do here what it always does, and adapt to the new market conditions.

In considering such an action, I would point to several precedents in federal law. First, as you are undoubtedly aware, the Copyright Act of 1976 expressly precludes copyright protection to works produced by the federal government and its employees. Thus, for example, papers published by NIH intramural investigators are not subject to copyright, and are thus freely available to the public. Although the law did not apply this exemption from copyright to works funded by the federal government but not carried out by government employees, Congress carefully considered the matter, and wrote the following in the report language accompanying the bill:

Copyright Law Revision (House Report No. 94-1476)

A more difficult and far-reaching problem is whether the definition should be broadened to prohibit copyright in works prepared under U.S. Government contract or grant. As the bill is written, the Government agency concerned could determine in each case whether to allow an independent contractor or grantee to secure copyright in works prepared in whole or in part with the use of Government funds. The argument that has been made against allowing copyright in this situation is that the public should not be required to pay a "double subsidy," and that it is inconsistent to prohibit copyright in works by Government employees while permitting private copyrights in a growing body of works created by persons who are paid with Government funds. Those arguing in favor of potential copyright protection have stressed the importance of copyright as an incentive to creation and dissemination in this situation, and the basically different policy considerations applicable to works written by Government employees and those applicable to works prepared by

private organizations with the use of Federal funds....Where, under the particular circumstances, Congress or the agency involved finds that the need to have a work freely available outweighs the need of the private author to secure copyright, the problem can be dealt with by specific legislation, agency regulations, or contractual restrictions.

The scenario envisioned in 1976 – “the need to have a work freely available outweighs the need of the private author to secure copyright “ - surely applies now.

Finally, I find the following metaphor useful in thinking about the absurdities of the current publishing system: Consider the process of bringing a new baby into the world. Few would dispute that obstetricians play a significant role in the healthy delivery of a newborn baby. In exchange for their service they provide, they could demand ownership of the baby, and charge the parents a monthly fee to access their child. After all, the doctor “added value” to the baby by ensuring that the birthing process went well, and they deserve to be compensated for it.

Of course everybody recognizes this is absurd, because, while the doctor did do something of value, their contributions were trivial in comparison to those of the mother who carried the child for 9 months and did far, far more work during the actual delivery. But it is precisely this logic that leads publishers to assert the right to control permanently and restrict access to the primary record of publicly funded scientific and medical research.

Some additional references:

“Research Bought, Then Paid For”, an OpEd I wrote on this issue in the New York Times

<http://www.nytimes.com/2012/01/11/opinion/research-bought-then-paid-for.html>

“Why PLoS Became a Publisher”, an essay written by myself and the two co-founders of PLoS explaining open access.

<http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0000036>

January 12, 2012

To: Office of Science and Technology Policy Executive Office of the President
725 17th Street Room 5228
Washington, DC 2050

From: Nathaniel Hoffman

Re: Response to the White House RFI on OA publications

I am a videogame developer who makes frequent use of government-funded research in my work. I am also active in my professional organization, the Association for Computing Machinery (ACM), including volunteering for publication-related activities such as journal paper review and conference organization.

Overall, I largely agree with Harvard's position on this issue (<http://osc.hul.harvard.edu/stp-rfi-response-january-2012>). My strongly-held opinion is that every federal agency funding non-classified research should require immediate free online access to the full-text, peer-reviewed results of that research, without any time delay. The government should also provide the means (e.g. web-accessible database archive) to provide this access, as it does with PubMedCentral. Furthermore, these means should be made available to the copyright holders of any properly peer-reviewed non-government-funded research publication on an elective basis. This is the minimum that US taxpayers deserve - full and immediate access to the research they have paid for.

The US government should also encourage other government funding bodies (e.g. the European Union) to incorporate similar mandates, and should strive for mutual agreements to automatically make the publications from each government's archive available in the others, or possibly even set up a shared archive. To reduce inconvenience to researchers, ideally such agreements would also state that submitting the work to one of these archives counts as fulfilling the mandates for all the governments in the agreement.

Given that the vast majority of peer-reviewed research is funded by some government or other, my expectation is that even the few papers not covered by this mandate would eventually be submitted to these archives, making them complete repositories of all mankind's research. This would result from pressure by the authors of this research, who would wish to maximize the availability (and thus impact) of their papers. The end result would be universal open access.

It is hard to overstate the benefits of universal open access. The reduction in costs to educational institutions (who now spend huge amounts on journal subscriptions) would be, although large, one of the least significant benefits. Far more important would be the reduction in research friction - any researcher, practitioner, student or hobbyist could immediately access any research results. Even researchers in large institutions and industry practitioners in relatively deep-pocketed companies do not currently have access to all research, since there are many publishers and professional societies, each with their own paywalls and separate non-open archives. For example, I work at a very large game company and have access to three different paid archives, and I still regularly encounter papers I do not have access to. Since the additional cost of each such paper is not negligible, I then need to weigh carefully whether the value of the paper exceeds its cost - something that is often hard to determine without reading the paper in question! This introduces a huge amount of friction and limits my

productivity when doing research and development work. How much worse must it be for small companies, self-employed people, people working in third-world countries, etc.?

This problem is especially bad for anyone doing cross-disciplinary work (for example, my own specialty, computer graphics, involves elements of computer science, optics, electronics, and others) since each discipline typically has its own set of archives. Some publishers (like the ACM) allow authors to post "preprints" (trivially different than the official published version of the paper) on their own websites, and this is indeed preferable to not allowing such. However, this is not enough; although most authors do so (to maximize availability and impact of their papers), there are always a few who do not.

Since most technology-driven businesses are based in the USA, the US would benefit from universal open access more than other countries.

However, even if this were not the case it would still be highly beneficial to the USA to institute such policies. Research is not a "zero sum game" where every benefit to one party implies a corresponding loss to another. Improvements in technology will increase productivity and economic output worldwide, benefiting the USA as it benefits other countries.

There will be scientific publishers which will claim (most likely in response to this very RFI) that these open access mandates are unfair, that they will cause economic hardship and result in job loss. This claim should be ignored, for two main reasons. One is that through a series of historical accidents, scientific publishers found themselves in a position where they extract all the value from peer-reviewed scholarly publications while contributing a negligible amount to their creation. I know full well (from personal experience as well as that of many of my colleagues) that besides the research itself (which is typically funded by government bodies and for which in any case the publishers can take no credit whatsoever), all significant parts of the process of creating a peer-reviewed paper are performed by unpaid volunteers; researchers and members of the scholarly community performing a type of community service for largely altruistic reasons.

Paper reviews, organizing paper reviews, final decisions on acceptance, are all done by unpaid volunteers. There was a time when the scientific publishers would add some value in typesetting, printing and distribution. However now that electronic distribution via the Internet is the rule and most of the typesetting is done on computer by the paper authors (as can be seen by anyone comparing an "author preprint" to the final paper - the differences are negligible).

Open access mandates may or may not adversely affect the financial situation of the scientific publishers. However, this is immaterial.

The benefits of universal open access to all sectors of the US economy, to the advancement of science and technology, indeed to the betterment of all mankind far outweighs the profits of a small group of companies which once served a valuable purpose but do so no longer.

