

Subject: Response to RFI

Date: November 7, 2011 3:13:24 PM EST

OSTP:

The Federal government, and by extension the U.S. taxpayers, generously fund basic science. In order that the government receive the highest possible return on its investment, and in order that taxpayers can read the results of science that they pay for, I think that the government should require that all scientific papers generated from federally-funded research be freely accessible to any and all members of the American public.

- Jason W. Barnes
Moscow, Idaho

Subject: RFI: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Date: November 7, 2011 6:07:14 PM EST

To Whom It May Concern:

California Sea Grant receives federal funds, which it awards for peer-reviewed research. In order to make the resulting scholarly publications available, we post abstracts, and when allowed by copyright, full text documents on one or more websites:

California Sea Grant website: www.csgc.ucsd.edu<<http://www.csgc.ucsd.edu>>

National Sea Grant Library website: <http://nsgl.gso.uri.edu/>

University of California eScholarship website: <http://escholarship.org/uc/csgc>

We believe that posting these documents on public websites is a cost-effective means of sharing them broadly with the research community and the general public. We also highlight research news in news releases that are widely distributed by email to the news media and research community, and which reference any peer-reviewed publications.

Please feel free to contact me if you would like additional information.

Sincerely,

Marsha Gear

Communications Director, California Sea Grant

From: Jonathan Graehl
Subject: **open access to publicly funded research**
Date: November 8, 2011 2:12:26 AM EST

All publicly funded research ought to be:

1) published in open-access journals (i.e. papers must be available on the Internet for viewing without cost or registration), as a requirement of receiving funding.

2) reproducible - policy should encourage sharing of experimental data (modulo privacy or other necessary licensing restrictions) and the methods and source code used to process and analyze it. As I'm no policy expert, I leave the details to our law/rulemakers.

Subject: Open Access of Federally-funded Digital projects

Date: November 15, 2011 3:11:39 PM EST

I think open access is a terrific idea—in theory. Where it runs aground is where digital projects/publications have a mandate both to be open access AND self-sustaining (I.e. Have a business model that can support daily and expanded operating costs). I don't know what the answer is but perhaps forcing all digital programs into granting open access to their final product has unfortunate consequences as far as making sure that the content produced today will still be available tomorrow.

--

Matthew Gibson
Director of Digital Programs
Virginia Foundation for the Humanities

Subject: Request for Information: Public Access to Peer-Reviewed Scholarly

Date: November 16, 2011 8:33:47 AM EST

To the OSTP –

An open access endeavor can only help the Nation as a whole grow in our learning and expand the possibilities of future advancements in science, medicine, history, or any of the other major areas of study that might be accomplished within the purview of the inquiries put forth at the Federal level. Knowledge is power and if the taxpayers are funding this research, then it should be made publically available for the public to benefit from its completion. While this may only interest a limited portion of the population at any given point in time (I believe that research institutions have the greatest potential to benefit) it could still potentially lead to the next technology which would keep the United States at the front of the global stage.

I believe that information should be, as much as is possible, made openly available to the public.

a servant of Jesus Christ...

a servant to all...

Mark Hanson Cedarholm Library | Circulation Supervisor

Bible Department | Adjunct Professor



"Let your manner of life be worthy of the gospel of Christ" - Phil 1:27

"Credit belongs to the man who spends himself for a worthy cause... (he) shall never be with those who knew neither victory nor defeat." - T. Roosevelt

Subject: Response to question of free access to digital data

Date: **November 16, 2011 2:51:36 PM EST**

I am attaching my response to the question of free access to digital data from the perspective of a historian and an editor. With thanks,

Holly C. Shulman

--

Holly C. Shulman
Editor, Dolley Madison Digital Edition
Founding Director, Documents Compass
Research Professor, Department of History
University of Virginia

The Cost of Free
Holly Cowan Shulman
26 October 2011

Digital publication is neither free nor cheap. Despite enthusiasm for scholarly publication at no cost to the end user, free publication will not solve the dilemmas that have upended the ecosystem of scholarly communications since the invention of the World Wide Web. As a community with many different stakeholders, ranging from scholars to universities, libraries to publishers, and funders to programmers, we need to think more clearly about the future of scholarly electronic publication.

The World Wide Web burst into public awareness as an open space for freely available communications in 1993 with the introduction of Mosaic, the precursor of Netscape. By the end of the 1990s that space had been further transformed by increasingly effective search engines, especially Google in 1998. As we all know, the impact was, and remains, a virtual tsunami on the world of communications. The WWW has already changed the environment for newspapers, magazines, music, political organization, and personal communications. It threatens cable and broadcasting. It is in the process of upending publishing and scholarly communications as one piece of this global shift.

The impact of this revolution has played out differently in various arenas. The cost of journals – especially those in Science, Technology, and Medicine (STM) – has skyrocketed while the budgets of libraries and institutions of higher education have decreased. In STM fields the importance of early publication is important, and publication online is a solution to the time factor. In some areas of the humanities online publication has posited a solution for the difficulties of finding traditional scholarly outlets. As the wages of most Americans over the past generation, including most academics, has flattened, the prospect of reducing the cost barriers of purchasing books, subscribing to journals and buying newspapers has been seductive, especially as new reading devices such as Kindle and iPad appear and their prices drop. The role of agents is threatened as new organizations are formed for authors who seek to self-publish their books. We are now at the point of vertical integration of the publishing industry as Amazon expands beyond a virtual bookstore to become a publishing house. As a recent report published by the British Research Information Network (RIN) said: “At a time of financial stringency for universities, research funders and publishers, it is important that all the stakeholders in the scholarly communications system work together to find the most cost-effective ways of fulfilling their joint goal of increasing access to the outputs of research.”

The United States government has weighed into this debate. Research funded by public monies, they argue – especially STM – should be freely available to the American public. STM journals have responded by charging their authors to publish, a practice now known as “author-side payments.” Already 40% of biomedical journals work this way. There is also the policy of online publication after a time barrier, which most of us know through JSTOR and MUSE as two examples of aggregators that sell their product rather than give it away. Some members of the digital humanities community passionately believe that online scholarly communications should be freely available to all as a matter of policy in order to open academic discourse to the widest audience possible.

Libraries are integral to this new mix. Where once they provided shelves, cataloging, reference, and an occasional rebinding services, they now must subscribe to new online publications and host their own catalogs and other works. In some cases libraries embrace their university publishers or have become close working colleagues of the university publishing house. There is a movement for open repositories where any scholar in that institution is expected (if not required) to post their work. Libraries have also hosted a world of experimentation such as the Institute for Advanced Technology in the Humanities and the Scholars’ Lab at the University of Virginia. Adding to the seduction of library publication are copyright rules that differ for educational versus commercial use. If a humanities scholar wants to publish images, for example, she or he may find it far simpler to stay within presently defined academic boundaries and publish their work through their library.

Within this shifting ecosystem of scholarly communications, lies the small world of reference works, and within that world the smaller arena of documentary editions, and even within that the very small field of documentary editions in history. This is the world of the “the papers of...”: Thomas Jefferson, Albert Einstein, Elizabeth Cady Stanton, Margaret Sanger, George Washington, the Freedmen and Southern Society Project, Thomas Edison, James Madison, Eleanor Roosevelt, Naval Documents of the America Revolution, Frederick Douglass, Abraham Lincoln, Cordell Hull, Andrew Jackson, Martin Luther King, and dozens more including my own Dolley Madison Digital Edition. While some have private funding, most operate through grants received from the federal government, specifically the National Endowment for the Humanities and the National Historical Publications and Records Commission. These funders are now considering the pros and cons of demanding not only electronic publication, but also free access to those electronic publications. The implicit question is: what are the costs of free? Do we need to unbind ourselves from the world of publishing in order to nurture creativity and open dialogue – or are these illusory goals that will destroy the armature of editions by replacing cost of purchase with cost of production, hosting, and maintaining?

*

It is my belief, to begin with, that a documentary edition should be neither a website nor an electronic book. Information on a website moves through a series of links that allows the reader to go from page to page. Its basic language is HTML. It is better suited to an electronic exhibition than an electronic archive. An electronic book transforms the printed page into displayable and searchable pixels that visualize a book for the consumer. It adheres closely to the traditions of print, plus perhaps a slider on the bottom, a page marker, a percentage number to locate the reader in the book, and various ways of highlighting the text and taking notes.

As a reference work and research tool, the documentary edition is best suited to the format of a digital archive, or a storage repository. After all, the edition may include thousands, if not millions of documents. Unlike a novel or an essay or a scholarly monograph, there is no argument being made, no narrative, no real spine at all on which to hang the bones of the edition. It is a collection

that is searchable through a process of identification and constraint. Each piece of information is like a pebble on a beach, and when you do a search you pull up from the beach a certain number of those pebbles. When you're done, those pebbles drop back to the beach. The scholar or student can read it from end to end, or browse through it, or search for what they want through different search tools. This representation, however, requires that whoever is publishing the work tag it in XML, most likely using the encoding guidelines of the Text Encoding Initiative (TEI), in a manner conformant to a whole series of programming demands. It also means that its display be reasonably intuitive so that the reader can navigate the text. In sum, it means that a well-done digital documentary edition should not be composed of PDFs taken from a printed book, or a series of pages simply coded in HTML, and posted on line. The internal structure of a digital edition is important. It may be altered; it may be expanded; but it cannot simply be thrown up on the web as can a blog, for example.

The central argument, however, revolves around what are the advantages and disadvantages of digital versus print. After all, a print edition is a wonderful object, replete with nicely formed pages, all kinds of fore matter such as preface and introduction, bibliographical information and a table of contents. It has an index that allows the contents to be searched. Why bother to move over to an electronic format? Are the reasons cost or scholarly benefit? And do these align on one side or the other of this argument over free?

There has long been an assumption that digital is cheap, in fact so cheap it can be given away for free. The hypothesis is in part built upon the way the medium grew in the 1990s when newspapers, for example, started publishing for free electronically as well as for a price in print. They rue the day. Today newspapers are establishing cost firewalls to protect their bottom line, while they fight off competition from bloggers and radio and television online products. Who wants to pay for The New York Times if you can get it for free? The print world of daily news and magazines is demanding subscription fees while also adding new content to give additional value to the online subscriber reader. The Times is a leading example of value added electronic materials rendering a superb electronic publication – at a cost. We all know that the print world is struggling to balance production costs with revenue. Amazon.com set a radically low fee to entice readers to purchase e-books so they could sell the Kindle, and as observed above are now trying to further cut their own costs through vertical integration of production by starting their own publishing organization. Meanwhile, the rest of the book-publishing world struggles to keep up. The American Association of University Presses recently issued a report tackling the problem of communications transformation entitled Sustaining Scholarly Publishing in which they lay out their own issues and explore viable solutions.

The cost of creating quality electronic scholarly publications in the field of documentary editions is, in fact, not cheap. Currently, funding grants carry a project through to the point where they are assumed to hand it over to someone who will then take care of it and somehow get it online. For starters, this model portends the elimination of a number of important editorial interventions that publishing houses undertake. They copyedit. They provide peer review. They help develop a product. They market. All of this has long been true in print. In the new world of digital, presses must also check data for consistency, extend code for new features, create new data interfaces, and so on.

They must have someone who can deal with the challenging world not only of markup languages but also of the specifications provided by the TEI. Without these standards there can be no consistency. Without this regularity there can be no interoperability or cost scaling or industrial output; electronic publication would instead remain a craft in which each producer created their own unique widget that would not play with their neighbor's or peer's widget.

Every edition needs to be “served,” or hosted on a server. A good one is costly; even purchasing a license to use a piece of it is expensive. Moreover open source software needs customization. Most editors simply cannot open up a box entitled Drupal, an open source content management system, and make it work for their project without external help. There are start up costs and there are platform costs and there are new content costs. And there are the continuing costs of maintaining a product in a world of constantly shifting technology in which what works today will not necessarily function tomorrow, of sustaining the machine and the human resources to maintain that content.

Imagine that you are the editor of a small project, and you have decided – eagerly or reluctantly – to be born-digital. That said -- to whom would you go? Who would you find as a programmer, an information systems person, or a designer? You would need a new way to control and track your work and your deadlines. Where would you find an appropriate, and free, content management system? What would your overall costs be? Whereas in print you would have simply handed over the manuscript to your publisher, who would have counted on recouping all costs through the sale of books, now you can no longer do so. You may fare well at a large university, but suppose you are an editor residing at a small institution. Or perhaps you are a retired professor and want to edit an edition as a retirement project. To whom do you turn for advice; what questions do you even ask? And if you have conquered the challenges of DTDs, XML, and TEI for encoding ordinary text, then what do you do when you later want to add different kinds of records such as inventories or ledgers? Two years on, can you find the same team who first set you up? Without a publisher, who guarantees that the work is not only quality scholarship, but even legitimate? Large-scale, well-funded projects can hire developers. Small projects usually run on a shoestring. And yet to date there has been NO talk about adding subvention costs to grant funding. The bottom line is that digital publication is neither free nor cheap. Three months of a programmer’s time might cost \$25,000. The cost of a production-level license for a first-rate XML publishing environment like MarkLogic Server runs upwards of \$30,000 at current prices.. And every time you wanted to add a new installment you would need editorial and technical work that would easily cost \$5,000 to \$10,000. And that is assuming there is a structure somewhere out there that can do for you what you want.

Electronic publication has important advantages: but only if done well. To begin with there is the compelling allure of cross-searchability, interoperability, and aggregation. The execution of this vision is, to date, tied to server and tagging issues. Thus Rotunda, the electronic imprint of the University of Virginia Press, can publish compilations such as Founders Online and make these editions cross searchable. There is a huge scholarly reference advantage here, but had Hamilton, Madison, Washington, and Jefferson all been published in discrete models this integration would not yet be possible. The Dolley Madison Digital Edition is now beginning to include lists of 300 names or so that will require a new kind of tagging; Rotunda is there, ready and willing and able to help accomplish this goal. Single volumes in a series such as the Papers of George Washington are now tied together and the researcher can do one simple search that carries her or him through all 60 volumes. Documents Compass, a unit in the Virginia Foundation for the Humanities, has been working on collecting all of the names and biographical references in all the founding editions and creating both a biographical dictionary and a prosopography. Funded by the Andrew W. Mellon Foundation, this will be a Rotunda publication.

The model of author-side payment is not a solution for documentary editions. Nor may individually built editions have the resources or technical infrastructure to create quality publications. I know of no digital humanities tools that confront these basic issues, so while it may be useful to have citizen transcribers using a tool known scripto (scripto.org), it does nothing to solve the basic

problems of publication, while it is unclear that it resolves issues of costs in order to professionally establish the texts at hand. While technology is upending the world as we once knew it, we should approach the problems of electronic editions more carefully and beware of solutions that promise free – at great cost.

¹ There were other web solutions in 1993 besides Mosaic but it is this author's memory that Mosaic stole the show. By 1994 a number of search engines started up including Yahoo! and Lycos, which the following year were joined by AltaVista, Magellan, and others. Google was launched in 1998. The combination of browser and search engine transformed communications.

¹ STM fields tend to be rapidly developing areas of research in ways not true in the Humanities and Social Sciences, and even before the WWW depended upon prepublication of journal articles.

¹ Julie Bosman, "New Service of Authors Seeking to Self-Publish E-Books," *The New York Times*, 2 October 2011, <http://www.nytimes.com/2011/10/03/business/media/perseus-creates-new-service-for-authors-seeking-to-self-publish.html>

¹ *Heading for the open road: costs and benefits of transitions in scholarly communications*
<http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/heading-open-road-costs-and-benefits-transitions-s>

¹ *Heading for the open road.*

¹ There is the Coalition of Open Access Policy Institutions (COAPI) in the United States. See Jennifer Howard, "Universities Join Together to Support Open-Access Policies," *The Chronicle of Higher Education*, 27 October 2011, <http://chronicle.com/blogs/wiredcampus/universities-join-together-to-support-open-access-policies/32632>

¹ *Sustaining Scholarly Publishing: New Business Models for University Presses* (AAUP, March, 2011).

¹ Matthew Gibson to Humanist Discussion Group (Willard.mccarty@mccarty.org.uk), 26 October 2011.

Subject: Comments on RFI

Date: November 17, 2011 10:54:35 AM EST

I am responding to the RFI entitled "Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research."

I am a NASA-funded astrophysicist, and a Senior Scientist at the Planetary Science Institute based in Tucson. I have been a principal investigator on over \$1M in NASA funding in the last decade. I have been an author or co-author on approximately 35 peer-reviewed publications.

As with any scientists, I spend many hours reading papers and articles from other researchers. While some articles are easy to obtain, others are not.

The internet has made it easy to locate potentially relevant research articles, using services such as the Astrophysics Data System (<http://www.adsabs.harvard.edu/>) and other online search engines. Many mainstream articles published within the last several years are available easily online. However, it is often quite difficult to access older articles, or articles published in smaller journals, or those in journals outside of my area of research.

My institution, the Planetary Science Institute, has a budget for online journal subscriptions and chooses to subscribe to journals most likely to be of interest to its employees. However, because of the diversity of scientific interests, there are always bound to be some journals which we do not have institutional subscriptions to.

These articles are often available on a pay-per-download basis, at prices typically \$30-\$50/article. This price becomes a substantial barrier, because even for low-cost items, institutional reimbursement policies are such that it may take several hours of work to get reimbursed for such a purchase, and such payment may be several weeks or even months down the road. Therefore, the mere fact of charging a per-article cost is a strong disincentive to read such articles. The research within them is ignored, and essentially wasted.

Of course it is possible to contact the author personally and request a copy. But researchers move, e-mail addresses change, messages are ignored, and eventually this may take several hours of work to to avail.

These issues are particularly frustrating when the work I am interested in was done by fellow American scientists, paid for by US tax dollars, under contract to NASA and other agencies just like my own research.

In response to your RFI, I suggest the following:

- 1) Existing publishers do an excellent job of reviewing, editing, producing, and indexing existing research. They provide a valuable service, and deserve to be compensated for this.
- 2) Some journals (such as Astrophysical Journal) have relatively open publication policies. They provide free access after one year, and subscription rates within the one-year proprietary period are reasonable for both institutional and individual subscribers. Page charges are high, but as a fraction of the total research cost, are on the 1% level (\$100K in research funding and \$1K for publications charges, for instance). I encourage this model. Articles published here are easy to access.

3) Other journals (such as Icarus, or Nature) are much more closed about access. These journals assess essentially no page charges, but retain proprietary rights. Subscriptions are difficult and/or expensive. Articles published here can be quite difficult to access, and as a result their research is under-utilized. I do not encourage this model.

4) Conference proceedings and books are a different category. These formats have not adjusted well to the electronic era. In my perception, fewer scientists are interested in publishing in these formats, because of their more difficult accessibility. I encourage new formats to be explored here. In particular, conference proceedings are often supported financially by a conference, so putting these (valuable) results online and accessible is extremely valuable.

Please feel free to be in touch with me if you would like additional feedback on this RFI.

Best regards,
-Henry Throop

Dr. Henry Throop
Senior Scientist
Planetary Science Institute
Washington, DC / Tucson, AZ

Subject: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Date: November 17, 2011 1:46:10 PM EST

I am responding to your RFI. I am a US citizen and taxpayer. I operate a private radio astronomy observatory in Anchorage, Alaska. I DO NOT receive, nor have I ever asked for, funds from the US government for my work.

I think it is imperative that all US taxpayers have free access to any and all US government funded research, and there should be no delay in providing that access.

I do not like having to pay for subscriptions to get US funded research data and papers when I already paid for all of it through my hard-earned tax dollars. Those subscriptions generally are outrageously priced. Some of these organizations claim to provide free access after one year. I do not like having to wait more than one day, more or less one year, to get this work from organizations who sell those subscriptions.

I have little sentiment for the organizations that sell my already-paid-for data through subscriptions. Essentially, I am being double-taxed, and I object to that. If they want to sell that information to foreigners, that is fine with me. But as a US taxpayer I should never have to pay for it, in any way, shape or form.

Whitham D. Reeve
Anchorage, AK USA

Subject: FW: Sharing of NASA research with the public
Date: November 21, 2011 11:37:28 AM EST

Per the request of Max Bernstein, I am forwarding this information.

There is and has been more than one mechanism in place to carry out the distribution of NASA science and technical reports and publication. There are probably too many outlets and the information is getting lost.

Sent: Monday, November 21, 2011 11:27 AM
Subject: Re: Sharing of NASA research with the public

Max,

I saw the request for info in the PEN.

Here is a thought, why create some new sight when there already is a sight, which is already tasked with the job.

NTIS National Technical Information Service.

The National Technical Information Service serves as the largest central resource for government-funded scientific, technical, engineering, and business related information available today. For more than 60 years NTIS has assured businesses, universities, and the public timely access to approximately 3 million publications covering over 350 subject areas <<http://www.ntis.gov/about/coverage.aspx>> .

Our mission supports the Department of Commerce mission to promote the nation's economic growth by providing access to information that stimulates innovation and discovery. **Public Law 102-245, Section 108 American Technology Preeminence Act of 1991** This includes almost every area of science and research that NASA carries out.

It is under the Department of Commerce, so ITAR and other export information should be addressed as well.

It also already contains some NASA reports, NASA publications in other journals should be able to be placed there as well due to the copyright exclusion for NASA funded research publications.

Subject: Public access to federally funded research

Date: November 21, 2011 11:58:06 AM EST

Further, There are existing federally funded information distribution organizations such as DITC, NTIS , NTRS, GPO and others, who should have specific guidance and distribution information concerning such reports. As Copyrights are forbidden on NASA and other federally funded research, the reports and journal articles should be available at distribution cost.

Failure of the availability of federally funded research information that is not classified, is a failure of the organizations to meet the letter and intent of the law or, a lack of appropriate oversight of publications, or a lack of funding to disseminate the information.

John S. Canham, Ph.D.
Senior Staff Scientist
Systems Engineering Group
ATK Space

Subject: RFI: Public Access to Peer-Reviewed Publications

Date: November 21, 2011 12:38:40 PM EST

In response to the Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research, I would like to briefly offer a perspective from a small college library in Minnesota.

At Gustavus Adolphus College, students and their faculty conduct research in part to extend our knowledge and in part so that students will be equipped with the habits and skills of inquiry so that they can be engaged citizens. In the sciences, in particular, our small library has difficulty providing access to essential research. We have had to cut journal subscriptions three times in the past ten years and this year have had increases in subscriptions to two essential journal packages from the American Chemical Society and SAGE amounting to 22% and 25% respectively.

In the face of these costs we have reduced individual journal subscriptions and purchasing books and videos, cuts that are particularly hard on our humanities and social sciences programs. Because publishers make it difficult for libraries to share electronic versions of articles, we find we are forced to buy articles published in journals we don't have directly from publishers at costs typically ranging from \$30 to \$50 per article. Though we are staggering along, this is difficult to sustain.

Liberal arts colleges like ours disproportionately educate future scientists. Many of our students go on to graduate school and careers in medical and scientific research and practice. We can't do this without access to the primary literature, which increasingly is only available at well-funded research institutions. The heroic stance taken by the NIH to develop a path forward in the face of commercial publishers' hostility, making good on the extraordinary and necessary investment we collectively make toward funding basic science, should be emulated by all federal agencies that invest in research. Without the results of that research, most of us are left with severely constrained access to knowledge that we have paid for in order to advance our understanding.

As a citizen, I feel closing the loop of funding research by including the dissemination of results is the only way our investment in basic science and other funded research will bear fruit. It's important for our national prosperity and important that we share American science with the world because global scientific and medical advances will help the world's well being. As a librarian at a small undergraduate institution where future scientists learn how to do research alongside their faculty who involve them in authentic research experiences, we need access to this material, and purchasing every article that might be relevant is not always possible.

Thank you for raising these questions.

Yours truly,

Barbara Fister
Department Chair
Folke Bernadotte Memorial Library
Gustavus Adolphus College

Subject: Input on question of Public Access to funded research results (8 comments)

Date: November 22, 2011 3:30:45 AM EST

This input is from:

Dr. David E. Hill

(no affiliation with public institutions or private corporations)

Each of the 'comment areas' has a lot of separate questions, some with leading questions that make a lot of assumptions with respect to values. I will try to address each of these separately.