If they manage to find some significant value that they can add to the scientific process, they will survive and even flourish despite no longer being able to continue their current rent-seeking behavior. If they do not find some way to provide significant value, then their demise need be of no concern.

Some professional organizations (like my own, the ACM) also serve as scientific publishers, in addition to their other activities. These non-profit organizations were created, and continue to exist, only to advance the science and practice of a given field of human endeavor (computing, in the case of the ACM). Unfortunately, many of them (sadly, including the ACM) have management that has grown accustomed to the revenue streams attendant upon their publishing operations, to the point that they oppose open access despite its obvious benefits.

As a member, I strongly feel that this position - which values publishing revenue over the advancement of the computing field, which is the very purpose of the ACM's existence - is proof that the ACM's management has been sadly corrupted by reliance on these revenue streams. If these revenue streams disappear as a consequence of open access mandates, the ACM will be a better organization for it, more responsive to the desires of its members and the advancement of the computing field. I expect officers of the ACM to respond to this RFI with claims that open access mandates cause the ACM damage, and that if they were extended the damage could very well render the ACM unable to continue its various beneficial activities. I have two answers to those claims - that they are almost certainly untrue, and that even if they were true it would not matter. The claims are almost certainly untrue because some combination of other revenue sources would most likely be found to make up any shortfall in publishing income. The claims do not matter because all the good that the ACM (or any other professional society, or all of them put together) do pales in comparison to the benefits of universal open access to scholarly research.

Here are my answers to the specific questions asked in this RFI:

(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?

Answer: A full and immediate open access funding mandate by the US government, combined with encouragement of other governments to follow suit, will lead to universal open access and greatly increase the productivity of the scientific enterprise. The costs are negligible, and the benefits are immense. A model similar to PubMedCentral, but covering all federally-funded research and without the 12-month delay, is the best policy.

(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?

Answer: The intellectual property interests of publishers should not be a concern - their unique contributions to the scientific process (typesetting, printing and physical distribution) are now all irrelevant. The scientific publishing enterprise can continue very well without them. The intellectual property interests of scientists are served by maximizing the distribution (and thus the impact) of their research. The intellectual property interests of Federal agencies and other stakeholders are also served by maximizing the distribution and minimizing the friction of access to research. Policies which prioritize publisher profits over scientific advancement should be avoided.

(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?

Answer: The PubMed Central model is successful and proven. The US government should maintain its own archive with a copy of all research, easily accessible and searchable. This minimizes burdens on researchers, and maximizes long-term archiving stability as well as ease of use. However, there is no reason to prohibit additional, decentralized repositories. Many authors will probably continue storing a copy of their work on their own institutional web pages - the difference being that they will no longer be forced to do so in order to keep their work accessible, as they are today. Ideally, there would eventually be multiple government archives (one for the US, one for the European Union, etc.), each of which has all the papers present in the others as well (automatically, via exchange agreements).

(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?

Explicit public-private partnerships are not needed - the government can handle archiving. Existing publisher archives can definitely continue to exist alongside if the publishers wish to keep them, but they would no longer be the sole source of research results.

(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?

I recommend to continue building upon the successful PubMed Central model in these matters. A centralized government archive, with similar metadata to that used currently in PubMed Central.

(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?

Full and immediate open-access mandates maximally benefit all stakeholders except for the (now irrelevant) publishers.

(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?

Yes, definitely.

(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?

Zero - there should be no embargo period. Any embargo period is to the detriment of all stakeholders except the (now irrelevant) publishers.

Dear Science and Technology Policy Office,

Thank you for extending the deadline for comments on Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research. The Research Works Act has only very recently come to the notice of scientists, and it is because of this extraordinary proposal that it is now apparent to us that we need to reaffirm what we thought was settled: that OF COURSE scientific work funded by the public should be freely accessible to the public. I do not understand how this can even be a matter for discussion. The public pays: the public should benefit in every way possible.

The language in the RWA is highly misleading, attributing to publishers far more input into the scientific process than they really have. The truth is that scientists (often funded by public money provide the underlying research, the writing and the figure preparation that result in a manuscript submitted for publication. Other scientists then provide the editorial services and (contra publishers' claims, as can be easily verified) the peer review. Publishers' contributions are limited essentially to typesetting, the provision of web hosting, and sometimes a very limited amount of compensation for senior editors only (usually not the handling editors who actually deal with authors' works). The notion that such a minor contribution should suffice to hand publishers, rather than the public, the right to determine how, where and under what regime the resulting works are disseminated, is ludicrous. It would be laughable if it were not so iniquitous.

For instance studies show hundreds of people die if their physicians don't know about the latest literature. Remember these physicians cannot read the literature which scientists give the publishers. So misdiagnosis is common, and avoidable by access to the literature.

The process of publishing knowledge in closed journals is not the most beneficial way. Why do we still predominantly rely on these outdated mechanisms in the age of crowdsourcing, innovation in rating systems, social media and Google. Especially if we know that closed access means that people die.

Andrew E Hospador

Retired Scientist, Chief Engineer, et. al.

MEE, BSEE, Cornell, '62

From: Patrick Brown

Sent: Friday, January 13, 2012 12:42 PM

Subject: Re: Response to RFI on public access to scientific research

The federal government has a responsibility to the taxpayers who are paying for the federally funded research and on whose behalf it is being conducted to ensure that they receive the greatest possible return on their investment and that they have unrestricted freedom to read, use, evaluate, criticize and benefit from the results of that research.

Both of those responsibilities are best served by removing all barriers to access or use of the formal published record of the research.

Before the internet, those barriers were a necessary evil. The most efficient way to maximize access to research results was to print and distribute research reports in peer-reviewed journals, each copy of which cost money to print and distribute, necessitating limits on distribution and charges for each copy. The traditional scientific publishers (generally) did this well.

With digital publication on the internet, the costs of publication are completely independent of the number of copies distributed or the number of readers who have access. Restricting access and distribution to preserve a pay-per-copy business model that has been rendered obsolete by the internet makes no economic sense.

Anything short of immediate and complete open access to all published results of non-classified, publicly-funded research results would be cynically putting the narrow interests of a scientific publishing industry dominated by a few multinational megacorporations ahead of the vitality of the scientific research endeavor and the benefits it brings to public health, quality of life, job creation and economic growth. Any policy that in any way favors delayed or restricted access over immediate open access to scientific research would sacrifice jobs that would otherwise be created by technology- and research-intensive businesses whose competitiveness depends on access to the latest research results. Any policy that in any way limits any practicing physician's access to the most up-to-the-minute published results of the research that the citizens of the United States have funded would be a betrayal of their trust.

No self-serving argument put forward by the scientific publishing industry (of which I am a part) can possibly outweigh our fundamental responsibilities to the citizens who fund our research. The "old school" scientific publishers fear open-access publishing as a disruptive threat to their huge profit margins, but open-access publishing is a thriving business and may well have been the greatest source of growth and job creation in the scientific publishing industry in the past few years.

I am a full-time research scientist at Stanford University and a co-founder of Public Library of Science, a non-profit open-access publisher of scientific research. I have benefitted greatly from the generosity of the citizens of the United States who have had the extraordinary vision and altruism to support the greatest scientific research system in the world. It is my privilege and my obligation to make any results of my research unconditionally available for the benefit of the people who made it possible.

AIChE

December 29, 2011

Dear Sir/Madam:

The American Institute of Chemical Engineers (AIChE) is pleased to respond to OSTP's November 3, 2011 Federal Register notice requesting comments on "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research." We are gratified by the administration's consultation with stakeholders in the engineering community.

Founded in 1908, AIChE is a professional society of more than 40,000 chemical engineers in 92 countries, with approximately 90 percent of these members based in the U.S. Our members work in corporations, universities, and government using their knowledge of chemical processes to develop safe and useful products for the benefit of society.

Through its publications, meetings, research efforts and other programs, AIChE is a focal point for information exchange at the frontiers of chemical engineering research in such areas as nanotechnology, sustainability, new energy sources, biological and environmental engineering, and chemical plant safety and security. Our publications program includes four peer-reviewed journals, a monthly member magazine, books, databases, and electronic newsletters. Like many other scientific and engineering societies, we depend on the revenues generated by these publications to support activities that benefit practicing chemical engineers, researchers, students, and society.

While we certainly agree that taxpayers should benefit from the results of federally-funded research, learned societies like ours add significant value to the research papers submitted to our journals through rigorous peer review, quality control, and other activities in the publication process. Taxpayers may fund the research, but the societies and technical publishers fund the publication of this research. What we copyright and own is the value-added article, which is our work product. Expropriation of these value-added articles by the government without compensation would undermine copyright, intellectual property rights, jobs and exports, as well as the societies' ability to carry out important programs that benefit the nation.

Depriving the learned societies of revenue from their publications will reduce our resources for—and efforts to—grow and strengthen the science, technology, engineering and mathematics (STEM) workforce that the nation so desperately needs to meet important challenges and remain competitive. In fact, we have been increasing the resources we devote to preparing the next generation of chemical and biological engineers. Decreasing our ability to support the future STEM workforce seems to fly in the face of stated national priorities.

Additionally, there is no evidence that making access to the journal literature free will improve research productivity or the public weal. On the contrary, free access, like copyright piracy, may well have the opposite effect. Making societies' content freely available is likely to stifle opportunities as customers choose to access

free versions of journal articles rather than pay for the Version of Record. Just how will U.S. taxpayers benefit from the transfer of intellectual property owned by U.S. not-for-profit societies to every researcher, business and government abroad?

We strongly believe that the Federal Government, currently under such significant financial stress, **not** mandate the deposit of journal manuscripts in what will likely be a very expensive, if freely available, archive, regardless of format, process, or timing. Rather, the Federal Government should strive to provide public access to the information that it already controls and has a right to distribute — for example, research summary reports regularly received from grantees.

Learned societies and journal publishers have been good stewards of the literature. Indeed, since journal content delivery via the Web began in the mid-1990s, we have invested not only in the technology necessary to deliver increasing amounts of content via the Web, we have also worked to digitally recover existing print material, in many cases back to the first issue of a title. The online availability of so much content has accelerated and broadened availability of the peer-reviewed literature. Most academics and researchers access the necessary literature on their desktops via subscriptions or licenses maintained by their institutions.

Mandating a single approach to public access will stifle innovation in what is a rapidly changing environment, as engineering researchers explore new ways and new media to enhance and communicate their findings. The proposed action may well reduce the development of new tools, delivery vehicles and functionality to advance the future of engineering and the many enterprises our members serve.

We feel that we must ask if the government is—and should be—a credible provider of the kinds of publication services that not-for-profits societies and commercial publishers have created. Given overwhelming budget constraints, why would the government consider using taxpayer dollars to duplicate existing, well-functioning services?

Take PubMed Central, the repository for mandated NIH grantees, as an example. It is not a simple archive of articles but a sophisticated publishing platform requiring millions of dollars of investment. Have the full costs of similar repositories been developed in any consideration of an expansion of the NIH mandate? Will each funding agency develop its own processes and create its own platform to serve its special needs?

Rather than imposing an unfunded mandate on societies and publishers and taking on such very substantial costs, the Federal Government could:

- Make funds available for purchase of open access to published articles. (Several research funders already do this.) These costs are a small fraction of the investment in the research itself.
- License content from learned societies and make it available to specific audiences.
- Make the agency-collected and maintained output of taxpayer-funded research, including grant reports or research progress reports, freely available to the public. Work with learned societies and private sector publishers to make that content findable and link it to the journal literature.

What the Federal government should **not** do is take accepted or published articles from learned societies directly or through new mandates placed on grantees.

Most researchers acknowledge funder support they have received in their journal articles. Publishers are working to develop a means of standardizing funder information so it could be made easily available to funders. We believe that such community-wide solutions will be simpler and far less expensive to construct. Experience shows that publishers and the learned societies, and the partnerships they have created through

DOI and CrossRef are well positioned to ensure that journal articles are accompanied by standardized, high-quality metadata. Extending this approach to provide information about the agency, program, and even specific grants that funded the research seems like a straight-forward solution. Visitors will be able to follow links (enabled through the DOI) to and from the agency's or the publisher's platform, finding research reports, article abstracts that are freely available and the peer-reviewed, quality-controlled Version of Record.

Perhaps research funders would better serve the community by also partnering with publishers and learned societies to provide access to raw research data and to enable content mining that can drive interdisciplinary research and support the identification of new areas of discovery.

Unless research funders provide the resources to the societies and other publishers, we believe that these learned societies and publishers should continue to manage access to peer-reviewed papers within the duration of copyright. For accepted author manuscripts and published journal articles, in which we have made substantial investments, learned societies should be free to determine the business models under which their publications operate, including the time, if any, at which the final peer-reviewed manuscript or final published article are made publicly available.

Thank you again for this opportunity to comment.

Sincerely,



June C. Wispelwey
Executive Director
AIChE

Mon 1/2/2012 7:20 PM

Public Access to Federally Funded Research

I am an ordinary US citizen, 67 years old, who is most DEFINITELY in favor of public access to ALL federally funded research in ALL areas of endeavor. I also think it should cost very little to make such access available digitally at this point in time since most people now have computers and access to the internet. I think had we had access to federally funded medical research, in particular, in past years, there would not have been so many reports after the fact of drugs and medical procedures that have harmed people who used them.

Access to information, PUBLIC access made available to anyone desiring it with enough time and interest to read it, is one of the hallmarks of this republic, and most certainly should be allowed as much as possible when the means is available to make it accessible to as many as possible.

I doubt these comments by an ordinary citizen will make one fig's worth of difference to whoever is collecting these "comments," but I'm sending them anyway.

Gail Kearns

**Before the
Office of Science and Technology Policy
Executive Office of the President
Washington, D.C. 20502**

*In the Matter of Request for Information: Public Access to Peer-Reviewed Scholarly
Publications Resulting From Federally Funded Research*

Comments of Wolters Kluwer Health

Wolters Kluwer Health is pleased to respond to the Request for Information (“RFI”) from the Office of Science and Technology Policy on approaches for ensuring long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally-funded scientific research.¹ Wolters Kluwer Health believes that recommendations on such approaches must be discussed in the context of the Administration’s goals on economic growth and innovation – the strategy for which identifies health information technology as a national priority, but also notes that intellectual property enforcement is critical for continued innovation. Therefore, in addition to commenting on the questions raised in the RFI, Wolters Kluwer Health will provide herein information on some of its activities in healthcare IT innovation.