1a. Yes there are ways to grow markets, but I don't know what interest the public would have in 'growing new markets.' The meaning of this question is impossible to fathom.

1b. The part about 'growing the economy' makes no sense at all. And, improving the productivity of the scientific enterprise? If journals were open access, people could share information more freely, and access information that is presently too expensive to access. This improves public participation in science. After all, Einstein was an amateur in 1905.

1c. Relative costs and benefits? Academicians generally don't get paid by publishers for their writing, they get paid by the government if the work is publicly funded. So, if libraries and individuals don't have to pay to read their work, everyone saves. The change is simple- it cuts out the profit-minded corporations that are trying to create scarcity out of abundance, so that people will pay them to access information that they have already paid (through taxes) to produce. The game only works because academics are afraid that their work will not be recognized if it does not appear in these journals, to the point where they give their copyright away to them for no compensation at all. But if the public paid for the work, they are giving away something that the public paid for, to a private party. Costs and benefits? Obviously, open access reduces costs for all, and can eliminate an unneeded middle-agency that is no longer needed. But then, what about ethics? Is the present system fair at all?

1d. Maximize economic growth? Is this a goal shared by people, or just one that is widely publicized by corporations and agencies that want to grow their budgets? As Robert F. Kennedy said, more crime, and more prisons, and the GNP goes up. Charge people for air, the GNP goes up. The measurement means nothing.

1e. Improve productivity of the 'scientific enterprise': Open-access allows a wider community to participate, not just those with access to university libraries. Isn't this what democracy is all about? Less time and effort spent trying to obtain material to read, and more time reading it. For-profit publication by its nature reduces readership, to maintain revenues.

2. 'Intellectual property interests' represent an important area to consider, but one that can be resolved easily. We already have copyright laws to protect intellectual property, and they are sufficient. However, when someone is writing on the taxpayer's dime, whether as an employee of a federal agency, or a person paid by a federal agency, then that work belongs in the Public Domain. It's as simple as that. It does not belong in the asset list of a private publisher, to be certain. As noted above, publishers take copyright wherever authors are foolish enough to give it away. Authors give it away, but it should not be theirs to give if the work was paid for by the public. For older works, publishers routinely and illicitly assume copyright (e.g., for works of old or deceased authors), when none was originally given. What a person does on their own time, on their own money, is their own and they can have copyright. That is fair. It's not complicated. Anyway, you can't copyright facts or ideas, just presentations.

3. You can do both. There is no reason why private archives should not mirror public archives of open-access documents. The cost of data storage and retrieval is now so small that this is not an issue at all. It will cost far less in the near future. Centralized administration of a national resource can ensure that it is done well and reliably, but this can also be supplemented by decentralized libraries or archives of on-line documents. The centralized administration can provide universal indexing services, to help everyone find relevant documents, and it can take this role out of the 'obligate cost' mix of other groups, for example public university libraries. At the same time, the presence of distributed libraries and related private indexing services can also help people to find and use these resources, but it should not be an obligate function.

4. We have seen the highest degree of innovation in non-profits like Wikipedia and Internet Archive. Publishers of print publications have no financial interest in furthering public access to publications, unless they can charge high prices for that access to maintain their position with university libraries that pay high subscription fees. Some companies like Google, with revenues not derived from print publication, have tried to facilitate access, as part of an advertising-based business model. Publication standards are presently rooted in the PDF, based on an Adobe standard that is available, and electronic document standards need to evolve to do a better job with recorded media, video, etc. The government can help with recognition of standards, and encouragement of the private sector to develop these. As noted, print publishers have no financial interest in promoting and enhancing these standards.

5. One of the best things the government could do would be to recognize and to promote new standards for early publication in advance of review, followed by continuous review of published work. Some open access publications have been very innovative in this area. Presently, peer review is of little use other than to keep a lid on the size of printed publications, and to delay access to all but a few who share manuscripts. So, putting out versioned manuscripts and adding comments as time passes is the way to go, and reflects what people really do, outside of the time constraints of the publishing systems. Likewise, the government just has to put published 'metadata' (e.g., gene sequences) into the same category and written text.

6. I think this question has already been answered. Since the solution that reduces costs to the public, and costs to public institutions like libraries, is the best one, there is no advantage for for-profit publishers to accept these obvious solutions. Cost of electronic publications is cut drastically because authors can format these, and they don't need to be printed and distributed, reducing our use of fossil fuels. This will reduce the GNP, but a point that I made earlier is that the GNP is not a useful measure of anything. I hope the government of this country does not exist to maximize its own tax revenues by raising costs.

7. The criterion needs to be, if the work was funded by the public, then it belongs to the public. If an author's writing is funded by the public, that work belongs to the public. If an author's research is funded by the public, but that author writes something else on his/her own time, then that author has the same right to use the publicly-funded work than anyone else, but cannot restrict access to it by others.

8. Embargo periods, like delayed publication in general, are against the interests of academicians who want to share their current work with the larger community. They simply restrict communication of timely work in a relevant time-frame. These are only useful to publishers that want to guarantee a revenue stream by restricting access. There is no secret about this. Some societies are very explicit in saying that they embargo to keep their revenues coming in. My personal view is that, you want to embargo your work for 12 months, fine, I won't

read it for 12 months. This has actually been a reality for me, except where I have been able to get authors to send PDFs and work around the system (almost all do). So, the embargo just keeps the 'casual reader' out, that person who the author would really like to communicate with to share his discoveries more widely, outside of the predictable audience. Delayed access is just that, it restricts readership and maintains revenue streams collected from those who think they have to be 'current'.

(9). Other issues. The heart of this 'problem' lies in the fact that, from the perspective of authors, academic publishing has always been a non-profit exercise. Publishing and distribution used to represent limited and costly services, and since people would not pay for the results, authors had to find ways to pay for the publishing directly. Over the last 30 years, societies have turned their publishing programs over to 'for profit' companies in the private sector, and these companies have taken advantage of the professional 'need' of academicians to 'get published.' They now routinely require that authors give them all copyright to their work, something that was not done many years ago, because they now see a continuing revenue stream can be obtained through direct sales, by preempting the public library goal of public access. The current situation represents a lot of 'conflict of interest' impacting the oft-stated goal of communication. And, regardless of actions taken by the US government, this situation is falling apart. Academicians really don't want it to continue.

(10) The US government could greatly facilitate certain areas of scientific research by creating or recognizing open-access repositories, such as a repository of chemical formulas, or scientific names. In Zoology, with only a few gigabytes of data in all to be concerned with (at 5 cents a gigabyte storage costs!), the ICZN can't seem to get its act together on an on-line repository of scientific names. It's just a political stalemate that print publishers would like to maintain as long as possible, as silly as this situation is. The government could help us to move out of the dark ages in this area.

(11) It is important to recognize, honestly, that there are serious conflicts of interest here. The bottler wants to own the well, to sell bottled water, and the taxpayer wants government employees to drink the tap water that they already paid for. It's that kind of problem.

Subject: Public access to gov't funded

Date: November 22, 2011 9:45:21 PM EST

I have been a PI on many Federally funded grants, and I am on the editorial boards of several journals. I would not have a problem with a policy that requires that papers supported by Federal funding be either 1) published in journals that makes them open-access one year after publication or 2) "republished" 1 year after publication to accomplish the same goal. NIH does this. Publishing is changing very rapidly. In my opinion a policy that requires open access in under 1 year is likely to put journals at a greater risk of going out of business than they are already facing.

--

Barbara Turpin, PhD
Professor, Environmental Sciences
GH Cook Campus Dean for Undergraduate Education

Subject: Public Access to Peer-Reviewed Scholarly Research

Date: November 29, 2011 1:56:15 PM EST

Tax-funded research findings should be in the public domain.

M M

MICHAEL MARDIKES

Subject: Re: [sparc-advocacy] Call to action: 2011 White House RFI on public access (deadline Jan. 2)

Date: November 30, 2011 3:07:48 AM EST

Dear Task Force on Public Access to Scholarly Publications: please see my responses to your questions in the [Federal Register Volume 76, Number 214 \(Friday, November 4, 2011\)](#) 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 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 increase access, simply introduce APIs (application programming interfaces) to allow people to write programs which can easily download the data (note - I have not reviewed existing APIs). And when I say data, I mean more than descriptive data - encourage actual scientific data to be structured into articles; for example, in the medical context this means numbers in trials, results of various tests, codes for randomization procedures, etc. As far as maximizing economic growth - making this information public may not be reflected in increased GDP, since GDP is a function of spending. However, it would certainly increase public welfare and provide great benefit to human beings in general by steadily increasing the scientific literacy of the general populace. This has additional effects of spurring innovation through inventions and also potentially greatly increasing the health of the citizens, as citizens are able to bring medical research to the attention of their providers or apply it to themselves.

(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?

This question is somewhat oddly worded - both questions are asking the same thing; thus, there is no need for two questions. Also, it uses the word "conversely" incorrectly.

The intellectual property question is a tricky one. As far as I'm concerned, the efforts need to be focused on minimizing the way that large corporate publishers, which act largely as just warehouse and copyeditors but, through holding title to a journal name, extract rent from the individuals which are creating the true value: the scientists, peer reviewers, etc. The government should not permit scientists to relinquish their and the government's intellectual property rights to a journal company like Elsevier, LexisNexis, Wiley, etc.

(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 pros of decentralized approaches outweigh the cons, but the best might be a mix of the two. Obviously, a highly reliable centralized source is necessary. as the goto source. But public data, far from being unusual, is *typical* for government-funded work. And the fact that government data is public record has been enormously helpful for the public. In most important areas, data is actually public: medicine, weather, countless sciences, and even corporate finance (look at SEC filings). Even in insurance, the industry I work in, data is aggregated and much is public as it is filed with the government.

Wikipedia and open-source software provide examples of what happens when you have ultimate decentralization, although these are fairly different in that they don't get tax revenues. There are plenty of examples of open-source software allowing groups to free themselves from inflexible bureaucracies and innovate. In theory Wikipedia faces the same thing and may some day "fork". The framework could theoretically be actually structured in such a way that this government institution could be subject to the same competitive pressure.

(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?

Finance has lots of examples, and the finalization of XBRL will bring even more. Obviously, there are winners and losers - and the existing companies are likely losers, just like the cable companies can be losers to say Netflix. On net I would expect overall administrative costs to decline, which means lobbying companies lose jobs and revenue, while a higher percentage goes into actual research and data analysis. Not sure if this qualifies as research per se. Otherwise, I don't know - altho public government data is used by private companies all the time, often with a little bit of "value-added" and then sold for a big markup.

(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 this will vary by fields, and I'm not qualified to speak in detail about any real fields - altho in the field I'm mainly interested, medicine, I would point again to numbers like numbers of individuals in a trial and compliance with various evidence-based medicine guidelines like blinding, randomization, intent-to-treat, etc. There's a vast amount of historical data that needs to be structured, and the project could open this up to the public to edit while allowing employees to supervise and do their own work.

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

Not sure. My impression is that the large publishers get such a big slice of the pie that if you cut their slice, the rest of the groups won't feel so much pain. Perhaps I'm wrong about that.

(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?

I don't know about this. I believe that scientists and peer-reviewers should be well-compensated for the work they do - I do not think that science is free or that it should always be free. My impression is that the current system does not do this efficiently. Embargo periods encourage payment, which is OK, but the payment is to the publishers, who then inefficiently transfer it to the scientists.

I am not a scientist. I got a couple BAs (economics, philosophy) in 2008, and have since worked as an insurance regulator in Alaska. I edit Wikipedia which is partly what my interest in publications stems from. I've benefited enormously from public access to articles, although to be fair the mere fact that abstracts are available online is probably even more of a benefit.

Regards,
Ben Creasy

From: Joe Hahn

Subject: **Public Access to Peer-Reviewed Scholarly Publications**

Date: December 1, 2011 11:35:52 AM EST

Affiliation/Organization Space Science Institute

City, State Austin TX

Comment 6

Federally funded scientists need online access to the journals in their field; they are much more productive with that online access, and less efficient without it. I suspect that many such scientists that are affiliated with smaller universities and research centers don't have online access to all of the journals in their field. Although my current employer does provide me with online access to these journals, past employers couldn't afford the cost of that access, which hampered my access to scientific publications, and made me less efficient.

My field is astronomy research, which is supported by grants from NASA and NSF. I would recommend that these agencies negotiate agreements or contracts with the major the publishers in the relevant fields so that these publications would provide open access to all readers to those journal articles authored by federally funded scientists.

I would also recommend that the relevant funding agencies also pay these publishers to provide an open access to past journal articles as well.

James Ryley, PhD, RPA

December 1, 2011

Mr. Ted Wackler, Deputy Chief of Staff
Office of Science and Technology Policy
Executive Office of the President
725 17th Street Room 5228
Washington, DC 20502

Re: RFI: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Dear Mr. Wackler,

The following comments are made in my capacity as a scientist, the President and Founder of SumoBrain Solutions, and a Registered Patent Agent. It will help put my comments in context to explain my background, what our company does, and why.

I am a biologist by training. As such, I am quite familiar with biomedical research and NIH's PubMed data, which will be discussed later. I am also a Registered Patent Agent with the US Patent & Trademark Office. And, in addition to being a Registered Patent Agent, I have been immersed in field of Intellectual Property for the past 7 years at SumoBrain Solutions. SumoBrain Solutions provides software, data and data analysis, and consulting in the Intellectual Property and Business Intelligence arenas.

SumoBrain Solutions also creates and maintains free web sites aimed at serving various research communities. For example, www.FreePatentsOnline.com (FPO) provides a free patent search engine, and www.BioMedSearch.com (BMS) provides a free biomedical search engine.

Both web sites mentioned above rely upon public access to large government document collections. In the case of FPO, it is patents and patent applications from many patent offices around the world. In the case of BMS, it is NIH's PubMed database.

Our mission is to add value to important sets of documents like these, in many ways, including the creation of federated databases, advanced search tools, data analytics, and the creation of user communities.

FPO was our first, and continues to be our most popular, web site. FPO has existed for over 7 years, and is arguably the most popular patent site in the world. We serve over 10 million page views per month. Our audience consists of attorneys, patent searchers, academics, librarians, independent inventors and others.

By improving upon the access to these documents offered at government sites, we have been able to provide a great deal of value to the patent community. We have over 1 million

registered members, and I think the fact that so many people choose to use our web site instead of the free sites offered by the USPTO, EPO, WIPO, and others is a testament to the amount of value.

And, though many large companies use our site, I like to think that it is particularly helpful to small businesses that may otherwise be unable to afford expensive subscription-based access to patent databases (which frequently cost between \$150 and over \$1000/month per person).

BMS also serves a substantial number of users each month. However, it has not attained nearly the community acceptance in the biomedical field that FPO has attained in the intellectual property field. There is one main reason for the different levels of community acceptance of these two sites: Patent documents are truly public access, while biomedical documents are not.

Patents have no embargo period, no restrictions on reuse and every document (going back to the founding of the US Patent Office in 1790) is available. The result of this unfettered access is that we have been able to add a tremendous amount of value to the patent documents without any of the drawbacks that would occur were the documents subject to any kind of restrictions (e.g., missing documents, lack of full text, or lack of ability to create derivative works).

In contrast, NIH's PubMed database, while also operating under a form of public access, is not freely available in a practical sense. With substantial embargo periods, reuse restrictions, and a great deal of historic data that has not been grandfathered into public access, comprehensive data is not available to the public. And comprehensive data is what is needed; no one wants to search a database where they might be missing crucial documents, no matter how good its other features may be.

Consequently, BMS has not been able to, and never will, add the amount of value to its community that FPO adds to the intellectual property community.

The contrast between FPO and BMS provides an important case study in the effect of different levels of public access: restrictions on public access, even "minor" restrictions that may seem like reasonable compromises, can completely negate innovation and competition in providing access to the documents that are the product of federal research grants.

The reason is simple: Users will always prefer a database with 100% of the documents over a database with less than 100%, even when the incomplete database might be better in every other way. Said another way, if a government or publisher web site, due to restrictions on public access, has documents that are not available to third parties, you will effectively have public access at only that government or publisher web site. This is not conjecture; as explained above, we have done it both ways.

My answers to specific RFI questions are below. Please note that my ideas on this topic are hardly revolutionary. So, I will frequently not go into great detail and instead provide references to those who have already eloquently expressed these concepts. I do not believe that the value of my input is in originating new solutions to the problems involved in implementing public

access. I think the solutions are already known. The problem is the number of possible solutions and the competing interests of various stakeholders.

Therefore, I see the task at hand as choosing among the many possible solutions already proposed in order to establish a public access policy that provides the most benefit to the stakeholders, and the nation as a whole. I hope that my input as someone familiar with public access from a variety of different viewpoints will help with that goal.

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?

Provide true, unfettered public access. With the appropriate data available, access and analysis will occur through competition. Our own company and the many companies which develop search, data mining, and data visualization tools are evidence of this.

The appropriate data and access to the data means full text and meta-data, with the ability to reuse in its native form, create analytics based on the data, and create derivative works.

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?

Public access can only help science and the economy if it truly adds value beyond the methods of accessing these documents that currently exist. That means that the private sector must be willing to provide search and analysis tools which improve access to information and productivity across many fields of science. Encouraging the participation of the private sector means lowering barriers to entry, both technical and financial, and allowing for viable commercial products.

While I will go into more technical detail elsewhere, lowering technical and financial barriers means that public access policy should provide for free bulk access to normalized documents and meta-data, preferably in a single repository [1], without an embargo period. The viability of commercial offerings based on this data, on the other hand, hinges largely upon ensuring that there are no reuse restrictions or embargo periods and 100% participation. A commercial search and analysis solution must rest upon complete data to be viable.

What are the relative costs and benefits of such policies?

I see very little cost and huge benefits. The costs to set up the actual document access are vanishing small compared to the cost of the research represented by the documents and data in question. For example, Houghton estimated that the cost of a national system of repositories in Australia would average \$10 million per year over 20 years.[2] While certainly the cost for US institutions would exceed that number due to differences in research spending and literature output, the point is that the cost of an open access system would be measured in millions per year, versus billions per year in research spending.

And, while a complete summary of the benefits of public access to research and the economy goes beyond the scope of this reply, I believe it is fair to say that the benefits are substantial and already well-documented.

For example, the ability to discover and access academic research is paramount to many research-related businesses, the cost savings from Open Access can be substantial, and the burden of locating the full text of papers that are not Open Access is estimated at 60 minutes per problem-publication, creating substantial lost productivity.[3]

Further, a benefit to cost ratio of 37:1 has been estimated for enhanced access to higher education research, a ratio far in excess of the ROI on research itself.[2] Given this information, it would seem that increasing the efficiency of access to research documents may be the single most cost-effective measure that can be taken to bolster research ROI.

The sole drawback I see is that publishers will be negatively impacted. However, I believe that this cost is small when compared to the potential benefits to science and the economy, and that alternative business models, such as author publication fees, can be used by publishers to offset the negative impact of robust public access.

Note that, while my understanding is that NIH's public access policy has not had a negative effect on publishers, for reasons already stated the NIH model of public access is insufficient to realize that vast majority of the benefits that could be reaped from true public access. Therefore, I assume that under future policies allowing for true open access, publishers will need to adopt new revenue models.

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

From the point of view of the private sector interested in adding value to the data, the requirements are free, instant, bulk access to structured data and meta-data, preferably from a normalized, federated database, with no restrictions on reuse. This creates the lowest barrier to the private sector, resulting in more competition to promote access and analysis.[1]

Our experience with patents makes it very clear that public access lacking any one of these attributes greatly increases the barriers to entry. Some of these drawbacks can be overcome. For example, multi-point access instead of single-point access drives up technical costs, but could be dealt with. However, an embargo period or restrictions on reuse rights cannot be "programmed away." If such restrictions exist, third-party entities will not be able to provide competitive solutions and so will shy away from entering the field.

With respect to the access that end-users of the data require, it would be the availability of fast, federated databases with state-of-the-art search and analytics tools (some of which do not currently exist, and which may be area-specific) which allow researchers to save time by quickly sifting through millions of documents to find those truly relevant to their needs. It is widely acknowledged that time savings (as opposed to monetary savings via the elimination of subscription fees) is one of the most important benefits of public access.[1, 2, 4]

- 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?**

These groups do not have aligned interests, so this question requires two answers.

With respect to intellectual property, let me address patenting first. I have heard concerns voiced that public access somehow compromises the patenting rights of researchers. This is absolutely untrue. The patent law is clear on the fact that once information has been made public, anywhere, it serves as prior art. There is no difference between a journal article appearing on a publisher's web site, in print, or in a public access database. Once the author decides to publish, public access is irrelevant.

Open Access is, however, very relevant to copyright. Full public access requires authors, institutions and publishers to give up some copyrights. However, most of the stakeholders in this equation are not adversely impacted by this.

For example, it would be rare that a researcher would want to limit the distribution of his manuscript, or would attempt to profit from it by selling copies (if we are talking about peer-reviewed articles, as opposed to, e.g., book chapters). On the other hand, the researcher certainly benefits from wide distribution of his work in terms of reputation, citations (and public access articles are cited 45-100% more than non-public articles [5]), collaborations, and possible licensing opportunities. The same comments could be applied to Universities. And, of course the public benefits from increased efficiency which impacts them more in terms of advances in science that improve the economy and quality of living, rather than directly through the availability of documents. On balance, the intellectual property picture for most stakeholders is extremely positive.

The unaligned interest is, of course, the publishers. The publishers wish to retain copyright to documents to which they have added value via the peer-review and editing process. And in the past, limitations on public access were accepted to protect the publishers' interests. For example, embargoes and a lack of reuse rights are key drawbacks in NIH's current public access policy. While effective in protecting the publishers' interests, these limitations (as I described in detail above) severely reduce the benefits of public access to all other stakeholders.

I believe that there is a true conflict here which cannot be avoided with compromises and clever legislation. If the full benefits of public access are to be realized, it must be accepted that such policies will indeed force change upon the publishing industry by requiring new revenue models.

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 believe the response to the above question addresses this. Immediate, complete public access does undermine the interests of the publishers (though perhaps not technically if one were to assume that an Author's Copy of each manuscript was made available via public access instead of the Publisher's Copy). However, immediate, complete public access is by far more advantageous than any compromise position to all other stakeholders.

- 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?

We aggregate data from the US Patent & Trademark Office (USPTO), the European Patent Office (EPO), the World Intellectual Property Organization (WIPO), the German Patent Office (DPMA) and others. After years of dealing with data coming from multiple agencies, and in multiple formats even when from the same agency, I can say with certainty that a centralized approach to public access would be preferable.

If the data is decentralized, particularly if each agency is allowed to create its own data standards, the task of aggregating and normalizing the data (and simply dealing with each individual agency when problems arise) can be substantial. This increases the barriers to entry for private entities wishing to provide access to, and analysis of, the data.

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 see no reason to allocate stewardship to private sources. The creation of value-added tools and services surrounding the content is an appropriate task for private entities, as it is in their interest to do that to the best of their abilities. Long-term stewardship however, is not. Nor is stewardship a burdensome task for the agencies. Storage space is inexpensive, as is distribution. The costs associated with creating and maintaining such a system would be trivial in Federal research budget terms, and any savings realized by allocating this role to the private sector would not be worth even a slight possibility of data loss or inaccessibility.

Note that several agencies, including USPTO, NIH and SEC already provide such stewardship of their own public data (although the USPTO shares such stewardship with Google, as discussed in Question 4). Our experiences with access to these repositories have been more than satisfactory. So, existing models would seem to support stewardship by the government as well.

I should clarify though that data access and data quality are two very different things. Our experience with access has been good. Our experience with data quality from various government entities has not been nearly as good, which is why I strongly suggest a centralized repository with stringent requirements for data, meta-data, normalization, OCR, a strong preference for “born digital” data, and other technical issues.