Introduction

Wolters Kluwer Health (“WKH”), a medical and health care publisher, is a leading global provider of medical information, workflow solutions, and platforms for research and development, as well as business intelligence tools. As a publisher, WKH is dedicated to broad access to peer-reviewed scholarly publications. WKH’s mandate includes providing access not just to the publications themselves, but to an array of medical information, tools, and solutions. These are used by healthcare professionals and organizations worldwide, improving clinical practice, and raising access to quality and cost-effective healthcare.

¹ See *Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research*, Notice of Request for Information, 76 Fed. Reg. 68518-01 (2011).

Investment in On-Line Tools to Improve Medical Research

WKH has invested substantially in recent years in new productivity workflow tools and access options and expanded its clinical content offerings to improve medical research.

The company recently introduced workflow improvements to its flagship research platform, OvidSP. These enhancements focused on improving the researcher's productivity with the My Project and the Ovid Toolbar, each of which helps users effectively manage their search strategies and research projects within a single, integrated platform. WKH has helped researchers not just in the United States, but in other countries as well, through the launch of local language search portals for use in many emerging markets including those in Asia and the Middle East. For instance, WKH has partnered with the Egyptian Ministry of Higher Education and State Ministry for Scientific Research to collaborate on the development and publication of leading medical and scientific journals and to help create a 'Center of Excellence' in the region.

WKH, through its Lippincott Williams & Wilkins ("LWW") business, has continued to enhance its e-journal platform with new features and functionality for journal readers and editors, including new email alerting features, "publish ahead of print" content alerts, topical collection alerts, as well as personalization features.

Investment in On-Line Tools for Education and Service Delivery for Healthcare Professionals

WKH is dedicated to expediting the results of medical research into the practice of healthcare. Through its investment in digital and human resources, the company can now, on average, translate medical research into practice guidelines for use by physicians, nurses and physician assistants within 10 weeks, in almost 20 practice areas, with almost 9,000 topics, by over 5,000 authors and editors. WKH has invested significant resources in expediting research results to physicians' point-of-care. The company is dedicated to ensuring research does not sit unused in an archive, but is applied in the delivery of care giver training and patient services. It should be noted that in the past it could take years for clinical research to impact actual patient care.

As the Chief Executive Officer of Wolters Kluwer N.V., parent company of WKH, has recognized, by 2013, more people will access the internet on mobile devices than from desktops.² Wolters Kluwer N.V. has launched several initiatives across the company to take full advantage of this trend, including the development by WKH of mobile apps for care givers at the point-of-care. WKH has devoted significant time and effort to reshaping and transferring our products to mobile devices that reach our customers at their point-of-use.

² NANCY MCKINSTRY, 2010 ANNUAL REPORT: MESSAGE FROM THE CEO (2010), *avail. at* <http://reports.wolterskluwer.com/2010/ar/messagefromthecoo.html?cat=m>.

WKH is proud to be a leader in the development of mobile apps for healthcare professionals. WKH has launched 500 education titles in multiple e-book formats, including iPad® apps. Wolters Kluwer Mobility Lab supports the mobile development across the organization by focusing on the interface to multiple mobile operating systems, including for e-books, mobile apps, portable interfaces, and other workflow-centric online solutions. Going forward, WKH is looking at the next-generation of mobility products that are compatible with Blackberry, iOS, and Android systems and concepts based upon the professional user's workflow.

LWW – a Wolters Kluwer Health business - was the first medical publisher to introduce the ePub export format for retrieving articles for devices including the iPad®, Barnes & Noble Nook, and Sony Reader. WKH was one of the first biomedical publishers to introduce an iPhone® app, Journals@LWW, for mobile access to LWW published journal content. The app recommends LWW titles based on specialty preferences and pushes article content based on user selections. WKH is now introducing iPad® apps for its leading journals, which offer not just the text in an easily accessible format, but also include new multimedia content to enhance the communication of clinical information. As a partner for some of the country's leading medical associations, WKH is delivering real benefit to physicians and nurses in their daily practice.

WKH has invested in other on-line tools for healthcare professionals. The company developed an interactive testing platform for nursing students designed to help students learn and prepare for their national exam. This adaptive quizzing product delivers personal practice tests and allows students to track individual or class performance to optimize results. WKH's on-line tools also offer educators real time analysis at the class and individual student levels. Given the importance of electronic medical records ("EMR"), WKH has also invested in EMR educational software. One of WKH's products integrates a simulated electronic medical record into a learning tool for students to prepare future nurses for the demands of the evolving healthcare climate in the 21st century.

WKH has also developed tools for members of the general public who may be interested in particular healthcare developments in specialized fields. For instance, WKH has made its *Heart Insight* and *Urology* journals, published on behalf of its clients the American Heart Association and American Urological Association respectively, accessible on mobile devices by the general public.

Investment in mobile content and apps and other digital tools to improve point-of-care services requires that WKH earn a return on its investment in journal publishing. WKH would not be able to innovate and develop mobile tools for physicians to access up-to-date results from clinical trials, training programs, and other apps if it does not have sufficient revenue from journal subscriptions.

Recommendations on Approaches to Stewardship and Public Access

On behalf of the National Science and Technology Council’s Task Force on Public Access to Scholarly Publications, the RFI requests information and comment on eight questions regarding approaches for long-term stewardship and public access to the peer-reviewed scholarly publications that result from federally-funded scientific research. Wolters Kluwer Health will comment on several issues raised in the questions.

Embargo Periods

As raised in Questions 2 and 8, regarding embargo periods, it is not good policy for the United States Government to expropriate intellectual property of private entities from which it seeks innovation. The government should ensure that there is an embargo period prior to allowing users around the globe free, unpaid access to content for which publishers hold a copyright. Mandated free access to copyrighted material is inconsistent with the Administration’s goals of sustainable economic growth and technology innovation. Twelve months is not a sufficient amount of time for some publications to recoup the investment in scholarly manuscripts for which copyright is transferred from the author to the publisher. In no instance should an embargo period be shortened from twelve months, as some free access proponents have argued.

President Obama recognized in his Innovation Strategy that protecting intellectual property rights is critical to encouraging innovation.³ This is equally as important for scholarly publications as for other types of copyrighted content. Publishers invest significant editorial and administrative resources organizing peer-review of the many manuscripts they receive, and then invest additional resources after a manuscript is accepted for publication and the copyright is transferred from the author to the publisher. Editors, highly educated and knowledgeable experts in their discipline, bring value in their very selection of manuscripts to assign to peer-reviewers. That selection is intended to contribute to and advance the state of knowledge for that particular discipline. The final scholarly publication is the result of substantial investment by the publisher. At no point is it appropriate for the U.S. government to mandate that the public – including many overseas institutions for which the publication has economic value – should have free, non-paid access to the peer-reviewed manuscript accepted for publishing or the published article.