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?

With respect to ensuring stewardship, this seems rare, though one could look to the arrangement between Google and USPTO for one example. Google now provides bulk data access to some USPTO data. However, this situation is unusual in that Google seemingly did this “because they could,” rather than because there was a viable stand-alone business model.

Without a viable business model for private sector stewardship, this would generally seem problematic. For example, while SumoBrain Solutions could provide such stewardship (after all, we have to maintain the data internally anyway), we could not afford to do so for free. Perhaps there are efficiencies to be realized by assigning stewardship to companies like ours. However, I do not believe the cost savings would justify the possible drawbacks.

With respect to public-private models for encouraging accessibility and interoperability, this is not as rare and I would cite our company, SumoBrain Solutions, as a perfect example of this type of public-private partnership. We have over 50 million patent records from patent offices across the world (including the USPTO), 20 million biomedical records from NIH (PubMed), millions of trademark records from USPTO, millions of SEC records (EDGAR), and millions of documents from some of the largest scientific, technical and medical publishers.

We add value to these data sets in many ways, including the creation of federated searches, advanced search tools (including analytics for business intelligence), standardization of documents into downloadable PDF form, alerts and account features to help organize, annotate and share documents. In this way, we are doing exactly what public access advocates tout as the advantage of getting the private sector involved in public access. And, since there are viable business models surrounding the creation of new databases, search tools, and analytics tools, providing such services does not need to be legislated or funded. Assuming proper access to the data, market forces will ensure that such partnerships occur. They already occur where possible.

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?

I suggest a single archive with standardized data and meta-data. Multiple databases increase the technical barriers to entry, and the notion that there are clearly separable disciplines is problematic in itself.

For example, is a new type of titanium alloy with the promise to improve hip implants to be categorized as materials science or biology? Is a new compound with potential applications to cancer a chemical or medical topic? Worse yet, did the research I seek come out of NIH, NSF, or DARPA? Researchers should not have to worry about such questions.

There is no reason to insist on a one-to-one relationship between documents, areas of research and agencies. Yet, if multiple repositories exist such distinctions are implicitly being made.

In the end, the private sector will federate the databases if need be, since this adds value and therefore is a competitive advantage. But, the requirement that the private sector do this is another barrier to entry.

Regardless of the number of databases, standard formats and fields, both for data and meta-data must be agreed upon to facilitate cross-collection search and analysis.

To provide a simplistic, yet actual, example, consider that most documents have authors, while patents have inventors. Since that example is so simple, the solution is obvious: Consider inventors and authors to be synonymous in a federated search engine.

However, not all examples are so simple, and sometimes the meta-data necessary to make various distinctions does not exist. Consider a Word document which has headings that read “Title,” and “Abstract.” The meaning is very obvious to a human. It is not nearly so obvious to a machine. If the meta-data isn’t present to specify what text constitutes the title and what text constitutes the abstract, then we are reduced to language parsing, which cannot be done with 100% accuracy. And that is a simple case; in more complex cases, language parsing may not be feasible and certainly will not have a high degree of accuracy.

What are the minimum core metadata for scholarly publications that must be made available to the public to allow such capabilities?

This cannot be answered without taking the type of document into consideration. But, assuming a schema that caters predominantly to a typical peer-reviewed journal article, I would suggest the following fields.

Please note that I do not suggest these be the sole fields, or that they be in a flat structure. Rather these fields would probably be best incorporated into an XML structure that allows normalization and one-to-many and many-to-many relationships where appropriate. The table below cannot convey such structure, but given the fields and their relationships, creating an efficient XML version is fairly straight-forward.

I would also note that, due to the technical nature of this question and the extensive work that has already been done in the field (e.g., see Dublin Core and OAI-PMH), the actual implementation details go far beyond what can be provided in this response.

| Field | Comments |
|----------------------------|--|
| Abstract | |
| Author Affiliation | |
| Author Contact Information | |
| Author IDs | Unique author IDs to disambiguate authors of the same name |
| Authors | Separable, with a consistent order of representation (e.g., last name, first name) |
| Categorization | The actual data for, e.g., MeSH categorization |
| Categorization scheme | E.g., MeSH headings for biomedical documents |
| Chemical Formula | Standardized representations of chemicals are very valuable for machine |

| | |
|-------------------------------|--|
| | search. Patent agencies attempts this now, but current policies have created intractable problems in unambiguous interpretation of chemical structures. The technology exists; this is purely a legislative issue. |
| Date of Publication | |
| Date of Submission | |
| Document ID | |
| Federal Agency | |
| Format | |
| Keywords | |
| Language | In some cases, when not English, this can be hard to determine by machine |
| Publisher | |
| Rights | Specifies, in machine readable terms, reuse rights |
| Source | Generally a journal, with title, volume, and pages, but other sources, such as a conference proceeding, would have slightly different fields |
| Table and Figure designations | Allowing machine identification of tables and figures can be important to more sophisticated analytics |
| Title | |
| Type of Publication | E.g., conference proceeding, journal article, etc. |

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?

No response.

Questions 6-7: No response

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?

Immediate access is the only viable model if substantial private sector participation is desired. As I described in detail above, embargo periods are one of the most important aspects of public access policy with respect to private sector participation. This is because embargos are not technical issues that pose problematic, but surmountable, barriers to entry. Only legislation can eliminate embargos and so ensure that many companies in the private sector can create commercially-viable solutions.

I understand that this is not an empirical answer. This question is difficult to answer empirically because ideally what would be cited would be studies on the value added to the economy and the effects on all stakeholders under embargos of different lengths (including no embargo). However, since we do not have true public access at the moment, comparative studies cannot be done. And, while we do have open access journals, they co-exist with traditional journals. There are no major research ecosystems where open access has replaced the traditional model, allowing before and after comparisons to be made.

But, despite the lack of ideal data to help answer this question, I think it would be a mistake to think that the outcomes are not fairly obvious, even if not accurately quantifiable. I say this for several reasons.

First, I have spent the past seven year working in one of the few truly public access fields: Patents. At the same time, my company has attempted to add value to other data sets, including NIH's PubMed (operating under what is also ostensibly public access). The difference in dealing with truly restriction-free data sets versus data sets under some flavor of pseudo-public access is immense. Public access restrictions cannot logically do anything but hamper private sector participation.

Second, consider the research budgets and potential ROI involved. The research spending by the agencies in question has been estimated at \$60.5 billion annually, while total national spending on R&D is more in the neighborhood of \$378 billion. How much do you need to decrease the return on \$60.5 billion (or \$378 billion), through barriers such as embargos, before compromises that impair the efficiency of public access cease to make sense? Certainly it seems very plausible that the difference between truly effective public access and a lesser version of public access will cost the nation billions of dollars annually.

Third and finally, ignoring the involvement of the private sector and the important role it would play in the development of time-saving tools and analytics, consider the most superficial and obvious savings potentially afforded by public access: Reduced journal subscription costs for our nation's academic and research libraries. Those potential savings will almost certainly not be realized with any embargo period because libraries supporting research must allow access to

the latest publications. Even one month after initial publication is not acceptable in most fields of science; under any scenario with an embargo libraries will be forced to maintain essentially the same subscriptions they maintain now. Over the past two decades journal subscription costs have risen dramatically faster than inflation or library budgets, posing a burden on libraries and hamstringing researcher access to all the material they need to effectively do their jobs.

End Question Responses

In conclusion, I believe that this is an opportunity to bring research document access into line with what the internet and modern information retrieval technology have allowed for quite some time. The current system is an accident of history. To restrict access, in any substantial way, to the documents which are the products of billions of dollars in taxpayer money is not in the best interest of the people who paid for that research in the first place, much less most of the other stakeholders or the nation as a whole.

When thinking about these issues, I ask myself "How would this be done if we were setting up the system anew today?" I think keeping that question in mind helps avoid getting mired in how it has been done (which I believe resulted in a "splitting the baby" approach with NIH, producing legislation that was better than what existed before, but certainly not optimal), and rather focus on how it could be done.

Public access could be done in a manner that is much more effective than the current system, more effective than the current NIH public access policy, and in a manner that aids science and the economy to the highest possible degree by better leveraging the tremendous amount of research that is already being paid for, but not efficiently shared.

Sincerely,

James Ryley, PhD, RPA

References

1. Thakur, N. *Open access as a path to increased scientific productivity*. in *Berlin 9*. 2011. Washington, DC.
2. Houghton, J., *Exploring the Impacts of Enhanced Access to Publicly Funded Research*, in *The Socioeconomic Effects of Public Sector Information on Digital Networks: Toward a Better Understanding of Different Access and Reuse Policies: Workshop Summary*. 2009, Centre for Strategic Economic Studies, Victoria University, Melbourne: Melbourne.
3. Parsons, D., D. Willis, and J. Holland, *Benefits to the Private Sector of Open Access to Higher Education and Scholarly Research*, 2011, HOST Policy Research.
4. Houghton, J., B. Rasmussen, and P. Sheehan, *Economic Implications of Alternative Scholarly Publishing Models: Exploring the costs and benefits*, 2009, Centre for Strategic Economic Studies, Victoria University.
5. Houghton, J. and P. Sheehan, *The Economic Impact of Enhanced Access to Research Findings*, 2006, Centre for Strategic Economic Studies, Victoria University.

From: Warren G Lewis

Subject: **Pub access**

Date: December 1, 2011 11:24:50 PM EST

As a researcher at a biotech company, my access to published research was very limited. Being asked to pay \$30 to read each single article of published government-funded research reduced and slowed my access to cancer biology and biochemistry research.

Paywalls block learning, they block discovery, they block cures.

Now, at home or on the go, where much of my reading takes place, I can only read abstracts and news reports.

The data and figures, at minimum, should be provided free of charge.

_Warren

Subject: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Date: December 3, 2011 8:07:06 PM EST

I wrote an essay about that about the general topic of copyright policy for tax-funded and charitably supported research about a decade ago for the Markle Foundation when they requested comments on "Policy for a Networked Society:

<http://www.pdfernhout.net/on-funding-digital-public-works.html>

A shorter version of that is here:

<http://www.pdfernhout.net/open-letter-to-grantmakers-and-donors-on-copyright-policy.html>

Copies of both are attached.

From the executive summary of the shorter one:

"Foundations, other grantmaking agencies handling public tax-exempt dollars, and charitable donors need to consider the implications for their grantmaking or donation policies if they use a now obsolete charitable model of subsidizing proprietary publishing and proprietary research. In order to improve the effectiveness and collaborativeness of the non-profit sector overall, it is suggested these grantmaking organizations and donors move to requiring grantees to make any resulting copyrighted digital materials freely available on the internet, including free licenses granting the right for others to make and redistribute new derivative works without further permission. It is also suggested patents resulting from charitably subsidized research also be made freely available for general use. The alternative of allowing charitable dollars to result in proprietary copyrights and proprietary patents is corrupting the non-profit sector as it results in a conflict of interest between a non-profit's primary mission of helping humanity through freely sharing knowledge (made possible at little cost by the internet) and a desire to maximize short term revenues through charging licensing fees for access to patents and copyrights. In essence, with the change of publishing and communication economics made possible by the wide spread use of the internet, tax-exempt non-profits have become, perhaps unwittingly, caught up in a new form of "self-dealing", and it is up to donors and grantmakers (and eventually lawmakers) to prevent this by requiring free licensing of results as a condition of their grants and donations."

--Paul Fernhout (Edinburg, NY; Software Developer and Homeschooling/Unschooling Parent)

<http://www.pdfernhout.net/>

====

The biggest challenge of the 21st century is the irony of technologies of abundance in the hands of those thinking in terms of scarcity.

Subject: Lack of Public Access

Date: December 7, 2011 12:15:18 PM EST

Greetings! I am writing in regard to comment upon public access to federally-funded research through databases such as PubMed. When federal taxpayer dollars are used in research, the researchers have an obligation to ensure that their use of federal funds is accountable and transparent to the taxpayers by making the research freely and publicly available; this is especially so when the research involves a particular community, especially a community that cannot afford subscription access, either in print or electronically or both, to for-profit and non-profit research journals. For example, the following is unavailable to the average Leech Lake Anishinaabe citizen, despite the fact that the study involves Leech Lake and its citizens.

This seems an egregious disservice to the public, to Leech Lake citizens as both human beings and as health care consumers, to federal taxpayers, and to health care providers and educators and students in rural, economically-depressed areas and institutions.

Please consider the continued lack of economic and scientific opportunities on communities and populations with whom federally-funded unclassified information is shared. Miigwech / Thank you on behalf of my profession, the educational institution I serve and its stakeholders, and your constituencies.

The screenshot shows a PubMed search result for the article: "Demographic and genetic structures of reservation populations. 1. The greater Leech Lake (Ojibwa) reservation." by Robata DA, Polesky HE. The PMID is 4789268. The page includes options for "Display Settings", "Abstract", "MeSH Terms", and "LinkOut - more resources". On the right side, there are sections for "Related citations" and "Recent activity".

Melissa Pond, Director of Library Services
Leech Lake Tribal College



Leech Lake Tribal College provides quality higher education grounded in Anishinaabe values.

Subject: Survey of public access to government funded research and information

Date: December 7, 2011 12:58:50 PM EST

Good day,

As a life-long, professional librarian, I am concerned about public access to knowledge and information and wish to express my opinion for your survey. While I currently serve as a part-time college librarian at Curry College in Milton, Massachusetts, I am retired from a long career as a high school librarian in the town of Weymouth, Massachusetts.

During these years in libraries, I have observed that institutions struggle with limited funding, providing access to patrons at point of need and an ever increasing need to focus patrons on credible material as opposed to bogus information.

Below are some thoughts. Thank you for your consideration.

Sincerely,
Rosanne Aversa
Librarian
Levin Library
Curry College

(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?

(Forgive me for lumping all my responses into one answer.)

Professors, teachers and students need access to the newest research findings in order to keep the United States in the forefront as the leader of scientific research and development globally. Asking universities, colleges, high schools or even the public to use dated knowledge or information is unacceptable. In this time of proliferation of pseudoscience, the government has a responsibility to provide free and open access as much as is practically possible to peer reviewed, accurate data.

As more and more high quality, cutting edge research is published in electronic versions only, we risk restricting this knowledge to a limited few who are affiliated with institutions that can afford subscriptions to expensive databases if we do not publish publically. While the companies that sell this data are delighted to have exclusive access, this mode of dispensation profits only those companies and no one else. We may be restricting the information from the next Bill Gates or Steve Jobs. We cannot return to the Middle Ages when only the wealthy had access to books and print materials while the masses labored in ignorance because of limited delivery methods.

Open access offers a good return on investment as the broadest swath of people may use research findings to expand on the ideas presented, avoid duplication of effort and potentially innovate valuable

concepts. This is the vision of the future that the United States needs to pursue. The federal government is uniquely poised to provide the widest, free, public access possible to ensure equality and a more level playing field as well as to permanently preserve, archive, keep accessible and distribute information that taxpayers funded and should be able to read.

How to pursue this ideal I leave to those better versed than me is distribution methods, but I will always advocate for the widest possible access to the greatest number of people. Surely this vision will open the best possibility to expand use of research to grow our economy and ignite and inspire greater goals and productivity among thoughtful people everywhere.

Subject: Comment on public access

Date: December 7, 2011 1:20:28 PM EST

TO WHOM IT MAY CONCERN:

I am writing in support of the adoption of open access to all research resulting from the grants from federal tax dollars. It seems to me that the public has the right to access and reuse fully the results of public funded research. Although a short embargo may be acceptable to give publishers time to recoup their investment in publishing such research, the quicker the access the better in my view. Open access to research is an economic driver and helps driver the development of new jobs as new scientific breakthroughs lead to new technologies and new economic development generally. I that in recent years, the University of Pittsburgh and Carnegie Mellon University have combined to help create 300 startup companies based on federally funded research at these institutions. Open access helps to insure the broadest possible dissemination of knowledge arising from important federally funded scientific research and increases its value to the citizens of the United States.

We at the University of Pittsburgh have a long history of support for and implementation of open access publishing. The Library system has hosted several important discipline-based open repositories for a number of years and has a very robust institutional repository as well. In addition, we now are the publisher of more than 28 open access journals, headquartered at Pitt and other institutions in the United States as well as other countries. In our rather extensive experience, the utilization of research placed into open access venues is far higher than in traditionally published forms. Not only are these articles read more widely, they are cited much more often, and thus being incorporated into further research over time. In other words, the open access of this material makes it more useful to other researchers, and the faster it is available, the more useful it becomes.

I urge adoption of open access for all publicly funded scientific research in the United States to increase the return on this investment. The NIH and PubMedCentral is an excellent model to be applied much more broadly and the investment NIH has already made in creating an infrastructure for open access can be leveraged for other science areas. This requirement should be for open access with rights to re-use articles fully in digital form to maximize the value of the public investment in scientific research. I strongly urge that such open access placement carry appropriate licensing for re-use such as Creative Commons CC-BY which are compliant with the existing copyright law.

An alternative to a federal repository for this research could be a distributed archives involving institutions such as ours with an infrastructure to support such a system, so long as the system allows for access and use conditions that allow all interested parties to build on them as opposed to dark archiving this material. I am confident that Pittsburgh and many other research universities with extensive experience with these kinds of repositories would be more than willing to help develop a distributed archive system.

Finally, public access policies should ensure the requirements for open access of federally funded research be uniform across the government funding agencies and follow current best practices with regard to metadata, protocols, integration with things like grant management systems and assessment.

Publishers recoup the cost of publication and dissemination of research articles with a few months

of their publication. Libraries are the primary market for these journals and we all pay for subscriptions to them in advance. So the argument that a publisher needs one year or more before an article appears in a repository or is open to use in the repository is a highly questionable argument. Publishers do not make their money on archived journal articles, but on subscriptions. Open access in medical research has been in place now for several years with no negative impact on the sale of commercial journals in medicine. Any loss of revenues they might have experienced were much more like due to libraries having to cancel subscriptions because of the rising costs of journals coupled with reductions in budget support being experienced widely in higher education.

Again, I urge adoption of open access to all publicly funded scientific research.

Sincerely,

Rush G. Miller, Ph.D.
Hillman University Librarian and Director
University Library System



Subject: OSTP Requests for Information on access to scholarly publications and data

Date: December 7, 2011 1:42:51 PM EST

I am very concerned about the RFI that was posted in the US Federal Register surrounding "Open Access". At risk is serious compromise in the peer review system that is a principal safeguard for the vitality of our national biomedical research enterprise. While the peer review process is costly, the financial responsibility is primarily borne by publishers, thereby maximizing availability of funding for direct support of investigation. As a scientist with a large federally funded research program, I fully recognize and appreciate the importance of peer review and consider the protection of copyrighted publications as indispensable.

Open access can negatively impact on the capability of independent publishers to remain solvent. Beyond deterioration in the rigorosity of peer review, open access can have broad economic impact. The financial consequences can extend to scientific organizations that rely on copyrighted research publications as a key and rate limiting component of financial stability. In the current economic climate, the income generated from copyrighted publications cannot be overstated.

I am confident that enhanced mechanisms can be put in place for open access accessibility to research findings from federally supported programs without jeopardizing copyrighted publications. This position is directly supported by a series of ongoing initiatives, which are in part supported by publishers, to disseminate research findings in a responsible and timely manner. Equally relevant, approaches are being developed to translate complex research findings into formats and language that are accessible to both scientists and the general public.

I strongly urge retention of copyright protection and retention of intellectual property rights as a critical safeguard for the biomedical research community and the public, with the understanding that there are substantial economic implications.

Gary S. Stein

Gary S. Stein, Ph.D.

The Gerald L. Haidak, M.D. and Zelda S. Haidak Distinguished Professor and Chair of Cell Biology Professor of Medicine Co-Director, UMass Memorial Cancer Center

Department of Cell Biology

University of Massachusetts Medical School

Subject: comments on the 2010 America COMPETES Act

Date: December 7, 2011 9:30:39 PM EST

Dear Sir or Madm,

Is there any other realm of public spending where the public is not entitled to partake of the benefits?

The US government's massive spending on research is to be applauded and sets a high bar for every other country. However, the full benefits of this research, both to the US based researchers and to others, has been hampered far too long by privatizing this intrinsically public data and more recently, delaying it's release into to the public domain.

I have performed research at world class institutions in the US (MIT and U. Washington) and overseas (ICDDR,B in Dhaka, Bangladesh and CICAMS in Beijing, China). In each of those contexts I experienced some degree of restricted access to research findings which were funded by US federal grants. In the US it has meant wasted time trying to gather the articles from other sources and institutions. However, for my colleagues overseas, restricted access has often left them without alternatives and therefore may limit their ability to properly execute and interpret studies to their fullest potential. This backward pressure benefits no one.

It is not only logical that publically funded research should be made freely available to the public at the time of publication, but it is in the long term interest of all of us as well.

Kind regards,

Danny Colombara

Danny Colombara, MPH
University of Washington

Danny Colombara, MPH
Department of Cancer Epidemiology,
Cancer Hospital/ Institute,
Chinese Academy of Medical Sciences and Peking Union Medical College

Subject: Comment on Public Access to Scholarly Publications

Date: December 8, 2011 6:44:14 PM EST

Dear Task Force on Public Access to Scholarly Publications:

On behalf of the J. Willard Marriott Library at the University of Utah, we submit the following comment. Thank you for providing this opportunity.

Joyce Ogburn, Dean and University Librarian, J. Willard Marriott Library
Rick Anderson, Associate Dean for Scholarly Resources & Collections, J. Willard Marriott Library
Allyson Mower, Scholarly Communications & Copyright Librarian, J. Willard Marriott Library

December 8, 2011

Public access to scientific literature remains essential for our knowledge economy. Seamless, web-based access to scholarly publications maximizes the return on government investment in research, enables economic development, and benefits the tax payer. The increased productivity and drive for scientific innovation necessitates a slight shift in the research infrastructure. As information has rapidly become networked and digital, how it gets paid for and made accessible needs to come into the digital age as well.

Studies have shown that most academic researchers look for scholarly journal articles using the Web.[1] Access to the article depends on what the researcher's campus library subscribes to. Most library budgets remain flat or are in decline while journal subscriptions typically rise at 6-10% annual rates.

While library services such as journal subscriptions and interlibrary loan seek to meet research needs, non-availability still impacts researchers as they search for articles on the web.[2] Many of them abandon the article, email the author or hope that someone has publicly posted the publisher's PDF version for free usually at the detriment to a publishing agreement. This fragmented model makes searching onerous, inefficient and it works outside the legalities of the current system.

Question 1

Agencies could improve scientific productivity by pushing to make final, peer-reviewed versions of the research write-up publicly available. While this may not grow new markets, it will certainly improve the fragmented system described above. As researchers conduct their web-based literature search, they can click through to the relevant article no matter if their library has a subscription and without having to rely on extra steps such as interlibrary loan or emailing the author.

In terms of growing new markets, agencies can expressly support this by means of dedicated funding lines. We envision a model where the government supports not only the original research, but a certain level of the production and dissemination that allows publishers and librarians to collaborate for the benefit of the researcher and the reader.

For every \$1 spent on conducting research, another \$1 needs to be allocated for making the results widely and freely known to tax payers as well as those who benefit the tax payer--nurses, physicians, educators, entrepreneurs, social workers, engineers, demographers, scientists, land managers and the like. The scholarly communication infrastructure already exists to support the production, distribution,

attribution, and preservation of publications, but the funding mechanisms need improvement.

Question 2

As noted above, many authors post the published version to open websites despite the publishing agreement and many researchers rely on this when searching for articles on the web. As original creators, authors tend to see articles as their own intellectual property and the traditional transfer of rights to publishers becomes a mere formality. Intellectual property terms that come more in line with established perceptions and normative practices would make for a less fractured system. In a 2006 study, researchers found that most authors preferred to retain their copyrights and allow for any non-commercial reuse.[3] By means of Creative Commons licensing, authors can express such desires without government agency involvement.