³ See EXECUTIVE OFFICE OF THE PRESIDENT, NATIONAL ECONOMIC COUNCIL, OFFICE OF SCIENCE AND TECHNOLOGY POLICY, A STRATEGY FOR AMERICAN INNOVATION: DRIVING TOWARDS SUSTAINABLE GROWTH AND QUALITY JOBS 15 (2009), *avail. at* http://www.whitehouse.gov/assets/documents/SEPT_20__Innovation_Whitepaper_FINAL.pdf (“Protect intellectual property rights. Intellectual property is to the digital age what physical goods were to the industrial age. We must ensure that intellectual property is protected in foreign markets”); *see also* DEPARTMENT OF COMMERCE, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION, INTERNET POLICY TASK FORCE, COPYRIGHT (2010), *avail. at* <http://www.ntia.doc.gov/category/copyright> (“The Department of Commerce’s Internet Policy Task Force is conducting a comprehensive review of the relationship between the availability and protection of online copyrighted works and innovation in the Internet economy”).

Patient Access

For most biomedical publishers, patients – the primary beneficiaries - are not the primary market. While WKH has in fact developed tools for patients and other members of the general public to access scholarly publications on SmartPhones and tablets, its primary market is not patients. WKH is a supporter of *patientINFORM*, through which it and other publishers fund the development of free access to on-line material for patients and support links to our scientific literature of broad public interest. *PatientINFORM* is a collaborative effort by patient health organizations, medical societies, health information professionals, and scholarly and medical publishers like WKH.⁴ WKH clients the American Heart Association, the American Society for Nephrology, and the American Cancer Society, as well as other voluntary health organizations create on-line material for patients and their caregivers describing research developments and providing links to the full text of research articles selected from participating journals.⁵ The publishers allow readers following links from *patientINFORM* material located on the health organizations' sites to access the full text of these articles without a subscription, and these publishers provide patients and caregivers with free or reduced-fee access to other articles in participating journals.⁶

Relative to Question 1, on what policies federal agencies can take to grow existing markets related to the access and analysis of peer-reviewed publications that result from federally-funded scientific research, funding of *patientINFORM* from federal agencies such as the U.S. Department of Health and Human Services' Agency for Healthcare Research and Quality would help ensure that higher levels of content would be available to the public on *patientINFORM*. Federal funding of *patientINFORM* would realize a form of private-public partnership to encourage accessibility to peer-reviewed publications, envisioned in Questions 4 and 6.

Innovation Strategy

As noted above, President Obama has highlighted the importance of innovation to continued U.S. economic growth and competitiveness. WKH shares the goal of “harnessing the power of data and technology” to create “innovations in health care delivery”.⁷ As WKH's CEO has noted, “innovation for Wolters Kluwer Health is our history, our vision, and our future.”⁸ WKH drives innovation through continual research and development, unique industry partnerships and collaborations, strategic acquisitions, and investment in cutting-edge technologies and solutions. Innovation is the lifeblood of what we do to help our customers make healthcare better.

⁴ See *PATIENTINFORM, PARTICIPATING ORGANIZATIONS* (2010), *avail. at* <http://www.patientinform.org/participating-organizations/>.

⁵ *See id.*

⁶ *Id.*

⁷ THE WHITE HOUSE, *INNOVATION* (2011), *avail. at* <http://www.whitehouse.gov/economy/business/innovation> (“*The White House on Innovation*”).

⁸ See *WOLTERS KLUWER, 2010 ANNUAL REPORT: REPORT OF THE EXECUTIVE BOARD* (2010), *avail. at* <http://reports.wolterskluwer.com/2010/ar/servicepages/search.php?q=health&pageID=13078&cat=b>.

WKH is therefore committed to technology adoption, as a means to improve healthcare. With an aging population and a shortage of healthcare professionals, it is critical that technology be adopted that can help improve productivity of clinical service delivery. Technology can help caregivers reduce medical errors, eliminate waste, and improve patient outcomes. WKH is focused on extending its current products to new markets, as well as driving future content, tools, and software expansion, all of which will enhance workflow solutions. Mandated free, on-line global access to our copyrighted material reduces the very revenue that makes such innovation possible.

As the President has noted, intellectual property rights, including copyright, are critical for continued innovation and growth of the Internet economy.⁹ To ensure that publishers such as WKH are able to continue to provide innovation in health care delivery, government policies must protect copyright, both at home and abroad. Recently, the National Institutes of Health (“NIH”) acknowledged that over two-thirds of users accessing free, on-line journals from PubMedCentral are foreign.¹⁰ Policies that mandate free access to copyrighted biomedical content will undermine innovation, economic growth, and patient care.

Open Government

WKH supports the Administration’s goals of making government open and government data more accessible. It therefore supports the policies of the America COMPETES Act of making Project Outcome Reports on federally-funded research available on-line for the general public. Under America COMPETES, Congress directed the National Science Foundation (“NSF”) to ensure that “all final project reports and citations of published research documents resulting from research funded in whole, or in part, by the Foundation, are made available to the public in a timely manner and in electronic form through the Foundation’s Website.”¹¹ Currently, just NSF is required to make Project Outcome Reports of NSF-funded research available on-line. Relative to Question 1, WKH supports making project outcome reports on research resulting from all federal-agency funding, including NSF, NIH and others, available for free on-line public access. In addition to meeting the Administration’s goal of enhanced public access to the results of federally-funded research, such a policy would provide researchers both the positive and negative results of research, and the results of applied - not just basic - research. Importantly, public access to project outcome reports, a form of results of federally-funded research, would not undermine the copyright of publishers and their ability to innovate and compete in the global marketplace.

⁹ See The White House on Innovation (“We are applying these Internet innovation principles in our work on privacy, cyber security and copyright, to preserve the Internet as a source of growth and greater possibility for America and the World”).

¹⁰ See Letter from Susan R. Cornell, J.D., FOIA Officer, the National Institutes of Health, to Allan Adler, Vice President for Legal and Government Affairs, Association of American Publishers at 8 (May 19, 2011) (enclosure).

¹¹ America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (America COMPETES Act), Pub. L. No. 110-69, § 7010, 121 Stat. 572 (2007).

Conclusion

The President has stated that “We should be helping American companies compete and sell their products all over the world. We should be making it easier and faster to turn new ideas into new jobs and new businesses.”¹² A policy that mandates that copyrighted material should be made available for free on-line – including to foreign entities that would otherwise pay for that content – does not help companies like Wolters Kluwer Health compete and sell our products all over the world. In fact, the NIH policy facilitates global piracy. Wolters Kluwer Health and other publishers participate in a number of programs designed to enhance public access to the results of federally-funded research, including *patientINFORM*. Wolters Kluwer Health would welcome government funding of *patientINFORM* to expand the range of content accessible to patients and their caregivers on the site, such as lay translations of biomedical articles, in addition to providing summaries of research articles.

Wolters Kluwer Health also supports the America COMPETES model of making project outcome reports on federally-funded research available on-line for free public access. That model, expanded to NIH and other funding agencies, would not undermine copyright or innovation in healthcare information technology and would expand information to researchers to include applied research and negative research results.