Question 3

Health sciences researchers and librarians know to start any life science-related search in the National Library of Medicine's MEDLINE database. It is the one place anyone can go for relevant literature and, could be argued, a driving force behind the rapid growth of biomedical research. Without a national library of science, engineering, or education such disciplines are somewhat at a disadvantage when it comes to information access because there is no centralized index. However, with an index comes the expectation of full access to the related text and agencies may not be equipped with this expertise. Decentralized approaches, in this respect, may be a better approach. Academic and research librarians distributed across institutions of higher education have such expertise.

Question 4

It is interesting that publishers' archives are noted, but many archives of journals exist in libraries or are held by third parties on behalf of libraries and publishers. Often other related research material is held by institutions and the responsibility for archiving and stewardship rests with these institutions, not publishers. This information could be more creatively linked and combined than at present. But in so doing theneed to ensure that anything that has been publicly accessible and open to use remains. Having more open article access would facilitate these linkages and recombinations.

Question 5

Access points such as title, author, source, and subject are basic essentials for bringing together a range of information products. Using well-established schemas and controlled vocabularies further the goal of interoperability and cross-searching. For example, the National Library of Medicine employs bibliographers to describe items and assign appropriate terms from the list of medical subject headings. Agencies would need to use similar approaches.

Question 6

In order to account for differences in disciplines, allow authors to choose their specialized publisher, but give them express budget lines in grants for supporting the costs of publication and archiving. The funds would support publishers as they facilitate the production of discreet, citable scholarly objects (either articles or data sets) with equal amounts of support going to libraries by means of increased university facilities and administration rates for maintaining persistent access and/or supporting publication costs. Publishers and librarians have already partnered in this area with such initiatives as the Compact for Open Access Publishing Equity and, of course, PubMed Central. However, such partnerships are not yet systematic across federal agencies and stakeholders and a dedicated budget line would incentivize wider collaborations and potentially spur new business opportunities especially in terms of data publishing.

Question 7

Yes.

Question 8

The answer to the question regarding an embargo period is probably best tied to the nature of the various disciplines and how timely access and repurposing of information needs to be. For example, early access to scientific and medical information is critical. A parent who is seeking authoritative and current information to make a decision regarding the medical treatment of a child has an immediate and urgent need. Although you ask for empirical evidence, the question also goes to value and the insurance of long term and open access to the fruits of research. Libraries are one of the great equalizers in societies worldwide and library budgets are increasingly under extreme pressure due to publisher policies and pricing strategies. Nevertheless, and even in an environment of openly available research, libraries will need to acquire and preserve definitive versions of research, usually produced in the form of formally published journal articles and books. On the empirical side, there has been no evidence that demonstrates that public access policies and short embargo periods have negatively impacted library subscriptions.

[1] Bradley Hemminger, et al. Information Seeking Behavior of Academic Scientists. Journal of the American Society for Information Science & Technology. February 2007.

http://ils.unc.edu/bmh/pubs/Information_Seeking_Behavior_of_Academic_Scientists-JASIST-2007.pdf

[2] Research Information Network Report. How researchers secure access to licensed content not immediately available to them. December

2009. <http://www.publishingresearch.net/documents/RINOvercoming-barriers-report-Dec09.pdf>

[3] Esther Hoorn and Maurits van der Graaf. Copyright Issues in Open Access Research Journals: The Authors' Perspective. D-lib Magazine. February 2006

<http://www.dlib.org/dlib/february06/vandergraaf/02vandergraaf.html>



December 8, 2011

White House Office of Science and Technology Policy
Request for Information on public access to peer-reviewed scholarly publications
resulting from federally-funded research.

RESPONSE from the Duke University Libraries.

(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 most important and effective policy to accomplish these goals would be to require immediate public access to research publications that stem from federally-funded research via an easily searchable database or set of federated databases. Such public access should be free of charge to all users and should include a broad range of reuse rights so that users can build on and innovate from the research that they find. This greater access to the huge amount of scientific research that the federal government funds would result in faster follow-on research, new products and services and more jobs. It is a relatively simple and cost-effective way to dramatically increase economic growth and global competitiveness.

In his Directive on Open Government, President Obama stated that government should be transparent, participatory and collaborative. Increasing public access to federally funded research will serve all of these goals. It will also make government more accountable for how it spends taxpayers' money; such accountability is also a benefit the President promised as part of the open government initiative.

In order to achieve these benefits, a mandate for public access should be more comprehensive and include a broader set of reuse rights than the National Institutes of Health public access policy currently does. Commercial reuse, for example, is an important component in using public access to foster commercial innovation and growth, especially because it can benefit smaller and start-up companies. Reuse by computers - the ability to treat large repositories of text as data from which to discover new connections, collaborations and paths for exploration -- is another vitally important advantage of broader reuse rights.

A broad and comprehensive collection of federally-funded research publications across a wide variety of disciplines will foster discovery and thus support innovation. The

ability for more people to discover unexpected opportunities and interdisciplinary connections is a key to this benefit. Public access will foster subsequent research at a much faster pace than is currently the case. Also, public access takes account of the near certainty that there are unexpected readers and users who can innovate in ways not anticipated by the original researchers or their funders. These unanticipated users, who include commercial innovators, small companies, under-resourced researchers and potential collaborators, offer the best hope for true innovation. Nearly every experiment in open access exposes these opportunities, and federal agencies should design public access mandates to exploit them.

One research study that confirms the benefits of public access is the Working Paper by Heidi Williams of the MIT Economics Department called "Intellectual Property Rights and Innovation" in which she studied the follow-on R&D based on two types of gene-sequencing research, one proprietary and one publicly accessible. Williams found that the proprietary efforts suffered a 30% reduction in subsequent scientific research and product development.¹ This low-growth approach is now the norm for most federal agencies, and that must change.

Another study, by the Joint Information Systems Committee in the UK, found that the cost/benefit ratio for a static-state model of public access to funded research through national repositories was 122, or a five-fold return on investment.² Other countries are rapidly discovering the advantages of public access to scientific research and information, and the U.S. must do so as well in order to compete and grow.

These benefits can be obtained at a very reasonable cost. The National Institute of Health reports that the cost of making its funded research available to the public is approximately 4 million dollars out of a 30 billion dollar budget. Their database is used by over a half million users each day, suggesting that there is great demand for this information and that a very high return on investment can be achieved at modest cost. As access and reuse is increased beyond the NIH model, that return on investment will rise even higher.

(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?

When considering how best to manage the intellectual property rights associated with federally-funded research, it is important to remember that the initial rights holders,

¹ Professor Williams' paper, in the National Bureau for Economic Research working papers series, is found at <http://www.nber.org/papers/w16213.pdf>. (accessed Nov. 23, 2011).

² The JISC report can be accessed at <http://www.jisc.ac.uk/publications/reports/2009/economicpublishingmodelsfinalreport.aspx>.

for whom the system is structured, are the researchers and authors themselves. Attention to copyright, for example, is important insofar as it serves as an incentive for new discovery and authorship, but it must be recognized that copyright can also stand in the way of such innovation. A balanced approach that is focused on the incentives for new creation and discovery is called for.

In the academic world especially, the incentive purpose of copyright is short-circuited by the fact that authors are not paid when they publish their works. Instead, the rewards for doing research and publishing scholarship are all provided by the universities and the funders operating entirely outside of the structure of copyright. The benefits from the copyright monopoly thus are absorbed entirely by publishers who have a limited role in incentivizing new science. Whatever limited concessions must be made to publishing interests, therefore, should be weighed against the need for rapid and diffuse advancement of science, and should not be allowed to interfere with that goal.

The copyrighted works that would be the subject of public access policies benefit their authors most when they are widely available. For most scholars, copyright is a concern only insofar as it serves their desire to have the greatest possible impact on their research field.

For this reason, public access does not threaten the incentives for the creation of scholarship at all. In fact, it enhances them because it supports faster research, more innovation and greater impact on the field. Numerous studies have shown, for example, that open accessibility of scholarly articles increases the number and the pace of citations.³ Many scholars have recognized the benefits of leveraging their copyright ownership to provide greater access and fuller reuse rights than the copyright law itself permits (under fair use), often by using Creative Commons licenses on their works when they are able to do so.⁴ Thus the evidence suggests that the rewards system that actually provides the incentives for a large group of scientific researchers is supported by public access, not threatened by it.

(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 most important feature of a public access policy is the ability to search across a broad and interdisciplinary set of databases. These databases must be

³ The classic study identifying this citation advantage is **Eysenbach, G.** 2006. Citation advantage of open access articles. *PLoS Biology*, 4(5), 692-698. Available at <http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0040157>. (accessed Nov. 29, 2011).

⁴ <http://creativecommons.org/licenses/by/3.0/>.

crawled by search engines, since even highly technical scientific research these days often begins with Google Scholar, especially when a new field or sub-field is being broached. Under these conditions, multiple databases are fine. But it is important not to allow private entities to have too much control over access and functionality; doing so risks undermining the very benefits a policy is indeed to support.

In order to support long-term preservation and to provide the necessary level of accountability that is one of the hallmarks of a public access policy, at the very least the government should maintain a mirrored and accessible version of the databases, regardless of who the primary managers of the repositories are. And, of course, the government will be responsible for setting and maintaining the standards under which multiple repositories would operate. Again, government openness and accountability are a key part of public access policies, so ensuring interoperability, access, persistence (commercial firms do go out of business) and uniform standards are key.

In truth, a set of centralized databases may prove both more effective and more economical. The NIH has discovered that running a centralized database for funded articles is not very expensive and yields very high returns, as will be discussed more fully below. Also, a centralized database (whether maintained by the agencies or through partnerships with private entities) makes standardization easier, both for submission and for discovery and use. Authors or publishers would only need to learn (and support services like libraries would only need to support) a single or small set of submission workflows and APIs. And any number of third party services (free or commercial) could develop discovery and use interfaces against a single API, stimulating innovation in the discovery and use sphere.

Where agencies will create and manage repositories, it would also be preferable to use open source software. The ability to innovate the platform for searching and sharing such data may be critically important to enabling effective reuse of these contents. Alternatively, the ability to download the contents easily and to export to another platform might enable such innovation. If there were a way to prevent those who might develop private databases that incorporate this public information from laying hurdles, such as software patents, in the way of others who might engage in such innovation, that would be welcomed.⁵

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 models for public-private partnerships that are most relevant for supporting public access to federally-funded research are those with universities. It is an unfortunate fact that, in the area of scholarly communications, commercial publishers have often proved to be unreliable partners because the overall need for profitability has led to resistance to the very features that are key to a successful public access policy.

⁵ This point was made to us by Dr. Anthony So, Professor and Director of the Program on Global Health and Technology Access at Duke University.

Researchers themselves frequently express distrust of the publishers on whom they still must rely for dissemination of their work; it would be unfortunate for a public access policy to begin its life encumbered by this distrust. Universities, both public and private, have simply proven themselves to be better at cooperation, especially over long-term projects. The example of the ArXiv database, which was begun at the Los Alamos National Laboratory and is now hosted at Cornell University, is an excellent example of a working partnership between governmental and private agencies that has already provided a huge benefit to the fields of physics, mathematics, computer science and quantitative biology.

Regardless of where partners are found, standards for access, preservation and interoperability must be maintained. If privately managed repositories can be federated in a way that supports innovation, then such partnerships are fine. But federation does not merely mean the ability to search multiple platforms in a passive way. To really foster innovation, it is important that creativity and experimentation be allowed even as to the tools that will be used in searching the publicly accessible articles. The ability to create new kinds of search tools, and, critically, to compute on the contents of these repositories (text-mining, etc.), must be afforded to users in order to exploit the full range of opportunities for innovative research and economic 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?

It is important to understand that metadata is a key aspect of achieving the benefits of open access, since good, rich metadata facilitates the cross-disciplinary discovery and even the ability for machine-reading research that will improve scientific productivity. Metadata based on the Dublin Core should be a baseline standard, and it must be coupled with a linked-data API that supports standards-based data exchange. Again, developing a metadata scheme that can support machine-readability and interoperability is vital. Agencies that have long experience in the development of metadata schema, including NISO and the Library of Congress, should be involved in the creation of appropriate standards. In general the best approach is to look to models that are already successful. Libraries and other groups already have working metadata schemes, and using such a model means that researchers and librarians (as well as automated systems) will already have developed search techniques that work with the scheme. Thus the benefits sought by requiring public access can be accelerated or, properly-speaking, not delayed unnecessarily.

One important function of good metadata standards is the ability to track and analysis usage. The COUNTER project (Counting Online Usage of Networked Electronic Resources) is an important initiative to provide this kind of analysis and

should be a significant partner in any repository development.⁶ This kind of analysis will support accountability and allow federal agencies to quantify the return on investment benefits that will certainly follow from public access policies.

Another way in which metadata description can support the benefits sought through public access is by linking research publications with the data that underlies them. Again, this must be done in a way that is meaningful to readers but also navigable by computers. Unique identifiers and vocabularies that support semantic relationships will make these repositories much more valuable and, again, able to return the kind of benefits that science and commercial innovation depend upon.

(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 2006 report by Professors John Houghton and Peter Sheehan of Victoria University in Australia describes and models the dramatic return on investment that public access offers, showing impact ranges from 25 to 75% on research and development when access to funded research is increased.⁷ The simple fact is that public access is a very good use of tax money, offering an extremely high return on investment, as well as an increase in government accountability. The example of the National Institutes of Health shows that even a public-access policy that relies on an agency-managed repository can be done for a very modest cost, and the benefits of that investment have uniformly proved to be immense. Houghton and Sheehan demonstrated that increasing access and reuse beyond what is possible from the NIH will further increase these benefits.

The most important step that can be taken to minimize cost and maximize compliance is to develop consistent policies. Across agencies and across platforms, the requirements and procedures should be as consistent as it is possible to make them. Consistency will reduce the manpower necessary to educate users and to troubleshoot compliance problems

Another possible step to take – which would be difficult to maintain at scale but could serve as an initial step – is to consider the option of requiring deposit in any one of an identified set of vetted and trusted repositories, including, perhaps, disciplinary and university repositories, and then harvesting across these repositories in order to create access to the body of funded research. In this model consistency is also paramount, not only in the requirements and procedures for compliance but also in the identification of those repositories that meet the necessary standards.

⁶ <http://www.projectcounter.org/about.html>.

⁷ John Houghton and Peter Sheehan, “The Economic Impact of Enhanced Access to Research Findings,” CSES Working Paper no. 23, Centre for Strategic Economic Studies, July 2006. Available at <http://www.cfses.com/documents/wp23.pdf> (accessed Nov. 30. 2011).

(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 core of a public access policy should clearly be peer-reviewed journal articles, since this is a medium for which there are clear standards and expectations, whose use is familiar to most people, and for which the intellectual property rights situation is well-understood. As noted above, authors of these papers do not depend on royalty payments for an incentive, so participation by the funded researchers will be less problematic in this area than those in which commercial expectations are paramount. Other types of materials will have different conditions, and the IP rights can get very complicated in some cases. Educational objects -- digital objects created to teach a particular concept or point -- would be another type of material to consider, but the standards of peer-review are not yet clearly in place for those works, and the IP rights can be very unclear. So while these educational objects, if funded by federal research money, might well be the subject of a separate mandate crafted for the specific conditions, it is clear that journal articles should be the starting point. The complexity of other materials should not be allowed to delay the implementation of a mandate for public access to funded articles across federal agencies.

(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?

It is vital to understand that any embargo necessarily involves a tradeoff that diminishes the possibilities for innovation and growth that a mandate is trying to realize. In economic terms, each day of an embargo period represents lost opportunity costs that decrease the return on investment that is possible for the policy. For this reason, an embargo period is always a compromise, and they should be as short as possible.

Because the conditions and traditions of individual disciplines vary a good deal, it is probably best to allow authors themselves to determine appropriate embargo periods, or to permit immediate public access. In any case, this option should be preserved only with defined limits. For example, in order to avoid excessive negative impact on both the research community and the patient community, embargoes in the biomedical sciences should probably never exceed six months. We note that there is a strong sense among those who study global health issues that the current 12 month embargo allowed by the National Institutes of Health is too long, as indicated in a report funded by the Institutes of Medicine of the National Academies on "The U.S. Commitment to Global Health:"

Several prominent medical research funders have made open access a condition of grant support. The European Research Council, a funding body set up by the European Union to promote research in the region, has also put forward an open access policy requiring its grantees to post all publications to a research repository within 6 months of publication. This marked the first EU-wide open access policy and ERC has stated that it has interest in shortening the 6 month window period in the future. The Wellcome Trust requires submission of scientific publications resulting from its grants into UK PubMed Central within six months of the publication date and even provides funding for the upfront fees associated with publishing in such outlets. Grantees of the Howard Hughes Medical Institute also face a similar requirement to deposit publications in PubMed within six months of the publication date. By contrast, NIH's Public Access Policy remains at twelve months, twice the embargo period accepted by other leading funding agencies.⁸

It is worth noting that many journals, in a variety of disciplines, have adopted embargoes of less than a year, as publishers discover that the loss of revenues that they feared have simply not materialized.⁹

As far as empirical arguments are concerned, we should begin with the recognition that the evidence supporting the value of public access is plentiful and growing all the time. The burden of proof for the need for embargoes should therefore rest on those who would impose them. For example, in spite of repeated assertions and worries, there is no evidence at all that libraries have cancelled journal subscriptions because of publicly accessible funded research articles. Before such an assertion is allowed to impose an unnecessary burden on public access, evidence should be demanded of actual harm. Where no such evidence can be found, immediate public access should be the norm, because this is the best way to foster innovation, competition, economic growth and scientific progress.

These comments are submitted on behalf of the Duke University Libraries by:

Deborah Jakubs
Rita DiGiallonardo Holloway University Librarian & Vice Provost for Library Affairs

Kevin L. Smith
Director of Scholarly Communications

Paolo Mangiafico
Director of Digital Information Strategy

⁸ Committee on the U.S. Commitment to Global Health, Board on Global Health, "The U.S. commitment to global health : recommendations for the public and private sectors ," p.255. Available at http://www.nap.edu/catalog.php?record_id=12642 (Accessed Dec. 7, 2011).

⁹ A list of journals with the embargoes from Highwire Press illustrates that almost no journals are embargoed for more than one year, and some for substantially shorter periods. See <http://highwire.stanford.edu/lists/freeart.dtl>.

Subject: ACRA; Pub. L. 111-358)

Date: December 13, 2011 2:22:18 PM EST

ACRA; [Pub. L. 111-358](#))

[11/04/2011](#)

Dana L. Roth
Pasadena, CA

In order to achieve long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally funded scientific research, it is essential that the research be published in research journals published by recognized societies (e.g. American Physical Society, American Chemical Society, Royal Society of Chemistry, etc.)

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

Submitted by
James L. Mullins, PhD
Dean of Libraries and Professor
Purdue University
December 2011

Background

Access to scholarly literature and other outputs of research is of great concern to Purdue University and the larger academic community. At Purdue, we strongly support White House action in support of long-term stewardship and broad public access to the scholarship resulting from federally funded research.

Purdue University is a coeducational, state-assisted system in Indiana. Founded in 1869 and named after benefactor John Purdue, we are one of the nation's leading research institutions with a reputation for excellent and affordable education. Our West Lafayette campus offers more than 200 majors for undergraduates, over 70 master's and doctoral programs, and professional degrees in pharmacy and veterinary medicine. The University expended \$472.7 million in support of research system-wide in 2006–07, using funds received from state and federal governments, industry, foundations, and individual donors. The West Lafayette campus boasts more than 400 research laboratories and 116 University-approved research centers and institutes. Purdue's Office of Engagement matches Purdue expertise and resources with the expressed needs of business, governmental agencies, communities, schools, and individuals around the world, with a specific focus on improving the economic prosperity and quality of life for Indiana's residents. (<http://www.purdue.edu>)

At Purdue, we believe that the commercial academic journal pricing situation makes the current system of scholarly communication unsustainable. Scholarly journals from commercial publishers are too numerous and increase in cost at a rate that outpaces inflation and the consumer price index. For several years academic libraries have been working together and with university presses to explore alternative means of publishing scholarly content, with fundamental goals to reduce costs and increase access. In 2006 we developed Purdue e-Pubs, Purdue's digital repository for Purdue-based scholarship. e-Pubs, like other open access repositories, into which journal articles and other scholarly content may be archived, comprise an important part of the Open Access equation, particularly for libraries and institutions looking to preserve and provide access to scholarly outputs originating from the home institution. Open Access repositories are enabled in large part by the willingness of authors to negotiate terms of publication in ways that permit deposit of publication into a digital archive, often after an appropriate "embargo" period during which publishers can recoup their investments.

With respect to the specific questions asked in the Request for Information, we recommend the following:

(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?

Federal agencies such as the Institute for Museum and Library Services have supported a number of important initiatives to grow existing and new markets, and to increase the productivity of the scientific enterprise. Continuing support for IMLS is essential. Two recent examples of relevant projects in which Purdue University has been involved demonstrate the positive effect that IMLS funding is having in increasing efficiency and improving productivity.

“Library Publishing Services: Strategies for Success” (<http://wp.sparc.arl.org/lps/>) was a research project conducted in 2010 and 2011 that brought national leaders involved in developing library-based publishing programs together for the first time. A survey conducted as part of the project revealed that almost 80% of Association of Research Libraries member institutions are either developing or implementing publishing services, and that libraries are increasingly taking leadership roles in publishing scientific research, particularly in professionalizing the dissemination of “gray literature” such as technical reports and conference proceedings that have previously been difficult to access.

Databib (<http://databib.lib.purdue.edu/about.html>) is at an earlier stage in development but will offer an online, community-driven, annotated bibliography of research data repositories which will, among other benefits, increase efficient use of existing repository infrastructure for preserving and disseminating scientific research and better leverage past federal and private funding.

(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?

Intellectual property interests include copyright, trademarks, and patents. For the purpose of this RFI we will address the copyright issue with the understanding that trademarks and patents involve different ownership issues that are specific to institutions.

This public access policy should not be restricted solely to scientific research, but should extend to include all federally funded research. Scholarly publications can range from journal articles to books to book chapters to conference proceedings, etc. For the purpose of this response our comments relate specifically to scholarly journal articles.

A variety of approaches are possible, from the most radical -- of all publicly funded research resulting in a scholarly journal article being considered a government work and immediately accessible, with no copyright attached to it, i.e.: public domain, to the more moderate approach where all funded research would be deposited into a freely accessible repository, with a reasonable embargo period such as six months to one year but no more than that, whether at a university, government agency, etc. Would it be reasonable that Library of Congress as our national library be one of the repositories for publicly funded research?

Fundamentally, a policy not recommended in any form would be to allow a copyright holder, whether publisher or author, to restrict access to the results of federally funded research for the duration of copyright.

(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?

For long term stewardship, the Library of Congress could play a role, in that the Library of Congress's mission is "to support the Congress in fulfilling its constitutional duties and to further the progress of knowledge and creativity for the benefit of the American people." The mission of LC's Library Services unit "is to develop qualitatively the Library's universal collections, which document the history and further the creativity of the American people and which record and contribute to the advancement of civilization and knowledge throughout the world, and to acquire, organize, provide access to, maintain, secure, and preserve these collections."
(<http://www.loc.gov/about/>)

Pros and con exist for both centralized and decentralized approaches. With a decentralized approach there may be more quality control of content and organization or arrangement of content. With a centralized approach, the one-stop-shopping model to which most Internet searchers are accustomed to when searching the Internet may be more achievable.

(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?

HathiTrust, <http://www.hathitrust.org/>, a partnership of major research institutions and libraries could be considered one model of this type. HathiTrust states that it works to "ensure that the cultural record is preserved and accessible long into the future." Purdue is a member of HathiTrust and has made accessible through HathiTrust many Purdue-related publications.

The exponential increase in the amount of data generated by research in the digital age poses challenges to both publishers and libraries, and there are opportunities for joint initiatives that harness the expertise libraries have built in the organization and preservation of data to publishers' ability to effectively drive usage of it. Purdue University Libraries is one of the founding partners of the DataCite initiative, <http://www.datacite.org> that allocates stable digital identifiers to data so that it can be cited. The University has also been a leader in establishing a sustainable digital data repository, the Purdue University Research Repository (<http://research.hub.purdue.edu/>), the infrastructure behind which could be leveraged to create subject-specific collections of data in collaboration with specialist, especially learned society and university press, publishers.

There are as yet relatively few examples of this kind of library-publisher partnership to support a

disciplinary community, perhaps because society publishers in particular usually draw their members from multiple institutions and find it difficult to partner with one of them. However the work of the Digital Research and Curation Center (DRCC) at the Johns Hopkins University Sheridan Libraries in partnership with the American Astronomical Society (AAS) and its publishing partner, the Institute of Physics (IoP), has shown how digital archives, electronic publishing systems, and research communities can work together in a unified and efficient system.

(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?

Steps that can be taken to encourage search, discovery, and analysis include standardization of metadata for discovery; open sharing of core metadata; federal program or incentives to openly share citation data; standards development for the sharing of embedded and supplementary materials to enhance capacity for analysis. This would include usage of standard identifiers (e.g., DOIs) for such materials, but also the development of methods for embedding structured data formats from disciplines into electronic publications; and mandates for sharing of data (defined broadly) with publishers/societies more proactive about mandating data citation.

Minimum core metadata include metadata elements necessary for citation (see DataCite Metadata Kernel for an example) plus abstracts. Formal subject analysis would be nice, but is difficult and expensive to do in a cross-disciplinary manner. [The core metadata should include authors, title, date, source publication citation info \(ideally with persistent URL to access\) and keywords from controlled vocabularies \(not just discipline-specific, but also broader subject categories to be more widely shared across disciplines\) as well as source of funding \(agency\), this would allow linking to federal science funding.](#) While open metadata is important for search and discovery, open availability of full-text is necessary to allow disciplines to more fully engage the literature in their research (enables not only reading, but text mining apps).

Federal agencies can 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 by standards development and core metadata, as described above; mandates that publishers share core metadata (including abstracts) openly (perhaps exposing them using OAI-PMH, or a similar protocol); and stronger mandates for sharing of full-text and data. [A wide variety of metadata standards are emerging, with Dublin Core, at least at this time, offering the most flexibility in terms of use by a wide variety of systems.](#)

(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?

While there are pros and cons to both centralized and decentralized approaches, the distributed and collaborative nature of 21st century scientific research and the networked character of digital information distribution suggest that no one organization or agency can bear the burden of preserving and delivering the results of taxpayer-funded research.

In this environment, initiatives that encourage the development of a range of repositories organized at either institutional or disciplinary levels and the support of a diverse publishing ecology, not dominated by a few large players, will maximize the benefit of public access policies. Existing disciplinary repositories such as arXiv (<http://arxiv.org/>) suggest how large “big science” communities can be served, although they struggle to find sustainability models. There is a need also to support “small science” in a way that is sensitive to the needs of a particular discipline, and some innovative partnerships between libraries and information technology services on campuses are starting to emerge to do this. Purdue University now supports over thirty “hubs” hosted on the HUBzero “platform for scientific collaboration” which was developed with substantial support from the National Science Foundation (<http://hubzero.org>).

In practical terms, federal agencies that fund science need to ensure that sufficient funds are included in grants to sustain the network of disciplinarily-sensitive repository services, need to continue to support the creation of standards for repositories and new forms of digital publication, and need to encourage the development of directories and “refer-itory” tools that help authors and users locate the partners best suited to preserve and disseminate their research findings.

(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 believe that all publications resulting from federally funded research should be made available freely to the public.

At Purdue, e-Pubs contains many publications that are not articles in scholarly journals, but provide much beneficial research to the public. These including technical reports, white papers, and conference proceedings. Use of these non-traditional research publications in Purdue’s e-Pubs is extremely high, as illustrated in a recent paper by Purdue University scholars accepted by the Transportation Research Board of the National Academies for its 2012 meeting, http://docs.lib.purdue.edu/lib_research/146/. The paper also suggests that investment in updated systems and infrastructure more well-g geared to digital distribution of so-called gray literature produced with State as well as Federal funding is essential and offers a model for innovative partnerships between public agencies and university libraries.

As the division between book and journal publications reduces, the exemption of book chapters from federal funding requirements for Open Access becomes less meaningful. It is true that book publication operates according to different norms and has traditionally involved more investment by the publisher, but most of the added value lies in the apparatus surrounding the articles (such as the index, introduction, part openers, consolidated bibliography, and organization) rather than the chapters themselves which are analogous to journal articles.

In general, the digital environment is removing traditional distinctions between types of publications that were developed in the print age. Any policy that discriminates between

publication formats based on their physical characteristics will soon become outmoded in the virtual environment and may be subject to abuse if less scrupulous commercial entities manipulate the formats in which research is clothed to impede access to publicly-funded scholarship.

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

The appropriate embargo period after publication is dependent on disciplinary norms but we believe that this need never be more than 12 months. A system that relies on calculating the “half life” normal to each discipline and sub-discipline introduces unnecessary complexity and would not be scalable. The empirical evidence is divided and in any case is based on snapshots of research practice that are unlikely to carry forward in a digital environment where the normal practices of research, and understandings of what is “current” and what is “archival,” are changing at an astonishing speed.

Other types of publications resulting from federally funded research, beyond traditional journal articles, should also be mandated for open access for the public because if publishers believe they are losing revenue they may make business decisions to do away with traditional scholarly journals and instead publish under the guise of another format (i.e. book) to avoid the mandate of deposit.

Subject: Office of Science and Technology RFI

Date: December 15, 2011 10:36:58 AM EST

Comments:

Katie Porretta
National Safety Council,

As the managing editor of the Journal of Safety Research, I have 21+ years of experience with a peer-reviewed publication and, with that, federally funded research. Today's technology offers so much potential for easy international coverage for researchers to share their research, and translate this knowledge into action.

As such:

(a) Authors, editors, publishers are committed to providing the broadest possible access to our publications. Our goal is to expand access to the latest breakthroughs in the scientific research and developments in academic thought.

(b) The main purpose of a scholarly journal is to report on original research or experimentation in order to make such information available to the global research community.

(c) It requires considerable investment to maintain the highest standards for peer-reviewed scientific publication and sustainable mechanisms are required to enable this system to cooperate. This is threatened by access policies that do not take these costs into account, and it is critically important that any new policies do not damage the publishing institutions on which the Federal Government and science depend.

(d) Overall, access to research articles is good, particularly in the United States, where 97% of researchers are happy with access to journal articles.

(e) Publishers are committed to indentifying any gaps in access and have developed a number of mechanisms to close gaps in sustainable ways. This includes initiatives to broaden access for researchers in developing countries, patients, and the public

(f) Any effort by government to establish policies for peer-reviewed research should be done in consultation with all affected stakeholders, ensuring that such policies do not undermine the sustainability of the peer review publishing system which is necessary to ensure the quality and integrity of scientific research.

Thank you for your consideration.
Katie Porretta

Subject: Response to RFI

Date: December 15, 2011 6:51:02 PM EST

Public access to federally funded research is an important issue. However, the issue of "access" to research and published manuscripts that may or may not be directly related to federally funded research is a one that is adequately and appropriately addressed in the current marketplace. Indeed, authors and researchers have access to a wide collection of peer reviewed publication venues in the current market (including open access). As such, authors select journals based on their desired audience and any policy that limits the ability of researchers to select a journal (or dissemination format) arguably infringes on academic freedom. In my own experiences, I have published in open access journals (highly competitive journals) and done so with the specific intent of reaching a "different" audience. I also--and more regularly--publish in venues associated with global content providers such as Elsevier, Taylor & Francis, Wiley-Blackwell, and so on. I select these journals as they have sophisticated infrastructures that ensure the research reaches a target audience. Additionally, the quality of their product and their review processes are outstanding. Unfortunately, open access/public access venues seldom have these infrastructures and/or the resources necessary to invest in such infrastructures. As such, the major publishers play a critical role in the actual dissemination of research and their resources/infrastructures (which include partnerships with university and public libraries) ensure visibility. Similarly, major publishers continue to innovate and push the boundaries of research dissemination and electronic publishing.

Additionally, the boundary between "research" and "publication" is an increasingly blurry one insofar as research is often highly collaborative and often does not conform to research performed within a specific "time frame" outlined in a single grant or funded project. Insofar as knowledge is incremental and accrues over time, determining what portion of article X is a by-product of grant Z is highly problematic.

In the end, any effort to limit the markets associated with peer reviewed publications--while cloaked under the noble guise of public access--will harm knowledge industries; create confusion over definitions of "funded" v. "not funded" research; reduce the visibility of important research; and arguably stunt innovation across the publishing sector (i.e., limit the overall competitiveness of content providers/publishers). If authors want a "public option", they will create it and currently have access to such markets. Let the markets work.

Jay D. Gatrell, Ph.D.

Dean, College of Graduate and Professional Studies;

Professor of Geography; and

Editor, Applied Geography

From: Paul Krombholz

Subject: Availability of taxpayer-funded research papers

Date: December 16, 2011 3:04:02 AM EST

I feel very strongly that, if the taxpayers (that is all U.S. Citizens) funded the research, the resulting papers should be available on the internet for download at no cost.

When I was a graduate student at the University of Wisconsin, I browsed happily in the library for research articles. There was never an article that I couldn't get. A lot of my library research was based on my research project, but by no means all of it. Often I looked up articles just because they were interesting. That kind of free-browsing has supplied me with a lot of information that I use in teaching to this day, more than 40 years later.

Now I am teaching at a small private college that can not afford more than a few journals in its library. Journal costs have gone up even faster than medical costs and many libraries, even those of large state institutions can no longer afford to subscribe to everything they used to be able to afford. My school can not afford expensive reference journals such as Science Citation Index or Bio Abstracts, but Google Scholar is getting pretty good as a reference site. However, when I find an interesting article, all I can get is the abstract. If I want to see the article I have to pay \$20 to \$40 to download it. I can't do that.

I have read recently that there is a revolt at the University of California system over the costs of Nature Publishing Group journals being raised in one step by 400%. (<http://chronicle.com/article/U-of-California-Tries-Just/65823/>) Science publishers charge the author heavy page costs, and then they charge the libraries prohibitively high subscription rates. Science publishers have become the troll under the bridge.

There is no competition between publishers; each publisher is a monopoly selling access to its journals at prices as high as the market can bear. This is analogous to the government paying for the extracting and refining of oil, and then giving it to a single, unregulated monopoly to retail it.

There is a lot of hand-wringing these days about how badly educated American citizens are in the sciences, about how easily they are being misled by anti-science propaganda denying global warming, denying evolution, denying vaccination, promoting alternative medicine, promoting astrology, promoting witchcraft, alien visitation and any kind of woo woo nonsense. The Bush Administration prevented the launch of a satellite that would accurately measure the heat budget of the earth and answer the question whether or not global warming is occurring. Congress dismantled the Office of Technology Assessment about 10 years ago because they didn't want any advice from scientists any more. There are strong and well-funded anti-science movements in this country. Perhaps if our citizens had more access to science, they might vote for more sensible people to represent them.

From: Paul S. Brookes
Subject: Response to RFI for public comment
Date: December 16, 2011 9:26:34 AM EST

(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?

Yes! Encouraging journals to make their back-catalogs available (in many cases these articles are more than 50 years old and therefore standard copyright coverage has lapsed). A significant barrier to research productivity is obtaining scientific journal articles that are only in print form and not online. Some journals are doing a great job of making their archives available. Others not so.

(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 disagree with the term "intellectual property rights" applied to the publisher. The publisher has no intellectual property rights because they did not expend any intellectual energy; the scientists did! The scientists and the funding agencies own the rights to intellectual property. All the journals own is copyright over the presentation of the material. A big part of the problem involves pricing... I don't think anyone has a problem with the journals charging for access, since they have to invest significant resources in making the material available online. The problem is the actual price point. For many scientific journals, access to a single article is \$25-45. This is daylight robbery! A good policy would be to **set limits on the prices journals are allowed to charge for access to articles (say \$2 per article)**. They would actually make more money, because people would not try to bypass the system as they do now (getting the paper from a friend at another University whose library has access, or paying for interlibrary loan transfer).

(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?

NCBI is great. PubMed is fantastic. Please please keep on funding what they do. It is amazing.

(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 answer.

(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?

Big problem is **scientist IDs**. There are sometimes many scientists with the same name, and this makes searching very difficult. Same thing when someone moves institution - their contact email on a paper might be out of date. There really needs to be a centralized ID system, where everyone is assigned an ID # when they enter science (say, in graduate school), and it stays with them for life. Then, when I read a paper by someone from 10 years ago, I can find who they are, where they are now, and their contact details, just by clicking on their universal scientist ID #. I believe EMBO is starting to do this now, and Thompson Reuters also (www.researcherid.com)

(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?

ETHICS ETHICS ETHICS! I'll say it again... ETHICS! There is an absolutely scandalous amount of dodgy data in the peer-reviewed literature, from faked western blots to made up graphs and manipulated microscope images. ALL of this stuff is very easy to detect using simple software. ALL journals should be running every paper they receive through plagiarism software. ALL figures submitted should be subject to forensics to determine if they have been "altered". Just this last week, I reported a case to the NIH Office of Research Integrity (ORI), wherein a prominent cancer researcher has over 120 examples of faked data in more than 50 publications! This person has a program project grant and has been sucking up millions of dollars from the research funding pool for decades. Clearly, simple **ethics enforcement by journals** would have stopped this person in their tracks. It should not take the (unpaid) actions of a diligent scientist such as myself, to detect this sort of thing. Please, please, **increase funding to the ORI**, so they have the resources needed to detect and eliminate scientific fraud. Simple enforcement measures would get the journals to comply.... e.g. 1 point knocked off the impact factor, for every paper that gets retracted.

(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. Often these types of material are worked on outside of normal "office hours", so it is difficult to claim they were actually federally funded.

(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?

Current policy of 1 year embargo seems to be working well. Don't see any reason to change it.

Other comments... Scientific Journal Publishing is one of the most profitable enterprises in the world today. In the same manner that the administration has made it clear they do not like profiteering by the pharmaceutical industry, I think it would be reasonable to impose profit limits on the big publishing houses (Elsevier, Nature Publishing Group). Finally, **a distinction MUST be made between journals that are for-profit, and those which subsidize a scientific society**. Many scientific societies use revenues from their society journal to fund important educational and other activities. This is very different from publishing houses which simply keep all the revenue and plow it back into advertising and marketing and other initiatives. A 2 tier system, whereby society-backed journals would be allowed some leeway in pricing models, while private journals are held more accountable, would be great.

END OF RESPONSES

Subject: Response to RFI

Date: December 16, 2011 9:39:29 AM EST

To whom it may concern:

It is my opinion that the current system of open access/public access to published research findings funded by the US gov't,

is working well and there is no need for changing the system or for additional requirements upon publishers of science, technology,

and medical research findings.

Wilbur H. Campbell

--

Wilbur H. (Bill) Campbell, Ph.D.
The Nitrate Elimination Co., Inc.

publicaccess@ostp.gov

16th December 2011

**A response from The Publishers Association (UK)
to the Office of Science and Technology Policy Request for Information:
Public Access to Peer-Reviewed Scholarly Publications Resulting from
Federally Funded Research**

The Publishers Association is the representative body for the book, journal, audio and electronic publishers in the UK.

The PA welcomes this opportunity 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.

In framing our response, we feel it is important to be clear about the meaning we attribute to some of the terms we use or are used in the RFI. This is our understanding:

- **Research results** are not the same as scholarly publications. They are taken here to mean the reports by grantees to funding agencies.
- **Scholarly publications** are the final peer-reviewed, archival-quality outputs from research. They are the subject of considerable investment by publishers and those with whom we work.
- **Journals** are an aggregation of scholarly publications employed to disseminate research outputs. They represent quality assurance, provenance and brand recognition for the research communities that they serve.
- **Scholarly publishers** may be learned societies publishing journals in their field, or learned societies working with commercial publishers who publish their journals under contract, or commercial publishers publishing journals that they own or have developed on behalf of emerging research communities.
- **Access model** is a term used to encompass the wide variety of business models employed by publishers to sustain scholarly publishing. These range from 'toll access', funded by consumer-pays subscription models, to producer-pays 'Gold open access' models, including numerous hybrids in between such as 'delayed access' whereby the publication is available on subscription during an embargo period then is made open access thereafter.
- **Version of Record (VoR)** is the final peer-reviewed archival quality version of a scholarly publication available exclusively from a scholarly publisher.

We also wish to make clear the position from which we as scholarly publishers offer our observations and recommendations. Our position can be summarised as follows:

- As scholarly publishers we are committed to the widest possible dissemination of and access to the scholarly publications that we publish. We support any and all sustainable access models that ensure the integrity and permanence of the scholarly record.
- Taxpayers fund research through the US federal agencies, but they do not fund the processes that produce the peer-reviewed scholarly publications that result from that research.
- Taxpayers, the government, and the federal agencies have no de jure right of access to the Version of Record of a scholarly publication without compensation to those beyond the federally funded grantees who worked on it and the scholarly publisher whose investments made the VoR available for science.
- Public access does not mean free access or access free of cost. The costs of formal scholarly publication prior to any public access need to be recovered somewhere in the funding process.
- Taxpayers should have full access to the research results they have funded, namely the reports from grantees to the federal funding agencies. Grantees should also be required to deliver an abstract of their work in lay language that the agencies should make available to the public.
- Many activities working for the benefit of science, including the learned societies themselves, depend crucially on the revenues deriving from scholarly publishing and scholarly publishers. To undermine this revenue by appropriation of published outputs without compensation is to undermine the institutional network which supports the scientific academy.
- Scientists rely on certification through publication not just for dissemination of their results and for dialogue with their peers, but also for career metrics. To undermine the current system without an alternative in sight would undermine the sociological fabric of science.
- If it is to be public policy that the peer-reviewed published outcomes of federally funded research should be made available free at the point of use, it seems to us that there are two basic alternatives: either these agencies should provide adequate funds to cover the costs associated with one of the several variants to the producer-pays 'Gold open access' model; or the agency should license the Version of Record from the publisher for dissemination to the public.
- This is a global issue to which much thinking and research has already been applied. There remain many practical issues needing to be resolved, but scholarly publishers collectively remain ready to engage in all constructive dialogues and consultations.

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?

Growing existing and new markets is what scholarly publishers are all about. We work for the widest possible dissemination of the material in which we invest. We have every incentive to do so. Why should federal agencies be any better at this than scholarly publishers, whose whole purpose and mission is dedicated to this task?

Science is global, and scholarly publishers have a global reach. Through partnerships, journal development, and the application of technology, we are able to make available the peer reviewed scholarly publications that derive from the results of research throughout the physical and scientific world to all those who need it. Appropriating those scholarly publications to fulfil an obligation of US federal agencies to the US public that funds them is not only illegitimate, but undermines a system crucial to the health of global science.

Conversely, access to the research literature is not a constraint on science. Studies show that this comes low down on barriers to productivity. Research intensive institutions generally spend around 1% of their total budget on access to the literature.

The scientific enterprise depends crucially on stringent quality controls and reliable results. The publishing process is the principal means that delivers this control. To undermine or short-circuit that process in the name of public access potentially impacts the whole iterative basis of science itself and the quality standards on which science depends.

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?

Agencies should make funds available within (and after) research grants to fund the costs of scholarly publication. There are now many and various options available without the need to appropriate materials to which value has been added by publishers and those with whom we work in order to populate repositories managed by funding agencies.

These options include open access 'producer-pays' models, or alternatively the agencies could license access to the VoR of those scholarly publications relating to the research that they have funded from scholarly publishers

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?

Scholarly publishers are by definition involved in long-term stewardship. That is a fundamental part of our mission. Most journal archives originally published as print have now been digitised by their publishers so that the whole canon of scientific literature from 1665 onwards is available for researchers to consult. There is no evidence that the government needs to intervene to ensure that this task is completed.

Scholarly publishers already work together on common standards for interoperability, search and analysis, especially through CrossRef, an operation founded by publishers, which developed the DOI identifier that delivers seamless navigation across the scientific literature and which is also providing the basis for supplier-neutral analysis of usage and for open discoverability. As demonstrated by our commitment to DOIs and now prospectively the ORCID system for author and researcher attribution, we recognise that the scientific literature cannot operate via a series of proprietary databases. So we do not believe that further centralised repositories in the style of PMC are necessary or desirable to progress either the scientific enterprise or access to scholarly publications. Publishers compete in terms of services to the scientific community in order to attract the best authors to their journals. We should be supported and encouraged to invest in our platforms and to continue to deliver global-scale interoperability. We see no need for interventions by government agencies in these services.

Preservation in the print era was the preserve of library collections, but this task is beyond the means of all but the largest national libraries in the digital era, and the means to achieve this is still work in progress. Publishers of digital journal collections and archives are alert to our responsibilities in this regard and most now have multiple arrangements in place with national libraries and preservation initiatives such as Portico and CLOCKSS. It is neither necessary nor desirable for federal agencies to intervene in this work.

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?

We recognise that funding agencies currently find it difficult to track the scholarly publications that derive from the research that they have funded. Current metadata schema associated with published articles generally do not take account of the source that funded the research behind the article. Scholarly publishers through their channels of collaboration (CrossRef and the trade

associations) have this task in hand and expect to evolve a resolution through the CrossGrant initiative in 2012.

Once this work is complete, agencies will have access to the metadata for scholarly publications deriving from research they have funded. This metadata could be made available to the public allowing them to link to the article abstract on the publisher's site and thence to the Version of Record, enabled by whatever access model has been applied.

Scholarly publishers potentially have a valuable role to play in assisting access to research datasets. Significant amounts of data are already made available as supplements to scholarly publications. Several publishers are already linking articles to various datasets, and are willing to work with funders to develop further identifiers that can link articles to primary datasets.

Publishers are also engaged with the ORCID project to associate a unique identifier with each published researcher.

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?

A key development for cross-disciplinary research will be the capacity to mine content across large datasets of scholarly publications. Pilot work is already in hand with several scholarly publishers to advance this capability and to enable the managed access that this will entail. From this work it is likely that the industry will evolve mining-friendly formats, a shared content mining approach, and commonly agreed permissions terms.

The Publishing Research Consortium has recently published a survey of current practice in this area, with some recommendations for how to enable more access for this purpose in future. See [Journal Article Mining](#) by Eefke Smit.

Other potential collaborations to extend access to and the functionality of scholarly publication datasets are being explored, including metadata linkages between publisher and funder sites, and access via rental models.

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?

Quality control in scholarly publications is vital for the progress of science, so version control of the outputs of science is equally vital. We believe that the Version of Record deriving from formal scholarly publication should be the record of research that is made available for access and consultation, and that policies operating on authors' peer-reviewed manuscripts introduce unnecessary confusion into the progress of science. Federal agencies should negotiate with scholarly publishers to make the VoR of scholarly publications deriving from research that they have funded as widely available as possible rather than mandating authors to deposit an earlier version into a Green repository.

The current system for dissemination of science is not failing. There is right now more access to more research by more scientists than ever before, deriving from innovations in digital technology and in the access models adopted by scholarly publishers. This process of organic evolution is accelerating as publishers strive to extend access through further innovations, including many variants on the producer-pays 'Gold open access' model. For the government to intervene in this process at a critical time may undermine the incentive to invest and stifle a healthy and vigorous movement working for the benefit of science.

We accept that institutional repositories of universities and research institutes have a range of purposes that drive their mission but we see no concrete economic argument for funding repositories primarily as a parallel system to that established for the Version of Record by scholarly publishers. Instead, research funders wanting to pursue an open access policy to the outputs of the research they have funded should fund the producer-pays Gold open access model in one of its variants.

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?