Wolters Kluwer Health also recognizes that the government could provide an important role in data curation, including in standards-setting, since publishers engage in publishing and not curation. In an era of constrained federal budgets, and in light of the innovation the private sector has delivered in point-of-care content delivery, it is more efficient for the federal government to focus on data curation. Federal government leadership in curation standards, along with free public access to project outcome reports on the results of federally-funded research, would help the long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally-funded scientific research, without undermining the private sector’s ability to continue to invest in innovative ways to enhance access to such research.

Respectfully submitted,

/s/

Karen Abramson
President and Chief Executive Officer
Wolters Kluwer Health Medical Research

¹² BARACK OBAMA, THE WHITE HOUSE, TECHNOLOGY (SEPT. 16, 2011), *avail. at* <http://www.whitehouse.gov/issues/technology>.

333 Seventh Avenue
20th Floor
New York, NY 10001

January 11, 2012



Ecological Society of America
1990 M St, NW, Suite 700
Washington, DC 20036

December 20, 2011

Re: FR Doc. 2011-28623

Dear Madam or Sir:

As the world's largest organization of 10,000 professional ecological scientists, the Ecological Society of America (ESA) would like to provide feedback in regard to the Office of Science and Technology Policy's Request for Information on public access to scholarly publications. The Society has published ecological research in journals widely available to the public in libraries and universities for over 90 years. However, it is important to note that there is a significant difference between research results and peer-reviewed publications. Publishers such as ESA have a long record of reporting, analyzing and interpreting federally funded research. It is not appropriate for the federal government to expropriate the additional value publishers add to research results. A better course of action would be to allow publishers to continue to experiment with the best ways they can broadly disseminate materials that analyze and interpret research and to encourage federal agencies to make content they already own more visible and easier to understand.

ESA publishes four of the world's most highly cited journals in ecology and environmental science. Subscription revenue from these journals is crucial to ESA's publishing program. Without it, the Society could not continue to provide the peer-review and editorial services needed to produce high-quality scientific publications. Furthermore, subscription revenue helps to support other Society services, including scientific conferences, education programs, and the distribution of science information resources to policymakers and the public. The Society also publishes one open access journal under the author pays model. ESA's public access content includes:

- The *Bulletin of the Ecological Society of America*
- *Issues in Ecology*, a publication series presenting the scientific consensus on prominent environmental issues in language accessible to nonscientists
- A featured article in each issue of the four peer-reviewed subscription journals that ESA publishes
- All special issues of *Frontiers in Ecology and the Environment*, *Ecology*, and *Ecological Applications*
- The "Reports" section of *Ecology* and the "Communications" section of *Ecological Applications*, both of which contain concise papers on groundbreaking research
- *Ecological Archives*, which contains all appendices and supplemental material associated with papers published in the journals, including data sets, methodological and analytical detail, and computer code
- All ESA journal abstracts
- *Ecosphere*, a rapid-publication, online only, author pays, open access journal

In addition, ESA freely grants authors permission to post papers on their personal or home institution's websites. The Society also permits liberal use of ESA publications for educational purposes.

ESA continues to build its base of open content materials, but this evolving model must be allowed to develop with due care and should be orchestrated by the individual publishers. Moreover, different fields of science have different citation lives. Ecological research often examines changes that occur over long spans of time; findings frequently have a citation half-life of more than a decade. Papers published in ESA journals may therefore be just as relevant in several years as they are today, which means that any potential embargo period will do little to mitigate the financial losses that would result from full open access. Furthermore, journals in fields such as medicine and genetics, garner much of their revenue from advertising, whereas journals in other fields, such as ecology, must rely more heavily on subscriptions.

One way to make taxpayer funded research more visible and accessible to interested members of the public would be to require federally-funded grantees to provide a second version of the research summaries they already prepare, specifically for the lay reader. To aid in online searches, these summaries could also include the source of federal funding institutions and grant numbers. Publishers could also include grant information in paper abstracts which are usually available without a subscription.

Government mandates for publishers to make their work available online without compensation will endanger the U.S. scholarly publishing system. ESA respectfully requests that the Administration allow the scientific publishing community to continue to explore workable solutions that meet the dual goals of the scientific enterprise as well as provide resources to interested members of the public.

Thank you for your consideration,

A handwritten signature in dark ink, reading "Katherine S. McCarter" with a long horizontal flourish extending to the right.

Katherine S. McCarter
Executive Director and Publisher

Re: FR Doc. 2011-28623

American Mathematical Society Response to OSTP Request for Information:
Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded
Research

Prepared By:

Dr. Donald E. McClure
Executive Director
American Mathematical Society (AMS)
Providence, RI/ Ann Arbor, MI/ Washington, DC
donald.mcclure@ams.org

To:

publicaccess@ostp.gov
Office of Science and Technology Policy (OSTP)

January 2012

Dear Sir or Madam:

The American Mathematical Society (AMS) is a scholarly society whose primary mission is to foster research and scholarship in the mathematical sciences. We are a membership organization with over 30,000 members worldwide and over 23,000 members in the United States. Most of the members are engaged in research in the mathematical sciences and the majority of the members are affiliated with higher education institutions. We are a scholarly publisher of leading research journals, books (monographs and advanced texts) and a premier reviewing and indexing publication *Mathematical Reviews*, available worldwide on the internet as *MathSciNet*.

The AMS has a total staff of 210 located in Providence, Ann Arbor and Washington, DC. The majority of the employees are engaged in scholarly publishing.

The AMS is affiliated with the Professional/Scholarly Publishing Division (PSP) of the Association of American Publishers (AAP) and with the Association of Learned and Professional Society Publishers (ALPSP). Both of these organizations have responded to the OSTP RFI regarding public access and we endorse their responses.

I am writing now to provide some additional remarks pertaining to the mathematical sciences and how the use of the research literature in mathematics may differ from the literature in other disciplines. The numbers that I cite are based on (1) the *Mathematical Reviews* database

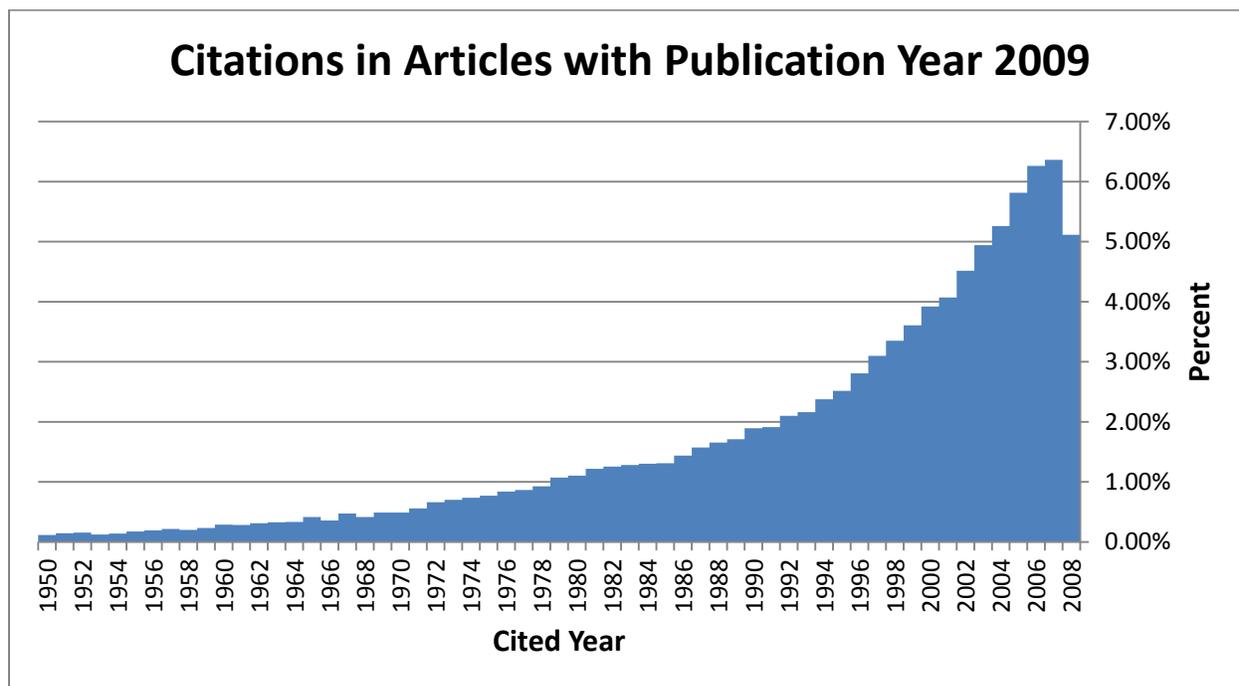
(MRDB) and (2) the AMS Publications Database (AMSPDB). Data from MRDB were extracted on January 25, 2011 and data from AMSPDB were extracted on February 2, 2011.

I. Question 8 in the OSTP RFI asks for input on the appropriate embargo period after publication before the public is granted free access to the full content of peer-reviewed scholarly publications.

The AMS makes all of the content of its journals freely available five years after publication. We believe that five years is an appropriate embargo period for the mathematical literature. There are two primary reasons:

1. Most of the literature is available much earlier anyway in public repositories or on authors' personal web pages; and
2. The "half-life" of a mathematics research paper is about eleven years.

The graph below shows a histogram of the *age distribution of cited articles* for citations made in mathematical sciences research articles published in 2009. The source is MRDB, which reported on 78,000 articles published in refereed journals in 2009.



For journal articles published in 2009, 50% of the citations contained in those articles were to papers published in 1998 or earlier. The usefulness and relevance of a mathematical sciences paper does not diminish appreciably over time.

Mathematicians customarily make their final manuscripts available as preprints. This is a tradition that goes back to at least the 1950s when it became easy to reproduce a research report. Today, many authors post preprints on ArXiv.org. To illustrate this point, of the articles published in *Journal of the American Mathematical Society* in 2009, 71% are available at ArXiv.org; of the articles published in *Transactions of the American Mathematical Society* in 2009, 57% are available at ArXiv.org.

II. In the mathematical sciences, many authors of research articles are not funded by a Federal Agency. The AMS has adopted a publication policy, the main point of which is to assure that all authors have the opportunity to publish in AMS journals regardless of their financial circumstances. The policy states:

“The American Mathematical Society strongly endorses and adheres to the principle that a paper in the mathematical sciences should have an opportunity to be evaluated and properly published without regard to the financial circumstances of its authors.”

A journal published under the Gold Open Access model might be regarded as discriminatory to unfunded authors under this policy.

In 2009, approximately 84% of the articles published in the AMS’s primary research journals had no U.S. Federal support. There are two reasons for the low percentage of Federally-supported articles: (1) scholarly research in the mathematical sciences is an international enterprise and the majority of authors are domiciled outside the U.S., and (2) many U.S. authors are not funded.

Dr. Donald E. McClure
Executive Director
Providence, RI

RSC Publishing

Response to OSTP Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research

On behalf of the Royal Society of Chemistry, UK

To: Office of Science and Technology Policy (OSTP)
Washington, DC 20502, USA

via e-mail to: digitaldata@ostp.gov

From:

James Milne PhD
Editorial Director & Acting Managing Director
RSC Publishing
Royal Society of Chemistry
Thomas Graham House
Science Park, Milton Road
Cambridge, CB4 0WF, UK

About RSC Publishing

RSC Publishing is one of the largest and most dynamic publishers of chemical science information in the world. We publish 34 international peer reviewed scholarly journals, almost 100 scientific books per annum, two highly acclaimed magazines, and a number of successful databases.

Not-for-profit

We are a not-for-profit publisher wholly owned by the Royal Society of Chemistry. Our authors, readers and customers are truly international and our publishing activity dates back to 1841.

Authoritative

RSC Publishing is a member of ALPSP, the Association of Learned and Professional Society Publishers, and we adhere to the ALPSP principles of scholarly-friendly journal publishing practice.

All research articles published by the RSC are peer reviewed. The journals are considered to be of the highest standards in their field, with an average impact factor of an impressive 5.4. Through the professional management of the publishing process, from submission through to publication, RSC content satisfies the pillars of scholarly publishing:

- Certification (validation of quality and integrity)
- Registration (recognition of achievement)
- Accessibility (unparalleled online access, worldwide)
- Archiving (reliable perpetual accessibility)
- Navigation (industry leading services to identify content)

Award-winning

RSC Publishing has been recognised by a number of prestigious awards, including the 2011 ALPSP Best New Journal Award for the high impact journal *Chemical Science*. The Award citation reads "Launched to present high quality cutting edge research across the chemical sciences, it has achieved swift success. There are very close links with the community and the journal is clearly defined by the science and the user,"

Professional

The publishing operation is based in Cambridge, UK, and employs around 275 people on the Science Park. These professional publishing staff engage in the preparation, peer review, selection, editing, production, marketing and distribution of information in the chemical sciences. Additional publishing staff are based internationally in Philadelphia and Raleigh, USA; Beijing and Shanghai, China and Tokyo, Japan.

Investing for the Research Good

As a Not For Profit organization, the RSC sustains its proven and established publishing activities primarily through subscription revenue. This model also enables the RSC to invest in new highly valued services for the community, generally at no additional cost to the user.

By way of example, during 2009 RSC Publishing acquired ChemSpider, a structure centric database for chemists. ChemSpider provides searchable access to over 26 million chemical structures and is considered to be one of the richest single sources of structure-based chemistry information worldwide. RSC Publishing provides free access to this service, as part of its publishing operations.

We welcome the opportunity to respond to the Office of Science and Technology Policy (OSTP) Request for Information (RFI): Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research.

The RSC is the largest organisation in Europe for advancing the chemical sciences. Supported by a network of over 47,000 members worldwide and an internationally acclaimed publishing business, its activities span education and training, conferences and science policy, and the promotion of the chemical sciences to the public. This document represents the views of the RSC. The RSC has a duty under its Royal Charter "to serve the public interest" by acting in an independent advisory capacity, and it is in this spirit that this submission is made.

Our comments are presented below, in response to the questions posed in the RFI.

(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?