A public access policy could apply only to such other types of peer-reviewed scholarly publications that derive from the direct initiative of researchers funded by a federal agency and not from a commission by a scholarly publisher who conceptualised the project in the first place, otherwise the government would be appropriating content to which it had no conceptual relationship. Books published out of conference proceedings tend to be complex projects. They are often subject to significant rework following discussions with a publisher in order to produce a thorough, coherent, and consistent publication and the contributors and their different levels of engagement are often difficult to project manage. Conference proceedings are frequently amongst the lowest sales volumes in academic publishing as they need to have specific grounds for acquisition. Moves to mandate the availability of precursor versions of such material may quickly result in books based on conference proceedings being uneconomic for publishers, with the loss of many of the benefits to scholars and the academy noted in the introduction to this response.

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?

There cannot be an ‘appropriate’ embargo period that would satisfy all potential impacts on the scholarly publishing enterprise. Clearly the ‘half life’ of scholarly publications (meaning the shape of the access curve after publication) differs significantly by subject, where the half-life for say mathematics articles is much longer than for medical science. To impose arbitrarily such a blunt criterion on the whole of science has the potential to seriously undermine the viability of some scholarly publications.

Any embargo period is in effect a shortening of the copyright term that protects the investment made by publishers in intellectual property. If scholarly publishers are unable to achieve an adequate return on investment within the embargo period then they will cease to invest in such scholarly publications in future. Thus will the source of publications needed for public access be diminished and the incentive to invest in further innovations be stifled, negating the purpose of the policy. The dramatic recent improvements in access for researchers would also be threatened when or if publishers decide to withdraw further investment in the face of damaging public access policies.

A better policy would be either adequately to fund Gold open access, or to license access to the VoR from the publisher.

Ends

Subject: RFI on Public Access to Scholarly Publications

Date: December 16, 2011 4:27:33 PM EST

Dear Sirs:

Federal Register Volume 76, Number 214

November 4, 2011

Pages 68518-68520

From:

Nancy Stellwagen

Adjunct Professor, Research Scientist

University of Iowa

Iowa City, IA

I would like to comment on some of the questions raised by the Task Force.

Comment 1. I think that the current model of both open-access and publisher-financed publications meets the needs of many scientific researchers. In some ways, "open-access" is a misnomer; it really should be called "author-pay". Most open access publishers require the author (or his/her grant or home institution) to pay the costs of reviewing, editing and publishing the article on-line. However, in these days of a scarcity of funds (especially from the National Science Foundation and the National Institutes of Health), such money is often not available. I myself, working with a very minimal NSF grant, do not even consider publishing in an open-access journal; the money could be better used to buy chemicals or pay for instrument repairs. I realize that some people refer to the publisher-financed journals as "reader-pay". While this is true, the cost is spread over many readers and many libraries and many funding sources.

The analysis of information from the exploding number of on-line journals presents a serious problem for the future. Often, the articles published in some of these journals have been rejected by one or more publisher-financed journals. There are already web sites that purport to mine all publications on a given topic and collate the data, presumably without any attempt to assess the validity of the results. If researchers begin to rely on such summaries, the scientific enterprise will suffer.

One helpful step in improving access to the scientific literature would be for all publishers to allow free access to specific articles in their back-files, without opening the whole archive to data mining. Some, but not all, publishers do this already. However, it is possible that various university libraries must pay for this privilege.

Comment 3. The problem with centralized access to the scientific literature is the continual need to update the method of incorporating and accessing the articles in the archive, depending on hardware and software used at the moment. One advantage of paper copies kept in a central location is that the articles can be scanned at any time in the future and made available to the public. Already, with the rise of the internet and automated searching technologies is that information published before ~1970 is almost completely lost to the scientific community. Notable exceptions are the journals published by some of the scientific societies, which have been scanned back to the original volumes and made available to the public.

Comment 7. Peer-reviewed publications such as book chapters and conference symposia should be covered by public access policies. At the present time, book chapters almost never can be accessed without the investigator buying the book; most university libraries have almost stopped buying books to concentrate their limited funds on scholarly journals. As a result, many investigators (myself included) no longer write book chapters, because any new summaries or insights are lost to the scientific community.

Comment 8. To me, a one year delay before an article published in the open literature seems reasonable.

Subject: FW: Response to RFI

Date: December 18, 2011 11:02:58 PM EST

Name/Email

A. Townsend Peterson

Affiliation/Organization

University of Kansas

City, State

Lawrence, Kansas

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?

- I am interpreting this question to focus on what federal agencies might do that would encourage markets related to an open access archive that federal action might create. I take it as a given that access to academic scholarship should be open access, and that the current closed-access system must come to an end, as lack of such access is closing off crucial avenues of insight and progress in research. I take this latter point as a given—and hope that those considering these comments will as well.

- As regards markets that could be developed or encouraged as a consequence of an open access world, clearly the most important one would consist of a new generation of journals that adopt a ‘gold’ level of access (i.e., free and open access to all papers published). Federal agencies can encourage such an ‘industry’ by means of subsidy grants that would assist such journals in getting off the ground, and up to the level of being indexed and accorded an impact factor. Once this status has been attained, the subsidy is no longer necessary, and permanent progress has been made towards establishing open access frameworks for academic scholarship. Such grants would be crucial because breaking into the high-end journal market will otherwise prove challenging for many open access journals. This shift is important because currently each university sends multiple millions of dollars ‘out’ to commercial interests to pay

for journal access. These funds leave the academic system, and frequently also leave the country (Springer, Cambridge, Oxford, Elsevier, Blackwell, etc.). A future system, in which access considerations are dealt with separately from profit margins, could potentially keep those same funds within academia in the USA, and thereby assist in making academia more solvent and self-sufficient.

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?

- I would emphasize the need for individuals to retain key rights to the work that were their intellectual products. That is, the current commercial system depends on publishers taking scientists' rights to their intellectual products in exchange for the admittedly expensive process of publication and distribution. The simple step of reserving key rights for authors makes an enormous contribution to the future of science, and yet has little or no effect on the solvency of the publishing industry. Quite simply, to publish, publishers don't need all the rights that they currently 'require' to be transferred to them.

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 am deeply concerned about the stewardship being distributed across multiple private sources. I can see clearly that there would result situations of limited or inconvenient access, even under a federal 'open access' mandate. The centralized model being tested in PubMed Central is an excellent model, and I have heard of no major drawbacks to this system. If it is not broken, why replace this system?

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?

- I think that the current system is an example of a public (i.e., academics) – private (i.e., publishers) partnership that has not worked well. Quite simply, the publishers have counted on enormous profit margins from publishing academic scholarship, to the point that they have lost sight of the mission of the endeavor. Forcing publishers NOT to force academics to give up copyright to journal-published scholarship would be a key step in fixing this broken system, and would make possible significant improvements in stewardship of the results of federally funded research.

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 minimum core metadata are well summarized in the Dublin Core that has been developed in the library community; these data can be supplemented lightly to standardize the information provided upon submission across agencies and repositories. Most important is the full-text searching capacities that one sees, for example, in Google Scholar. I believe that such open searching of full text is well within reach technologically, and will not require much in the way of new development efforts.

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 example of PubMed Central has been excellent. It has provided access to much of the biomedical literature in recent years and at considerable savings of time and effort. It would seem that broadening this effort to all federally-funded research would yield parallel, and significant, benefits as well, at similarly low cost. In the long term, as described above, such changes will eventually permit an open access, free-of-cost system in which universities do not have to fork over millions of dollars yearly to support a commercial, for-profit system that benefits neither them, nor the general public that funds the research.

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?

- Journal articles are a convenient place to start, because they do not involve profit or payment to the authors. In the longer term, however, I would love to see similar efforts in the realm of text books—text books represent a significant cost burden for university students, often far beyond what is necessary and prudent. In fact, in many cases, professors require their own students to purchase their own books, which—in my view at least—constitutes a massive conflict of interest. The text book question will—obviously—require considerable thought and consideration and discussion.

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?

I think that no embargo is most appropriate, at least in my own field. I make this suggestion based on the need for fast and flexible access to the entire scientific literature. Research is a tough endeavor if one is to remain current and not waste effort. Embargo periods act simply to delay the progress of research, and make the entire enterprise less effective. To my eye, this brake on science efficiency has significant costs associated with it, which justifies the idea of no embargoes.



American Society for Nutrition
Excellence in Nutrition Research and Practice

December 19, 2011

Office of Science and Technology Policy
National Science and Technology Council's Task Force on Public Access to Scholarly
Publications
725 17th Street
Washington DC 20502

Via Email To: publicaccess@ostp.gov

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

Dear OSTP Task Force on Public Access to Scholarly Publications:

The American Society for Nutrition (ASN) appreciates the opportunity to provide information to the Office of Science and Technology Policy (OSTP) National Science and Technology Council's Task Force on Public Access to Scholarly Publications regarding public access to peer-reviewed scholarly publications resulting from federally funded research. Founded in 1928, ASN is a nonprofit scientific society with over 4,300 members in more than 75 countries working in academia, practice, government, and industry. ASN is dedicated to bringing together the world's top nutrition research scientists to advance our knowledge and application of nutrition. ASN publishes peer-reviewed research publications *The American Journal of Clinical Nutrition (AJCN)*, *The Journal of Nutrition (JN)*, and the 2010-launched review journal *Advances in Nutrition*. *AJCN* and *JN* are the two of the top peer-reviewed scientific journals in the areas of nutrition science and dietetics.

ASN supports the principle of increased public access to scientific information that stimulates innovation, and voluntarily has taken the following significant steps to accomplish this:

- Since 1997, the Society has included free access to the online journal collection as a membership benefit.
- ASN has offered free public access to articles 12 months after publication on its website since 2000.
- ASN's entire journal collection, including over 110 years of archival content, has been online through Stanford University's High Wire Press since 2006. Approximately 98% of online journal content is freely accessible to both subscribers and non-subscribers.
- The content of all ASN journals is indexed in the National Library of Medicine PubMed/Medline database, and articles resulting from National Institutes of Health (NIH)-funded research are accessible from the online NIH open access repository PubMedCentral after an embargo period of 12 months.

9650 Rockville Pike | Bethesda, MD 20814

T: 301.634.7050 | F: 301.634.7892

info@nutrition.org | www.nutrition.org

ASN responses to select questions from 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?

Certain changes to the existing open access system may harm publishers, and therefore the U.S. economy. Any money and energy that would be put towards federal efforts to archive peer-reviewed publications, including those that result from federally funded scientific research, and to develop interoperability among multiple repositories would duplicate the archival efforts of many publishers.

ASN does not consider it advisable for the federal government to duplicate existing databases of published scientific information by establishing separate repositories for this information at multiple federal agencies or to invest time and resources to ensure the interoperability of these new systems. All final versions of scientific articles are available on publisher websites; these articles can be and are made publicly accessible after a suitable embargo period. Current search engines, such as Google Scholar, allow ready identification of and access to research articles published in scientific journals.

Federal efforts and funds could be better spent providing funding for and advancing new research, expanding efforts to translate research findings for the general public, specifically patients, and finding ways to help the U.S. public use this vast resource of scientific information to lead healthier lives.

- (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?

Typically the intellectual property right of copyright is transferred from authors to publishers, who make content freely available via their publisher platform and other databases 12 months after publication. Publishers and editors provide added

value such as peer review, editing and formatting to improve clarity, accuracy, readability, and discoverability of published research findings, and they ensure that content is archived on safe, reliable, and multiple sites. Copyright allows publishers and editors to produce journals and other publications to share federally funded and other research findings with a worldwide audience in both print and online versions which are indexed via multiple search engines and databases to enable the broadest access possible. Although ASN holds copyright, ASN provides authors with generous rights to reuse their material. ASN supports policies that allow publishers to retain copyright while giving privileges to authors. ASN also supports clearly identifying copyright holders in open access databases such as PubMedCentral. ASN does not support blanket requirements in grant contracts which have the potential to deny authors and publishers the benefits of their copyrights, such as how and in what form their works are distributed.

- (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?

ASN supports decentralized approaches to managing public access to peer-reviewed scholarly publications that result from federally funded research. To avoid significant and unnecessary costs for the government, any potential federal repository should link to the published research article on the publisher's site rather than post a duplicate copy in a federal repository. By developing and supporting search engines such as PubMed, the Federal government has already contributed greatly to the discoverability and accessibility of published research findings. ASN opposes publication of multiple versions of the same manuscript as this will confuse, and in some cases even corrupt, the scientific record. To minimize administrative tasks for authors, publishers, and the government, papers should be posted to and made publicly available via a single site such as a publisher's platform, rather than multiple repositories for different agencies or disciplines. We urge the government to work with publishers to provide public access directly from the article of record in the journal by providing links back to the content. In addition, the Federal government should not duplicate existing efforts supported by publishers and others to archive peer-reviewed scholarly publications, including those that result from federally funded research.

- (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 existing models of collaborative efforts amongst librarians, publishers, commercial entities, and others to maintain online archives of peer-reviewed scholarly publications, including those that result from federally funded research. Some examples of existing archives include LOCKSS (Lots of Copies Keep Stuff Safe); CLOCKSS (Controlled LOCKSS); Portico; and JSTOR; as well as others. Federal support for the archiving efforts already underway by public-private partnerships is critically important to maintain existing publisher archives. Federal support could also encourage innovation in accessibility and interoperability of these archives, 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?

ASN provides metadata to various journal content aggregators and databases, such as PubMed, to ensure searchability and accessibility of open access materials. ASN believes journal content aggregators are best suited to answer this 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?

Federal agencies can help to optimize search capabilities to allow the public to better access and understand federally funded research findings. Federal agency assistance in increasing public awareness and understanding of these findings will maximize the benefit of public access policies to U.S. taxpayers and their investment in the peer-reviewed literature.

- (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?

ASN does not believe that other types of peer-reviewed publications resulting from federally funded research, such as book chapters and conference proceedings, should be covered by public access policies. Typically these types of peer-reviewed publications do not contain original research that has not been or will not be published in a journal article already covered by public access policies. ASN supports covering book chapters and conference proceedings with public access policies only when they share original research findings that have not been or will not be published in open access journal articles.

- (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?

ASN believes that the appropriate embargo period before the public is granted free access to the full content of peer-reviewed scholarly publications resulting from federally funded research should be no less than 12 months after publication. This is in line with NIH policy which stipulates that all NIH-funded research must be made publicly accessible 12 months after publication. A shorter embargo period in public access policies devalues journal subscriptions, and therefore subscription revenue which many publishers rely heavily on to support publishing operations, including the costs of collecting, reviewing, editing, composing, disseminating, and archiving manuscripts. Two 2006 articles, “Self-Archiving and Journal Subscriptions: Co-existence or Competition?” (Publishing Research Consortium, http://www.publishingresearch.net/self_archiving2.htm) and “ALPSP Survey of Librarians on Factors in Journal Cancellation” (The Association of Learned and Professional Society Publishers, alpsp.org), demonstrated that an embargo period of 6 months or less would increase the likelihood that librarians may decide to cancel a journal subscription. A shorter embargo period can also more easily compromise the business models of small publishers, including many not-for-profit publishers, who publish scientific journals on a bimonthly or quarterly basis only.

Certain changes to public access policy may undermine the value of subscriptions and threaten subscription revenue, thereby interfering with the ability of publishers to recover costs for the peer review and distribution of published research. If publishers are unable to cover basic operating costs and/or sustain the rigor of peer review, this could easily lead to fewer journals and, thus, fewer routes for publication of scientific research findings. Consideration should be given to the potential economic impacts of any public access initiative on publishers, particularly not-for-profit publishers, and the fundamental roles and services scientific societies provide to their membership and the scientific community at-large. When ASN has implemented policies to improve public access to the research it publishes, the economic impact of these new policies was carefully considered. Considerations that influence how soon free public access is economically feasible for a particular journal include revenue sources, production costs, utilization patterns, time needed for cost recovery, and frequency of publication. For example, publishing costs are on average \$3,500 to publish an article in *AJCN* or *JN*. Certain changes to public access policy could lead some journals to greatly increase author fees in an effort to recover lost revenue from subscriptions. If these costs are covered by grant funds, this would reduce the amount of funding available for direct support of scientific research and could ultimately lead to fewer scientists sharing research findings in a timely manner. Before implementing a broad public access policy it is important to consider these factors so as to not negatively impact scientists, not-for-profit publishers, and others who already provide some form of open access.

ASN appreciates the opportunity to provide comments regarding public access to peer-reviewed scholarly publications resulting from federally funded research to the OSTP National Science and Technology Council's Task Force on Public Access to Scholarly Publications. We look forward to continued dialogue on this important issue and the opportunity to assist OSTP and other federal agencies in any other way deemed appropriate. We urge OSTP to fully involve publishers in the implementation of any public access policy. Please contact Karen King, Vice President for Publications, at 301.634.7053 or kingk@nutrition.org if ASN may provide further assistance.

Sincerely,



Sharon M. Donovan, Ph.D., R.D.
President, American Society for Nutrition

Subject: Public Access to peer-reviewed scholarly publications. . .

Date: December 19, 2011 10:51:12 AM EST

As Co-Editor of the Journal of World Business , let me briefly respond to your queries.

1. Open access is the way that scientific knowledge can be shared with others. Each publisher maintains their own data base and that's correct.
2. Intellectual property rights are very difficult to maintain, especially as scholars from various countries have vastly different interpretations of what this means. We require that authors clearly state that their manuscript is submitted as an original piece of research and that that manuscript is not at another peer-reviewed journal. Plagiarism is rampant in parts of Asia and difficult to discover. If we do discover it, we send a notification to the author and the dean/president of the school notifying these officials of the violation.
3. Keep everything decentralized. Editors are excellent gatekeepers.
4. No comment
5. Unfortunately, the reward system at most research institutions is such that cross-disciplinary research, while hoped for, is rarely rewarded and therefore, not achieved.
6. No comment
7. NONE.
8. There should be no embargo.

Sincerely,

John Slocum
Co-Editor, Journal of World Business

Subject: US Office of Science & Technology Policy issued a Request for Information

Date: December 19, 2011 11:55:27 AM EST

Open Access is a fine concept that has been unevenly implemented.

While some of the open access journals have high standards, there is an increasing proliferation of online journals that are poorly edited and accept poor quality work that would not be accepted anywhere else.

While the financial cost of obtaining published work may in some cases be an obstacle to access, the cost to the science community of having a flood of published misinformation to wade through is perhaps a greater problem. Furthermore, for the scientific community as a whole, publication in open access journals has a financial cost that, in my experience, is far greater to authors than is the cost of publishing in journals that delay open access for a year.

Henry Jay Forman, Ph.D.
Professor of Biochemistry and Chemistry
University of California, Merced
Research Professor of Gerontology
University of Southern California

President-Elect, Society for Free Radical Biology & Medicine
Associate Editor for Review Articles, Free Radical Biology & Medicine
Member, San Joaquin Valley Air Quality Control District Governing Board

**Subject: RFI responses Public Access to Peer-Reviewed Scholarly Publications
Resulting From Federally Funded Research**

Date: December 19, 2011 1:11:33 PM EST

Authors, editors, publishers are committed to providing the broadest possible access to our publications. Our goal is to expand access to the latest breakthroughs in the scientific research and developments in academic thought.

- The main purpose of a scholarly journal is to report on original research or experimentation in order to make such information available to the global research community.
- It requires considerable investment to maintain the highest standards for peer-reviewed scientific publication and sustainable mechanisms are required to enable this system to cooperate.
- Overall, access to research articles is good, particularly in the United States, where 97% of researchers are happy with access to journal articles[1].
- Publishers are committed to indentifying any gaps in access and have developed a number of mechanisms to close gaps in sustainable ways. This includes initiatives to broaden access for researchers in developing countries, patients, and the public.
- Any effort by government to establish policies for peer-reviewed research should be done in consultation with all affected stakeholders, ensuring that such policies do not undermine the sustainability of the peer review publishing system, which is necessary to ensure the quality and integrity of scientific research.

Roland Baron
Editor-in Chief,
Bone

Subject: [Federal Register Volume 76, Number 214 (Friday, November 4, 2011)]

Date: December 19, 2011 1:35:39 PM EST

As the Editor-in-Chief of Neurobiology of Disease (Elsevier; 2010 Impact Factor 5.1), I am writing in response to the Request for Information.

First, in my role as a federally (and non-federally) funded researcher who has worked at several academic medical centers, I do not believe there is a significant problem in public access to research reports that result from federally funded research. I have never heard this complaint, except when someone wants access to an article in a very obscure journal. This personal impression is supported by a recent survey (Access vs. Importance, A global study assessing the importance of and ease of access to professional and academic information. Phase I Results. 2010. Publishing Research Consortium http://www.publishingresearch.net/documents/PRCAccessvsImportanceGlobalNov2010_000.pdf).

Second, in my role as a journal editor, I point out that the main purpose of a scholarly journal is to report on original research or experimentation in order to make such information available to the global research community. Authors, editors, publishers are committed to providing the broadest possible access to our publications. Our goal is to expand access to the latest breakthroughs in the scientific research and developments in academic thought. Further, one of the metrics of a successful journal is the rate at which its articles are cited. Clearly, our articles must be accessible for them to be read and cited. Thus, we have an interest in ensuring adequate accessibility.

From my experiences at Neurobiology of Disease and other journals and with other publishers, it is clear that publishers are committed to indentifying any gaps in access and have developed a number of mechanisms to close gaps in sustainable ways. This includes initiatives to broaden access for researchers in developing countries, patients, and the public.

Lastly, I point out that it requires considerable investment to maintain the highest standards for peer-reviewed scientific publication and sustainable mechanisms are required to enable this system to cooperate with current and future mandates. As such, I would urge that any effort by the government to establish policies for peer-reviewed research should be done in consultation and partnership with all affected stakeholders, ensuring that such policies do not undermine the sustainability of the peer review publishing system, which is necessary to maintain the quality and integrity of scientific research.

Respectfully,

J Timothy Greenamyre, MD, PhD
Professor & Vice-Chair of Neurology
Director, Pittsburgh Institute for Neurodegenerative Diseases
UPMC Endowed Chair & Chief, Movement Disorders
University of Pittsburgh

Name: Dr. Joan R. Giesecke

Dean of Libraries, University of Nebraska-Lincoln Libraries

Organization: University of Nebraska-Lincoln Libraries

City, State: Lincoln, Nebraska

Comment 1

The complete collection of articles resulting from publicly funded research should be made freely accessible, so that the public can fully use them – (i.e. text mine, data mine, compute on them, create derivative works) without commercial restriction. Enabling full reuse of these articles enables innovative individuals and companies to construct new services and new products on publicly funded content. Open Access allows more users to stay on top of cutting-edge ideas and generate new uses and application for this research. Faster commercialization spurs economic growth, creating new jobs across broad sectors of the economy from the biotech sector to agriculture to energy to publishing. This is key for the economic growth of Nebraska.

Further, open access to research articles is a critical driver of scientific innovation and productivity. It:

- Increases citations and follow-on research
- Promotes diversity in follow-on research
- Increases the pursuit of new research pathways
- Encourages faster application of research

It opens up vast, previously unobtainable new research pathways, making new connections possible.

Information from Houghton reports shows that the benefits of an open-access policy similar to that of the 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 archive); and of that figure, approximately 60% is estimated to accrue directly to the U.S. economy. An effective, government-wide public access policy can be implemented in cost-effective manner by leveraging existing infrastructure to minimize unneeded duplication of efforts; utilizing the investments already made by the NIH with the annual operation of PMC; and supplementing existing access points with additional manuscripts which incurs only small incremental costs.

Enabling full reuse means we can do more with less; we don't have to duplicate research to be able to build on results, and we can continue to extract value from our initial investment for years to come.

Comment 2

Public access policies can be successfully implemented by respecting and working within the current copyright framework. The NIH policy currently allows uses of articles currently provided for under "fair use." Greater utility is needed for potential scientific and commercial benefit of this information to be fully realized. A useful strategy to consider to balance the interest of all stakeholders would be to take a stepped approach: First, provide an appropriate period of embargoed access where current, fair-use only rights apply; and second, after the expiration of the embargo period, provide full reuse right to the public under an appropriate open access license.