- Peer review publications are an international product, disseminating information globally in a highly effective manner. Many channels of access have been developed by publishers, to ensure content can be easily discovered and accessed by those with an interest in consuming the material
- Researchers in North America have exceptional access to research content, with a recent survey suggesting that 97% of North American researchers have very or fairly easy access to research journals (source: www.publishingresearch.net/projects.htm)
- Consideration should be given to who may genuinely wish to access research articles, with reference to differing subject areas and public interest.
- Publisher led initiatives have recently opened up access to content far more broadly than ever before. In addition to pay per view, services such as DeepDyve (www.deepdyve.com) provide a low cost article rental service extending accessibility to content more broadly.
- Initiatives supported by publishers also provide free or very low cost access to the lowest gross national income per capita countries. Several such initiatives exist, including Reseach4Life (www.who.int/hinari).
- It is important that each research community (subject) is considered individually, as each have their specific needs and opportunities. Policies should reflect such nuances.

- Consideration should be given to sustainable archiving and the long term preservation of content. Efforts to provide free public access to content should work with stakeholders to ensure any policies do not adversely impact on the integrity of content archiving.
- Consideration should be given to the international nature of science publishing. Providing public access in one country is likely to provide access to researchers and other users throughout the world. Economic benefits are therefore likely to be no different for North America compared to other nations.

(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?

- The long-term sustainability of the system must be preserved, changes which impact on intellectual property rights must consider the impact on existing policies, processes and systems
- Policy and statements should be clearly defined, to differentiate what content may be openly (publicly) shared, and what may not.
- Policies should acknowledge the inherent costs involved in refereeing, copy-editing, typesetting, hosting, maintaining, preserving and making available scholarly journal content. Those who invest in these processes should be entitled to recover their costs to ensure sustainability of the systems for future generations.
- Funding Gold Open Access for researchers, as authors, enables all such work to be readily available to the public, immediately upon publication, in an accessible and reliable way. Gold Open Access is the arrangement whereby 'author side' payments replace 'reader side' payments (subscriptions), enabling costs to be covered and content to be freely accessible to all. Consideration should be given to supporting this option, including hybrid Gold Open Access, which works in harmony with established publishing systems.
- The consequences of making content freely available, particularly with respect to potential piracy and unauthorized onward distribution, should be considered carefully, and steps taken accordingly.
- Alternative means of disseminating the results of Federally funded work should also be considered. An example may be to make interim and final research / project reports freely and publically available.

(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?

- It is important that consumers of research content are confident that the article they read is the final 'version of record'. This typically resides on the publisher's content delivery platform and is widely accessible.
- Creating an additional platform for one subset of content, could result in confusion as to where readers should go to access the authoritative version.
- Creating an additional platform, with the functionality expected from each research community, is likely to be expensive and would be a duplication of effort. Such a platform would require continuous development, to innovate and deliver value added services as these evolve over time. Supporting or utilizing established systems, which are often tailored to subject based needs, could be a highly efficient and effective way to implement a public access policy.
- Existing platforms are highly interconnected, through publisher led innovations such as DOIs (digital object identifiers), and CrossRef (reference linking). The investment in these activities are largely supported by Publishers, to aid navigation and discoverability of content.
- As mentioned in response to Q1, access to content for research professionals in North America is already good. Enhancing access, through making content publicly available, may be best and most effectively serviced, through the adoption of Gold or hybrid Gold Open Access. This would work with established and reliable systems, and enhance access worldwide.

(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?

- We, as with most publishers, are open to collaborating with Government and/or funding agencies, to explore how we can be part of the solution to increasing access to research content in a sustainable way.
- The RSC is a partner in the Open PHACTS consortium aimed at creating an open innovative platform, the Open Pharmacological Space, which will be freely accessible for knowledge discovery and verification. Funded by the Innovative Medicines Initiative (IMI), a private-public partnership, it will also serve other IMI projects, the broader pharmaceutical industry, and other public drug discovery efforts. This demonstrates

(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?

- Interoperability is an area that all stakeholders have an interest in continuously improving. It serves the community well, by increasing awareness, discoverability and (easy) accessibility to content.
- Linking mechanisms, developed by publishers, have set a benchmark for further development. For example, CrossRef, and utilization of DOIs.

- Additional interlinking between articles and key databases provides significant improvements for users. For example, linking from an article to the (free) ChemSpider chemical structure database (www.chemspider.com), can then take users to other articles with similar or identical chemical compounds.

(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?

- Federal agencies do not directly invest in peer reviewed literature. However, the output from agency funded research does contribute to the body of knowledge published in these journals. The investment costs are borne by publishers, with this investment recouped through subscription or alternative business models.
- Supporting Gold or hybrid Gold Open Access may be a cost effective solution to enhance public access to research content. This may also be a relatively quick route to providing public access to research articles derived from publicly funded research.

(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?

- There may be scope to consider these alternative publication types. However, they have quite unique characteristics and would need detailed investigation and evaluation to ensure any policies do not harm the sustainability of content dissemination for future generations
- It is unlikely these types of publication would have as much flexibility as journal content, such that care should be taken to avoid damaging the fundamental goal of content dissemination. By way of example, if a book of (say) seven chapters was made freely available to all, it is very unclear how the publisher could recoup their investment in commissioning the work, improving the content, copyediting, typesetting, publication, dissemination, and author royalties, if a reader can access these chapters without payment.

(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?

- Each scientific discipline has different readership behavior, and should be considered independently.
- The embargo period route (i.e. Green Open Access) is not supported by the RSC. The period for the publisher to recoup their considerable investment in managing the entire publishing process would need to be preserved, and there is likely to be pressure to reduce the embargo year on year. When this causes widespread cancellation of journals, no funding is then available to support crucial publishing activities required to prepare the articles in the first place. This could cause the

entire system could collapse to the detriment of science. Green Open Access is generally considered to be a parasitic model or approach.

- In order to preserve the opportunity for publishers to recoup their investments, typically more than 50% of article downloads should be retained within the 'subscription access only' period. For the chemical sciences, this typically ranges from 12-18 months.
- If the intention of this policy is to make publicly funded research available publicly, the embargo period does not fully satisfy this objective. Content will not be publicly available until months or potentially years after first publication.
- Following from the above, if the intention is to make such content publicly available, supporting the Gold or hybrid Gold option satisfies this goal. Content is then publicly available, with no barriers or delay, immediately upon publication. Many funding agencies already support this route, allowing researchers (as authors) to use part of the funding to pay the author side fees associated with Gold Open Access. These articles are then publicly available, worldwide, immediately upon publication.
- Almost all publishers already provide the option of (hybrid) Gold Open Access. All RSC journals have this option, allowing authors to choose whether they wish to make their article publicly available.

Please identify any other items the Task Force might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.

- Consideration should be given to the global impact of any policy or proposal. Science knows no boundaries, and science publishing is also a global venture.
- Publishers, including the RSC, are open to working with agencies, to enhance access to content. This is a shared goal.
- Of paramount importance is the long term sustainability of the scientific publishing framework: certification, registration, accessibility, archiving and navigation. These factors are as relevant today, as they were when the first research journal was published in 1665.

Dr James Milne

Editorial Director and Acting Managing Director

RSC Publishing, Royal Society of Chemistry, Cambridge, UK