Comment 3

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. To ensure this, any public-access policies that are developed must give the

federal government adequate rights to archive and distribute publicly funded articles. The federal government should, at a minimum, maintain an accessible, mirrored version of all content. Federal stewardship is cost-effective. NLM reports PMC costs less than 1/100th of one percent of HHH's operating budget to run. Simply providing government with a copy to put in a "dark archive" is not a viable solution; without regular access/use, archival veracity cannot be ensured. Library experiences have shown that regular access/use of digital materials is a crucial element in effective long-term preservation.

Comment 4

Publishers are one player that might be encouraged to participate in public/private partnership by providing approved repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles. However, none of the 50+ research funders who currently have public access policies are using publisher sites as the final archives. However, there are good examples of funders partnering with academic and research institutions in this role. Programs such as Hathi Trust provide a model for the deposit and preservation of digital information and provide an open access platform for this research.

Comment 5

No comment

Comment 6

For any public-access policy to be successful there must be consistency of requirements and mandates. Institutions often have researchers who hold grants from multiple agencies concurrently. Uniform requirements and procedures regarding deposit of peer-reviewed literature should be established across all funding agencies. Uniformity of deposit requirements will reduce the complexity and cost while, at the same time, increasing the rate of compliance. Effective implementation strategies for public-access policies can help maximize returns to taxpayers by ensuring that complete results are widely available in a timely manner. Policies can create opportunities to create/enhance productivity management tools for federal and internal reporting.

Comment 7

No comment

Comment 8

Immediate access is the ideal time to optimize scientific and commercial utility of information contained in these articles. However, to accommodate those journal publishers who continue to rely on subscription income, an author-determined embargo period of 0-12 months has proven effective across multiple disciplines. No data has been provided by any publisher that this embargo period (currently in use by NIH and numerous other funders around the world) has harmed them. Even publishers who previously expressed concern that opening access to back content to result in loss of revenue have now changed practices. All of these market conditions regularly contribute to journal cancellations and must be accounted for so that the effect of an embargo period can be adequately isolated.

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

Date: December 19, 2011 5:56:35 PM EST

Name/Email Joshua L. Rosenbloom / jrosenbloom@ku.edu

Affiliation/Organization University of Kansas

City, State Lawrence KS

Comment 1

I interpret this question to relate to the broader economic impacts of Federal policies that affect the availability of peer-reviewed publications that result from federally funded scientific research. Current restrictions on access to these publications limits their value to other academic and scientific researchers for whom they are a key input in further knowledge creation. As such it is desirable to make these results more readily available to the public. In addition such availability will likely stimulate new online services that provide ways of combining, accessing and analyzing such content. This is illustrated by the ways in which for example services combining GIS map data and other information have been developed, and in the value that out of copyright books scanned by Google has created.

Agency policies that require free access to scholarship produced with federal support appears generally to be a desirable characteristic, at least on the margin.

Comment 2

Development of common and commonly accepted templates for specifying intellectual property rights perhaps through a collaborative standard setting process would be helpful. Such templates would simplify the process of assigning copyright, for example, by insuring that publishers all adhere to the same guidelines and processes and that there is uniformity in the definition of Intellectual Property rights.

Comment 3

Pros of centralized management would include consistent treatment and ease of access. Cons are the likely reduction in innovation that would come from creating a single entity to manage these publications. Competition would tend to encourage improvements in interface, access, storage and other desirable characteristics. There are perhaps also issues of redundancy that would be better served by a more decentralized system.

Comment 4

University sponsored repositories, Google Scholar, and a number of ventures sponsored by scholarly societies could be interpreted as possible models of the provision of access to scholarly information.

More generally, the Internet, the federal highway system and most other modern infrastructure provide models that may be relevant to the issues at hand here. Specifically, the development of electronic access radically alters the costs of scholarly communication and creates a situation in which the processes of scholarly communication – viewed in total – are ripe for changes. The model of peer-reviewed science disseminated through print journals is under stress, and experiments to develop better ways of validating and evaluating the importance and worth of

scholarship as well as making new knowledge accessible are needed.

Viewed in aggregate a significant amount of resources are invested to support this system including subscription fees paid by university libraries, and the costs of faculty time devoted to editing and reviewing articles. It is reasonable to ask if the current mixed-model of major commercial vendors plus university presses and academic societies remains the most efficient way to manage this system or if other arrangements might achieve better results at lower costs.

Following the model of earlier investments, it would be valuable for Federal Agencies to support developing the infrastructure for new models of scientific communication. Exactly what these will be is not yet obvious, so investments in multiple competing models is desirable. These investments might well be made in partnership with existing private enterprises involved in the business of scholarly communication—including university presses, for profit publishers and scientific societies.

Comment 5

This is a technical question to which I cannot respond substantively, Federal agencies could facilitate the formation of appropriate standards-setting bodies to examine these issues and make recommendations.

Comment 6

Federal Agencies should acknowledge that science requires not just research, but the communication and curation of the results of that research. Supporting robust systems of scholarly communication and curation of research results seems to be an important implication.

Comment 7

It is not clear to me why one would distinguish between different modes of publication in developing policies for public access.

Comment 8

This is a good question, and seems to require further research

Joshua L. Rosenbloom
Associate Vice Chancellor for Research and Graduate Studies
Interim Dean, Graduate Studies
Professor of Economics
University of Kansas

December 16, 2011

**Re: OFFICE OF SCIENCE AND TECHNOLOGY POLICY (OSTP) – Request for Information:
Public Access to Peer-Reviewed Scholarly Publications
Resulting from Federally Funded Research**

On behalf of Carnegie Mellon University (CMU) and the roughly 4,000 faculty and staff we represent, I write to thank you for this opportunity and to share our perspective on public access to peer-reviewed publications resulting from federally funded research. Carnegie Mellon is a small, private university with over 11,000 students and 86,500 alumni. Recognized for our world-class programs in technology and the arts, interdisciplinary collaborations, and leadership in research and education, we are innovative and entrepreneurial at our core.¹

In September 2011, the Times Higher Education World University Rankings placed Carnegie Mellon at 21 overall and 12 in science and engineering. In August 2010, Unigo.com included CMU among the New Ivies. In 2007, the Academic Analytics Faculty Scholarly Productivity Index ranked us 4 overall, and in the top ten in nineteen disciplines, including 1 in information technology / systems; 2 in cognitive science, operations research, business, and public administration; and 4 in electrical engineering, computational science, and computer science.² Eighteen of our faculty and alumni have won Nobel Prizes. We firmly believe that open access to publicly funded research will increase productivity, innovation, and commercialization.

COMMENT 1

To maximize return on taxpayer investment in research, grow the economy, and improve the productivity of science, all federally funded peer-reviewed publications should promptly be made available to the public for full use without commercial restrictions. Full use entails not just the right to read and download the work, but the right to distribute, mine, crawl, create derivative works, and use in computations.

¹ See <http://www.cmu.edu/about/index.shtml>.

² See <http://www.cmu.edu/about/rankings-awards/rankings/rankingsroundup.pdf>.

Mandating prompt public access and full use rights – open access – to peer-reviewed scholarly publications will benefit not only higher education and research institutions, but small business owners, practitioners, and inventors currently hindered by cost barriers. Open access will grow existing and new markets by accelerating and broadening access and increasing use of the material. Faster access enables ideas to be generated and applied more quickly, speeding up the development cycle, creating jobs, and stimulating economic growth. Broader access enables more users and different kinds of users to stay on the cutting edge, encouraging innovative use and application of the work. As Google and other search tools demonstrate, public access and use rights encourage private investment and job creation. In contrast, restricted use stifles innovation and limits commercialization. Mandating public access and full use rights will maximize opportunities, encourage investors, and enable creative individuals and companies to build new products and services.

The factors that stimulate economic growth will also improve the productivity of science. Open access drives innovation and productivity by broadening access (and hence potential for interdisciplinary work), increasing use (evidenced by increased citations³), and accelerating development. Full use rights will add to this benefit by allowing scientists to use computers to process research findings, enabling them to use new findings in new ways (e.g., machine reading and computation) and to process more information faster, speeding and expanding opportunities for innovation and commercialization. The National Institutes of Health (NIH) reports that most of the 500,000 daily users of articles in PubMed Central are unforeseen users, i.e., not researchers in the .edu Internet domain.⁴ Enabling unforeseen users and uses of federally funded scholarly publications will inevitably increase scientific productivity and commercialization of the findings. To maximize the number of people with the opportunity to contribute, the U.S. government must adopt mandatory open access policies.

The social and economic benefits of providing fast, free access and full use rights will far exceed the financial cost of ingesting, maintaining, and providing access to the content. The NIH spends less than 1% of its annual budget processing submissions, refining the system, staffing a help desk, and maintaining its database of two million articles.⁵ **According to the Houghton report, *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*, over a transitional period of 30 years from implementation, open access to all federally funded scientific research modeled on the NIH policy will increase return on taxpayer investment by at least a factor of four and grow the U.S. economy by**

³ See, for example, C.J. MacCallum and H. Parthasarathy (2006), “Open Access Increases Citation Rate,” *PLoS Biology* 4(5): e176. doi:10.1371/journal.pbio.0040176. See also the bibliography available at <http://opcit.eprints.org/oacitation-biblio.html>.

⁴ D.J. Lipman (April 19, 2011), Testimony on Public Access to Federally-Funded Research. See <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>. Lipman reported that approximately 25% of PubMed Central use comes from universities, 17% from companies, and 40% from the public sector. The remaining 18% is government or other users.

⁵ Lipman, 2011.

roughly a billion dollars.⁶ Following the transitional period, the benefits of steady-state open access would be more than 50 times the cost. Federal public access mandates that expand on the NIH model to include full use rights will increase return on investment and grow the economy even more by allowing innovative uses and accelerated processing previously prohibited without permission.

Open access publishing and self-archiving are more cost-effective systems for scholarly publishing than toll access (subscription) publishing, offering cost savings throughout the scholarly communication process. According to the Houghton and Oppenheim et al report, *Economic implications of alternative scholarly publishing models*, “there are gains to be realized from moving towards open access publishing models and, despite the lag between the costs and the realisation of benefits, the transition may be affordable within existing system-wide budgetary allocations.”⁷ (See comment #9.) Transition cost has long been a key concern of academic administrators.

Federal public access policies can be implemented cost effectively by leveraging existing infrastructure and expertise. (See comments #3 and #4.) Furthermore, public access will provide data for federal agencies to assess return on investment and to target funding to the most promising research. Mandated public access is the appropriate federal response in an era calling for fiscal accountability, evidence-based decision making, and transparency.

COMMENT 2

Public access policies can be implemented within the current copyright framework. All that is needed to maximize return on taxpayer investment in research is an emphasis on author rights and appropriate licenses.

The NIH public access policy allows only limited use of the material. Greater utility is needed to realize the potential scientific and commercial value of federally funded research findings. Public access policies must enable full use of publicly funded work (i.e., the right to copy, distribute, re-use, crawl, create derivative works, mine text and data, and use in computation) in a timeframe much shorter than the copyright term. This can be accomplished under current copyright laws by implementing licenses that do not restrict use, for example, the Creative Commons CC-BY (attribution) license. To

⁶ J. Houghton, B. Rasmussen, and P. Sheehan (July 2010), *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*, Report to SPARC, pp. 7-8. Available at <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>. According to the report, if calculations are based on federal R&D lifecycle costing, the benefits of open access will be four times the cost. If calculations are based on the NIH's costing, the benefits will be eight times the cost. If calculations are based on arXiv costing, the benefits will be 24 times the cost. These calculations are based entirely on increased returns to R&D. Savings in time spent searching or acquiring materials or duplicating research are not included.

⁷ J. Houghton, C. Oppenheim, et al (January 2009), *Economic implications of alternative scholarly publishing models: Exploring the costs and benefits*, A report to the Joint Information Systems Committee (JISC), p. 14. Available at: <http://www.jisc.ac.uk/media/documents/publications/rpconomicoapublishing.pdf>.

balance the interests of various stakeholders, a reasonable compromise would be an embargo period of restricted access and use, followed by public access and full use under an appropriate license. (See comment #8.)

Federal agencies need to understand what is at stake in articulating a public access policy from different stakeholder perspectives, and where in the supply chain each stake is claimed. In short:

- The public's stake in federally funded research is rightfully claimed at both the beginning and end of the supply chain. Taxpayers have a stake in how their tax dollars are invested in research – a driver for transparency and accountability – and a stake in access to the research findings funded with their tax dollars.
- Federal agencies have a stake in ensuring that their investment of taxpayer dollars yields maximum return. (See comment #1.)
- Intellectual property owners have an important stake in public access policies. The initial copyright owner of federally funded research findings is probably the author or perhaps the author's employer. Publishers are not the initial copyright owners, though they have a stake downstream of the authors' copyrighted expression of research findings. Authors are primarily interested in securing recognition and a reputation for advancing knowledge. Publishers are primarily interested in revenue.
- Academic institutions bear all the real costs of performing peer review by paying faculty salaries and providing office and laboratory space. In return, they have a right to expect broad dissemination and innovative use of research findings, in keeping with their mission. Publishers only coordinate peer review.

Federal public access policies must require initial copyright owners to grant the government the right to archive, distribute, and enable full use of their publicly funded work. Creative Commons licenses should be considered standardized terms of use. The right to license full and public use is enforceable under current copyright law.⁸

Traditionally publishers required authors to transfer their copyrights exclusively to them, to ensure their investment in publication and dissemination was profitable. Authors complied because they had no alternative when work was disseminated in print. The Internet dramatically changed the situation. Now authors have open access alternatives. They can publish in journals that allow them to self-archive their work in an institutional or disciplinary repository (known as green open access), or publish in open access journals (gold open access). At Carnegie Mellon University, roughly 83% of the faculty and graduate students have self-archived their work or published in an open access journal.⁹ The Houghton and Oppenheim report indicates that both green and gold open access are more cost effective ways of disseminating scholarly work than subscription (toll access) publishing.¹⁰

⁸ See <http://creativecommons.org/weblog/entry/8826>.

⁹ J. Sutkus (2010), *Faculty Survey of Library Resources 2009-10 and Graduate Student Survey of Library Resources 2009-10*, Carnegie Mellon University: Office of Institutional Research and Analysis. See also D. Troll Covey (2011), "Recruiting Content for the Institutional Repository," *Journal of Digital Information* 12, 3. Available at http://works.bepress.com/denise_troll_covey/56/.

¹⁰ Houghton and Oppenheim et al, 2009.

As of December 2011, 57% of the 1,043 publishers with policies in the SHERPA RoMEO database allow self-archiving of peer-reviewed articles without an embargo.¹¹ A small percentage allows self-archiving after an embargo period. In addition, many publishers support public access policies by depositing work in designated repositories for the authors. In 2009, 40% of the articles submitted to PubMed Central were deposited directly by the publisher. In April 2011, the NIH reported having formal agreements for direct deposit with more than 900 journals, “a number that has doubled in the two years since the policy became mandatory.” The number is expected to grow.¹²

In addition to self-archiving, open access publishing is thriving. As of December 2011, the Directory of Open Access Journals (DOAJ) lists 7,000 journals; roughly four new titles are added daily.¹³ DOAJ statistics indicate 1,432 open access titles added this calendar year.¹⁴ Many open access journals are published by scholarly societies. According to the *SPARC Open Access Newsletter*, in 2007, 425 scholarly societies published 450 open access journals; in 2011, 530 societies publish 616 open access journals. Roughly 20% of these journals are published in the United States.¹⁵ Open access is a growth industry and grand opportunity in publishing. Even traditional toll access (subscription based) publishers – including Elsevier, Nature, Sage, Springer, Taylor & Francis, and Wiley-Blackwell – are joining the movement, either by providing a fee-based open access option for publishing in subscription journals (known as hybrid journals) or, more recently, by publishing fully open access journals.¹⁶ Open access publishing creates competition, ultimately yielding better (more effective) products at better prices. The Study of Open Access Publishing (SOAP) project found that

- At least 120,000 articles are published in fully open access or hybrid journals each year.
- Most open access articles are published by for-profit publishers.
- Two-thirds of open access journals and 75% of open access articles are STM.
- Fourteen publishers published roughly 72% of the open access journals and articles. A vast majority (82%) of these articles are published under a Creative Commons Attribution (CC-BY) license – granting the full use rights recommended here as critical to maximize the return on taxpayer investment in research.¹⁷

There is no evidence that libraries have cancelled journal subscriptions because some of the articles are available open access. However, should publishers see reduced revenue as a result of open access, “these reductions may well be offset by revenue gains from selling value-adding services to a larger number of readers and/or authors

¹¹ See <http://www.sherpa.ac.uk/romeo/statistics.php>.

¹² Lipman, 2011.

¹³ H. Morrison (September 30, 2011), *Dramatic Growth of Open Access*. See <http://poeticeconomics.blogspot.com/2011/09/dramatic-growth-of-open-access.html>.

¹⁴ See <http://tinyurl.com/85bvfrf>.

¹⁵ See <http://www.earlham.edu/~peters/fos/newsletter/12-02-11.htm>.

¹⁶ See <http://www.sherpa.ac.uk/romeo/PaidOA.html>.

¹⁷ S. Dallmeier-Tiessen, R. Darby, B. Goerner, et al (2010), *First Results of the SOAP project. Open access publishing in 2010*. Available at <http://arxiv.org/ftp/arxiv/papers/1010/1010.0506.pdf>.

and from alternative revenue streams.”¹⁸ Furthermore, open access creates jobs across a broad spectrum of content, from peer-reviewed scholarly communications to user-generated social networking, e.g., from the Public Library of Science (offices in San Francisco and the UK) and Hindawi Publishing (offices in New York and Cairo) to YouTube, Flickr, and Facebook.

Despite ongoing publisher push-back on open access to scholarly work, increasingly publishers are leveraging open access to their financial advantage. **The U.S. government should not concern itself with preserving business models that no longer serve the best interests of authors and readers.** It should focus on supporting new business models that serve these interests, for example, by funding Article Process Charges (APCS) for authors to publish in fully open access journals. Publishers have a legitimate claim to the formatted, copyedited, published version of a work. They might have a claim to the peer-reviewed author’s manuscript version, though given that publishers coordinate peer review (for which publishers expect remuneration) and academics perform peer review (funded gratis by their institution), the academy also has a claim to the peer-reviewed manuscript.¹⁹ **What publishers cannot rightfully claim is that federal mandates result in the government’s taking of private sector products.** If initial copyright owners choose not to transfer their copyrights to certain publishers, for whatever reason, that is their prerogative as copyright owners. If publishers choose not to support the rights authors need to comply with federal mandates, for whatever reason, that is their prerogative as publishers. Federal mandates do not take private sector products, but rather encourage authors to choose wisely in their own, their institution’s and the public’s interest.

Initial copyright owners – primarily authors – need to be educated, empowered, and required to steward their copyrights wisely. They must retain and license the rights necessary to comply with federal public access policies.

COMMENT 3

The federal government should be responsible for stewarding federally funded research findings. The government is in a unique position to ensure that the collections are permanently preserved, publicly accessible, and fully usable. The government should keep a copy of all research findings they fund. However, each agency need not maintain its own repository and additional repositories are necessary. These two

¹⁸ Houghton and Oppenheim et al, 2009, p. 16.

¹⁹ Houghton and Oppenheim et al provide an analysis of the cost of producing a journal article:

- 15% of the total item cost is the performance of peer review
- 55% is research and writing
- 30% is publisher related activities
- <1% is library acquisition and handling

A subsequent analysis reveals that universities fund 80% of the cost of peer review. Given that universities and funding agencies fund most of the costs of producing a journal article, they have a legitimate claim to the peer-reviewed manuscript. See Houghton and Oppenheim et al, 2009, pp. 6-7.

caveats insure the cost effective implementation of public access policies and the redundancy necessary for long-term preservation.

The government must maintain a repository of all federally funded content because current market attempts to archive collections comprehensively are inadequate. For example, Cornell University recently reported that only 13% of their journal holdings are currently archived by LOCKSS and Portico.²⁰ Many publishers do not participate in LOCKSS or Portico, either because they do not meet the technical conditions required by these archives or because the publishers do not respond to requests from librarians that they participate. Furthermore, though the cost of maintaining a LOCKSS archive is reasonable, licensing restrictions imposed by the publishers limit what can be done with archived copies. The limitations render actual use of the copies prohibitively expensive.²¹ Cornell and Columbia Universities currently treat their LOCKSS caches as dark, inaccessible archives.

Experience has demonstrated that maintaining a dark archive is not a viable preservation strategy. Without access and use, media degrade, files corrupt, and formats slip into obsolescence. Delays in discovering these problems render them difficult if not impossible to solve. Therefore repository content maintained by the federal government must be accessible and usable. Only regular access and use can ensure preservation and maximize the benefits of public access. Prompt access increases use in the short term, but long-tail use can be just as powerful, as evidenced by many successful Internet businesses, e.g., Amazon, eBay, iTunes, and Netflix.²²

The federal government need not allocate additional funds to each agency to develop infrastructure in support of public access. Large agencies, like the National Science Foundation, Department of Agriculture, and Department of Defense can, like the NIH, earmark 1% of their budget to develop and maintain an open access repository. Smaller agencies can piggyback on these efforts, increasing the return on investment.

Though the federal government must maintain an accessible copy, there must be other copies of the material. The government's investment in archiving will not unnecessarily duplicate the effort and cost of other providers. Additional repositories are necessary to provide the redundancy critical to long-term preservation and the decentralization critical to fostering interoperability standards.

These other repositories can be maintained by other organizations in partnership with the government. However, the government must stipulate the conditions or criteria that partner repositories must meet to support public access policies. Specifically, all partner repositories must

²⁰ Cornell University Library and Columbia University Library (March 2011), *Final Report of the 2CUL LOCKSS Assessment Team*, p. 17. Available at: <http://2cul.org/sites/default/files/2CULLOCKSSFinalReport.pdf>.

²¹ See Cornell University Library and Columbia University Library, March 2011.

²² See http://en.wikipedia.org/wiki/Long_Tail.

- Support public access to federally funded content
- Enable full use of the material as mandated by policy
- Commit to long-term preservation of the content
- Interoperate with other repositories of federally funded content

Complying with existing or emerging open standards that facilitate public access and allowing full and innovative use of the content will maximize return on taxpayer investment, grow the economy, and increase the productivity of science. As of January 2010, the interoperability developed by the NIH among its holdings was extremely beneficial, but the Scholarly Publishing Roundtable recommended interoperability with external repositories to maximize return on investment.²³ More recently, testifying before the United States House of Representatives, David J. Lipman, Director of the National Center for Biotechnology Information (NCBI), described the added value and enhanced return on investment derived from the NIH's interoperability with archives in the U.S., U.K. and Canada:

For example, during the recent H1N1 flu pandemic, NCBI was the major site for collecting all of the known flu sequences. Within months, NCBI had over 20,000 sequences from around the world. Taking advantage of the deep integration among NCBI systems, a researcher reading a paper on the spread of drug-resistant variants of the flu sequences could, with the click of a mouse, compare the new isolates to all other flu variants and gain insights into the epidemiological consequences. With equal ease, the researcher could map the variant viral proteins to known 3D protein structures to see how the mutations affect binding of the antiviral drug.²⁴

COMMENT 4

The work to date in support of a distributed system of interoperable repositories must not be overlooked. Federal agencies should leverage existing expertise, infrastructure, and standards by partnering with academic institutions and other organizations that have already invested substantial resources in open access repositories and long-term preservation of digital content. Publishers willing to provide repositories that meet the conditions for public accessibility, full use rights, interoperability, and long-term preservation of publicly funded work may also be suitable partners.

Under no circumstances should access to publicly funded content be restricted to a single site. This would conflict with full use rights, including the rights to copy and distribute (mirror or harvest) the content. The ultimate goal of public access policies is the creation of new knowledge, products, tools, and services achieved through enhanced dissemination, discovery, use, and preservation. Anything that disrupts or

²³ Scholarly Publishing Roundtable, (January 2010), *Report and Recommendations from the Scholarly Publishing Roundtable*, p. 13. Available at:

http://www.aau.edu/policy/scholarly_publishing_roundtable.aspx?id=6894.

²⁴ Lipman, 2011.

denigrates freedom of access and use will stifle innovation and commercialization and reduce return on taxpayer investment.

COMMENT 5

If federal agencies mandate and enforce deposit in open access repositories that meet established baseline criteria, the sheer volume of publicly accessible and fully usable material will drive development of interoperability standards and tools and, in the process, create jobs.

Federal agencies should require federally funded research results to be managed using open standards that are widely supported, regularly updated by the research community, and designed to facilitate interoperability and re-use of the material. Standards will no doubt evolve over time, but the government in collaboration with partner institutions can compile and maintain a list of accepted standards. For example, to facilitate searching, linking, mining, etc. and to optimize re-use, the storage format should use an open-standard mark-up like XML, with a standard document type definition (DTD) like the National Library of Medicine's DTD. The storage format must enable use by computers, not just people.²⁵

Metadata provides more than an item description. It is a tool to enable use, re-use, and content analysis. To maximize return on investment, metadata must be machine-readable and machine-interoperable. It must be coupled with an Application Programming Interface (API) to enable standards-based data exchange and harvesting. Repositories suitable for deposit of federally funded research results should support the Open Archives Initiative metadata harvesting standard (OAI-MH), the object reuse and exchange standard (OAI-ORE), and protocols like the Simple Web-service Offering Repository Deposit (SWORD) that facilitate automatic deposit of manuscripts in multiple repositories at once. The efficiency provided by SWORD will provide the necessary redundancy and increase compliance by enabling authors to deposit their work simultaneously in federal and institutional repositories.

Dublin Core is the minimum acceptable core descriptive metadata for publications. Explicit information about use rights must be included, preferably using a controlled vocabulary. Articles should be assigned a Digital Object Identifier (DOI) to ensure persistent identification, management, and linking. Similar unique identifiers are being developed for authors and institutions (e.g., Open Researcher and Contributor ID or ORCID). Descriptive metadata should also include attribution for funding organizations and grant IDs.

Developments in semantic metadata need to be encouraged, monitored, and adopted over time, for example, metadata that describe relationships between entities (using the

²⁵ PDF may be used for display purposes, but PDF is not an acceptable storage format because it does not support robust searching, linking, text mining, reformatting, etc., i.e., it reduces the utility and value of the work.

Resource Description Framework [RDF] or Web Ontology Language [OWL]) or build bridges between publications and underlying data. Federal agencies can support these developments by working with agencies invested in improving metadata interoperability, such as the Library of Congress (LOC), the National Information Standards Organization (NISO), and other standards organizations.

COMMENT 6

To minimize the burden on researchers, offices of sponsored research, and academic departments, federal agencies should mandate consistent public access policies and uniform deposit requirements across agencies. Consistent expectations and conditions for managing grants and research publications will reduce costs and complexity, save researchers and institutions valuable time, and increase the rate of compliance.

To maximize impact, federal public access policies should do everything possible to ensure compliance and to encourage innovative use and enhancement. **Author compliance should be quid pro quo for future funding.** Protocols and standards that streamline deposit and redundancy (e.g., SWORD, OAI-MH) are also critical to compliance. Developers should be made aware of opportunities to add value, for example, by creating tools that help users compile bibliographies or create author profiles, or enable universities to easily measure their research output or apply their brand. Added value features and functionality will encourage compliance.

COMMENT 7

Federal public access policies should apply to all royalty-free, peer-reviewed publications of research findings funded by taxpayer dollars. Conference proceedings in particular should be covered by public access policies because in some disciplines, such as computer science, conference proceedings are more highly valued than journal articles.

COMMENT 8

Prompt, free access with full usage rights will unlock value in publicly funded research unrealized to date because of barriers to access and restrictions on use. The Houghton reports calculate the financial effects of an embargo period between publication and self-archiving an open access copy. The shorter the embargo, the greater the gain, such that eliminating the embargo maximizes return on investment.²⁶ A twelve-month embargo over a twenty year period can reduce return on investment by 2% during a

²⁶ Houghton et al, 2010, p. 8.

period of transition from subscription to open access publishing; eliminating the embargo during this period can increase ROI by 3.6%.²⁷

Immediate access to peer-reviewed work would maximize the scientific and commercial value of publicly funded research. An embargo of twelve months or less is an acceptable compromise for the near future while publishers transition from a toll-access to an open-access business model. A report released by the Scholarly Publishing Roundtable – a group of academic administrators, academic librarians, STM publishers, and researchers – recommends an embargo of from zero to twelve months, the period to be determined by the federal agency in consultation with researchers and other stakeholders.²⁸

Some publishers and research communities have expressed reservations about an embargo of twelve months or less, claiming that one solution does not serve all disciplines. However, funding agencies and hundreds of journals in many disciplines around the world have adopted an embargo of twelve months or less, with no evidence of harm to publishers.²⁹ Publishers that previously expressed concern about such a short embargo recently changed their position with no untoward effects. For example, *Molecular Biology of the Cell* reduced their embargo from twelve to two months.³⁰

Interviews with Carnegie Mellon faculty in 2006 suggested that researcher concern about self-archiving open access copies is driven by unfounded publisher laments that self-archiving will result in canceled subscriptions. Faculty concern is heightened for scholarly societies that rely on surpluses from subscription income to fund conferences and other initiatives.³¹ A survey of librarians conducted by the Association of Learned and Professional Society Publishers (ALPSP) in 2006 revealed that faculty need, journal usage and price were the top three factors influencing journal cancellations. Few of the librarians surveyed would cancel subscriptions unless the embargo period was three months or less and the published version of record and 90% of the journal's content were available open access. When asked to project their views five years out (2011), 98% of the respondents said price would be an important or more important factor and the availability of open access copies would not be among the top three factors influencing journal cancellations.³² A cost-benefit comparison conducted by Houghton

²⁷ Houghton and Oppenheim et al, 2009, p. 14.

²⁸ Scholarly Publishing Roundtable, 2010, p. 12.

²⁹ See <http://roarmap.eprints.org/> and <http://highwire.stanford.edu/lists/freeart.dtl>.

³⁰ Scholarly Publishing Roundtable, 2010, p. 12, ft. 25.

³¹ Denise Troll Covey (2006). Unpublished report. *Preliminary Findings: Study of Carnegie Mellon Faculty Rights and Scholarly Communication Practices*.

³² Among the librarians surveyed, 82% said the embargo would have to be three months or less; 92% said six months or less. Only 39% saw an open access copy of the author's peer-reviewed manuscript as an acceptable substitute for the published work. Almost half (48%) of the respondents wanted the entire journal to be available open access before they would see open access copies as a substitute for the journal; 76% said an open access archive would have to contain 90% of the journal content for open access to be an acceptable substitute for the journal. See Marc Ware (2006), *ALPSP survey of librarians on factors in journal cancellation*, Association of Learned and Professional Society Publishers (ALPSP), pp. 2-3. Summary and conclusions available at: <http://www.alpss.org/ForceDownload.asp?id=53>.

and Oppenheim et al in 2009 indicates that the benefits of open access resulting from increased access and efficiency are sufficient to cover the cost of open access without canceling subscriptions.³³

Federal public access policies can encourage the transition from toll access to open access by mandating an embargo of twelve months or less. If different embargoes must be considered for different disciplines, the appropriate embargo should be determined by the federal agency in consultation with the research community based on a thorough, objective assessment of market conditions that carefully isolates the effect of the embargo independent of changes in price, competition, licensing practices, library budgets, and revenue from long-tail citations. Under no circumstances should the embargo period be determined by for-profit publishers.

COMMENT 9

To facilitate compliance with public access policies and the transition to a more cost-effective scholarly publishing system, federal agencies should fund payment of Article Process Charges (APC) for grantees to publish in fully open access journals. They should not fund payment to publish in hybrid journals. Hybrid journals constitute double dipping, with publishers securing both subscription revenue and an APC for the same content.³⁴

The Study of Open Access Publishing (SOAP) project found that scientists and scholars overwhelmingly support the idea of open access because of its many benefits, but only 8-10% of articles published annually are published in open access journals. A large survey conducted by SOAP found that 29% of the respondents had not published a single article in an open access journal. Among those who provided reasons for their delay, the biggest deterrent was funding to pay the APC.³⁵

Analysis conducted by Houghton and Oppenheim et al indicates that funding agencies and institutions could allocate 3.5% of their research funding to paying APCs without exhausting the net benefits of open access; 3.5% is much higher than reported allocations and much higher than what is required, based on careful estimates. The

³³ Houghton and Oppenheim et al, 2009, p. 12.

³⁴ As of early 2010, eighty publishers offered a hybrid option for a subset of their journals. However, few authors exercise this option, i.e., only 2% of the articles in hybrid journals are available open access. See S. Dallmeier-Tiessen, R. Darby, B. Goerner, et al (2010), *First Results of the SOAP project. Open access publishing in 2010*. Available at: <http://arxiv.org/ftp/arxiv/papers/1010/1010.0506.pdf>.

³⁵ S. Dallmeier-Tiessen, R. Darby, B. Goerner, et al (2011), *Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing*, pp. 3, 7, 11. Available at: <http://arxiv.org/ftp/arxiv/papers/1101/1101.5260.pdf>.

benefits of open access and the system cost savings far outweigh the cost of diverting funds to pay APCs.³⁶

Scholars in fields that are poorly funded need institutional support to pay APCs. Carnegie Mellon University is forming a Scholarly Communication Advisory Board to guide strategic initiatives. A high priority for the new Board is to explore how and under what conditions the University could underwrite APCs. Options include creating a fund to pay APCs (i.e., joining COPE, the Compact for Open Access Publishing Equity³⁷) or paying for memberships that offer discounted APCs (e.g., PLoS or BMC membership³⁸).

Open access journals are currently at a disadvantage compared to subscription journals. Publication fees levied by many open access journals are one of the reasons why authors who perceive the benefits of open access choose to publish in subscription journals. Subscriptions insulate authors from the attendant costs of publishing. The institution pays the toll by investing in research and in the library, creating what economists call a “moral hazard.” Moral hazards arise when individuals are isolated from the consequences of their choices. When someone else is held responsible for the consequences, people tend to act less carefully than they otherwise would. In the current context, the university (through funding for the library) suffers the consequences, paying escalating prices for access to subscription journals.

Stuart Shieber of Harvard argues that funding agencies must take responsibility not only for funding important research, but for the optimal distribution of research findings. “Part of that responsibility is putting in place an economically sustainable system of underwriting open-access publication fees.”³⁹ He outlines criteria that would enable funding agencies to underwrite APCs without introducing a moral hazard. For example, the government must:

- Level the playing field for open access journals by allocating funds that can only be used to pay APCs in fully open access journals, not hybrid journals.
- Recognize that research publications often occur after the grant period ends. Therefore the government must allow funds allocated for APCs to be spent after the grant period ends.
- Provide incentives for publishers to switch from toll access to open access publishing. This is why federal funds for APCs must **not** be used to publish in hybrid journals. To allow this would provide no incentive for publishers to transition to open access publishing.
- Avoid the moral hazard of insulating authors from the costs of publication. Fund APCs in full or in part up to a limit. A cap will encourage authors to choose low-cost, high-service open access journals.

³⁶ Houghton and Oppenheim et al, 2009, pp. 13, 15.

³⁷ See <http://www.oacompat.org/>.

³⁸ See <http://www.plos.org/support-us/institutional-membership/> and <http://www.biomedcentral.com/libraries/membership>.

³⁹ S. Shieber (November 16, 2011), “How should funding agencies pay open-access fees?”, *The Occasional Pamphlet*. Available at: <http://blogs.law.harvard.edu/pamphlet/2011/11/16/how-should-funding-agencies-pay-open-access-fees/>.

In closing, the federal government is in an ideal position to support key, evidence-based recommendations made by Houghton and Oppenheim et al.⁴⁰ The government can help overcome barriers by paying the APCs of grantees publishing in fully open access journals, by funding and encouraging the development of interoperable open-access repositories, and by advocating for alternative publishing models. The government can help realize system-wide cost savings, maximize return on taxpayer investment, grow the economy, and increase the productivity of science by mandating open access and full use rights to federally funded research.

Sincerely,

Gloriana St. Clair

⁴⁰ Houghton and Oppenheim et al, 2009, p. 18.

Subject: public access

Date: December 20, 2011 11:32:11 AM EST

I recommend you provide public access to scholarly journals in the fields of education, social work, psychology, mathematics, physics, and religion/philosophy.

I am happy to elaborate if needed. Thanks.

Shulamis Hes

Electronic Collections Librarian

Pollack Library

Yeshiva University

Subject: Public access to research funded by government moneys should be supported

Date: December 20, 2011 12:02:51 PM EST

As a taxpayer it is my right to have access to research I have helped support. As a scientist supported by the government I feel it is unethical to keep results from the public, or to make them pay for access.

--

Jenella Loye
Dept. Entomology, Univ. California Davis
Carroll-Loye Biological Research

Please continue policies that require researchers to make research results public through entities like PubMed , repositories or the like. As a librarian for a scientific research institute, our institute pays for subscriptions, we also pay article charges to get something published. As a taxpayer, I think all the research that results from federally – funded grants should be published in a way that provides open access opportunities to stimulate the development and application of scientific results that can better the human condition. I have seen first -hand the difficulties presented by research results that are published by private publishers with limited distribution to only those who can afford to pay.

Thank you.

Pat

Pat Weaver-Meyers
Director of Library Services

The Samuel Roberts Noble Foundation

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

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

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

Submitted December 23, 2011 via email: publicaccess@ostp.gov

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

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

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

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

The Medical Library Association (MLA) and Association of Academic Health Sciences Libraries (AAHSL) have been able to observe firsthand the significant benefit of public access to publications arising from National Institutes of Health (NIH) funded research including its cost-effectiveness. The NIH public access policy model has proven to be cost effective, representing about 1/100th of one percent of the overall NIH budget (\$3.5-4.6/million annually based on a \$30 billion budget). The PubMed Central database (<http://www.ncbi.nlm.nih.gov/pmc/>) contains over 2,000,000 articles and current usage is now up to 1,000,000 users per day (and growing!). The spread of usage across all the articles in the database is not what might

generally be expected (e.g. a small subset of articles getting the most usage), but rather consists of 99% of the articles being accessed on an annual basis. It is important to note that the Medical Library Association's premier journal, *Journal of the Medical Library Association (JMLA)* has been an open access publication since 2001 and provides immediate access upon publication through PubMed Central (back to vol. 1, 1898!) (<http://www.ncbi.nlm.nih.gov/pmc/journals/93/>).

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

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

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

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

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

fundors to make better decisions in the interest of scientific progress and applications that derive from it.

In summary, the importance of quicker and open access to new research findings cannot be underestimated in its critical role in spurring the faster application of research, thereby triggering a faster cycle for follow-up research, new findings, and new applications. All of these in combination provide significant economic benefits including new business opportunities, creation of new jobs, reducing duplication in research, and speeding the research and development process.

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

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

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

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

in a democratic society. Knowledge is indispensable in making informed decisions in health care, education and research.

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

In response to OSTP's question, "Conversely are there policies that should not be adopted with respect to public access to peer-reviewed scholarly publications so as not to undermine any intellectual property rights of publishers, scientists, Federal Agencies and other stakeholders?", MLA and AAHSL believe that all policies should be reviewed and evaluated for their ability to solve even a part of this many layered issue.

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

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

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

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

An example provided by a technological start-up investor highlights the frustration often experienced when entrepreneurs seek to use scholarly publications. The most recent scientific

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

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

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

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

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

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

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

The digital preservation strategies being undertaken by HathiTrust (<http://www.hathitrust.org/>) provide a framework for ensuring long-term integrity of deposited materials. For example, their use of standard and open content formats that meet community-accepted digital preservation standards, support interoperability among current and future systems, and their regular, automated system checks that verify the integrity of digital content and objects ensure quality control and long-term user access and preservation.

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

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

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

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

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

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

The Medical Library Association (MLA) and the Association of Academic Health Sciences Libraries (AAHSL) believe that immediate, free online availability of the results of taxpayer-funded research is optimal in order to maximize scientific productivity and accelerate commercial innovation. The Public Access Policy currently in effect for publications resulting from research funded by the National Institutes of Health (NIH) allows for an embargo period of one year. Though there was resistance to this policy from the commercial publishing community, there is no evidence that library subscriptions to scholarly journals have declined because articles resulting from taxpayer-funded research are publicly available in PubMed Central. Since journal content includes many other types of articles, such as editorials, commentary, letters, etc., libraries continue to subscribe to the widest selection of journal titles that their budgets permit.

It is important to note, however, that all types of libraries face unprecedented challenges to maintain their journal collections in the face of serious budget cuts, and subscription costs that continue to rise far in excess of the general inflation rate. Most libraries have been forced to cut journal subscriptions because of declining resources. In this grim scenario, access to the results of taxpayer-funded research will be increasingly restricted to a dwindling circle of

scholars who have immediate access to a well-funded library. Those who do not enjoy such access must do without.

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

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

We appreciate having the opportunity to provide comments on this timely and important initiative. If you would like additional information please contact Mary Langman, Coordinator, Information Issues and Policy, Medical Library Association or Gary Freiburger, President, Association of Academic Health Sciences Libraries

Subject: Response from the Consortium of Academic and research Libraries in Illinois

Date: December 20, 2011 4:57:30 PM EST

The Consortium of Academic and Research Libraries in Illinois (CARLI) welcomes the opportunity to provide information and recommendations on the **America Competes Reauthorization Act of 2010** (ACRA; Pub. L. 111-358) in order to ensure long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally funded scientific research.

CARLI is a consortium of 153 academic and research libraries that serve more than 880,000 students and more than 10,000 instructors, scholars and researchers. Member institutions include major research universities, five medical schools, over sixty schools training nurses and allied health professionals, a comprehensive network of forty eight community colleges, and an array of quality baccalaureate institutions. Many CARLI members will submit responses focused on the specific institutional needs, but there is also an important shared perspective to consider.

The economic argument for open access to publications derived from federally funded research can be clearly demonstrated through the examination of information resource sharing in a library consortium, such as CARLI. The identification, training and employment of scientists and technical staff comprise a continuum. Community college students and undergraduates become technicians and graduate students, who become public and private sector researchers and scientific scholars. In so far as the economic strength of the United States depends on robust and cutting edge scientific research, open access to the literature of federally funded research is both a social and economic good.

This progression is materially strengthened when the information resources available to students and faculty at each stage are consistent, relevant and of the highest quality. Making the most current and important research findings available to any student or faculty member in Illinois means that more and better scientists will be engaged in the next generation of research in the most timely and cost effective manner. In addition, open access to scientific information poses little, if any, threat to publishers' markets. CARLI's experience clearly shows that the majority of our 153 members cannot afford and subscribe to relatively few scientific publications of commercial publishers or scholarly societies. Only as part of consortial resource sharing can such information be made available. For our members who do not or cannot participate in the scientific publishing marketplace, open access represents no threat to the publishing industries interests.

Protection of intellectual property rights is not central to the discussion of open access for peer reviewed publications in modern science. In the majority of cases, the appearance of a published article represents a historical record. Publication has always trailed the conduct of actual science significantly, and the age of electronic information and instant communication has only widened the gap. Increasingly, the published article represents a mere synopsis of research, while electronic access to data sets, images and analytical results are the real intellectual property. The published literature of science is central to the promotion and tenure structure of American higher education, but plays a less vital role in scientific practice. Scientific literature is also vital

to the educational process and in this dimension open access accelerates the process, as the level of resource sharing conducted by a diverse consortium like CARLI demonstrates.

As a consortium devoted to the development of equitable and effective means of discovering and accessing information for 880,000 and 10,000 faculty, CARLI strongly believes that a centralized approach to managing the federally funded products of research is in the best interest of the scientific and educational communities. The ability to create clear and consistent metadata for these publications means that all of the relevant research can be efficiently discovered and made available to meet the instructional, scholarly and investigative needs of both public and private researchers and scientists.

CARLI has extensive experience in the challenges of meeting the diverse needs of our users with multiple electronic platforms and complex commercial licenses for information that was paid for with tax dollars and should be readily accessible. The efficacy of consolidated management of information resources has been demonstrated by the current NIH requirements for open access to the research it supports.

Since the published record compiles science already done, extensive embargo periods are not warranted. The investment in the peer review and editorial process of the publishing industry should be acknowledged, but they should not be overestimated. It is incumbent on the publishing industry to document the true cost of enterprises that utilize reviewers and editors from the academy.

In conclusion, CARLI, on behalf of its member libraries and the institutions they serve, strongly urges that the **America Competes Reauthorization Act of 2010** and any rules derived to implement the Act encourage, support and require that the literature resulting from federally funded research be made openly accessible to any citizen as quickly, transparently and inexpensively as possible. The approach taken by the NIH to accomplish this should serve as a model for all federally funded research.

Name/Email

Response to "Request for Information: Public Access to Peer-Reviewed Publications Resulting From Federally Funded Research, November 2011"
Prepared for and the official RFI response of the University of Oregon
JQ Johnson

Director, Scholarly Communications & Instructional Support
Professor, University of Oregon Libraries

University of Oregon

email: jqj@uoregon.edu

[corresponding author]

Dean Walton

Science Librarian and UO Libraries Biology subject specialist

University of Oregon

email: dpwalton@uoregon.edu

Deborah A. Carver

Philip H. Knight Dean of Libraries

University of Oregon

email: dcarver@uoregon.edu

Affiliation/Organization

University of Oregon Libraries

City, State

Eugene, OR

Comments

Prefatory note: The following comments begin from the perspective that the most natural and preferred federal policy for the next few years includes extension of a system similar to the NIH public access mandate to multiple federal agencies to promote public access, and that the major questions involve variants on that scheme. Our overall belief is that the NIH public access mandate was an excellent start in 2008, but that as the world has changed it requires extension and refinement to further meet the overarching public access goals and take advantage of changes in the scholarly publishing industry.

Comment 1

- (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 NIH mandated deposit policy has had a major positive impact on creating new markets and models for peer reviewed publication within the biomedical community, and needs expansion to additional agencies. For example, it appears to have been the major driver for the development of new publishing enterprises such as PLoS and BioMed Central. Overall, the rapid growth in the number of “gold” open access journals, particularly in the life sciences, indicates that the NIH policy is having a positive effect in creating new markets, and hence is one that should be emulated by other federal agencies. As the purchase of BMC by Springer last year indicates, these new enterprises are clearly seen as potentially valuable by more traditional publishers, and in fact the traditional publishers are rapidly moving to invest in their own open access journals under brands such as “SpringerOpen,” “Wiley Open,” “Sage Open,” the Taylor & Francis “Author Rights Initiative,” etc. Some traditional publishers are beginning to realize that public access to selected articles in their (hybrid) journals is essentially free advertising, whereas others appear to have identified open access publishing as a potentially profitable business model. Most importantly, note that these new commercial ventures are currently growing exponentially, with no evidence that growth is anywhere near the flattened top of the logistic (Laakso et al., 2011); we can expect substantial further expansion in this market as long as agency policies do not choke off this growth.

Although creating new markets and business models for peer reviewed publications is an important goal, it is part of a larger goal of using publicly funded research to spur increased commercialization (Houghton & Sheehan, 2006). This goal will be best served by creating opportunities for innovation in the use of scientific information, rather than focusing narrowly on peer reviewed publication. In particular, it is vital that the full text of articles, as well as other products of the funded research such as specimens and datasets, be made freely accessible without commercial restrictions. Having access to this corpus will allow new innovative uses such as text mining, synthesis and abstracting services, visualization tools that allow researchers to identify new interdisciplinary connections and reviewers to assess scholarly impact in new ways, etc. It spurs economic growth not just in the publishing industry but throughout the economy. Ultimately, the economic benefits accrue because open access drives scientific and technological innovation by increasing readership (e.g., as measured by increased citations to OA publications (Wagner, 2010)), decreasing time to impact, and promoting diversity in follow-up research by making research widely available outside of narrow disciplinary areas.

Estimates of direct R&D benefits to opening access to all articles derived from U.S. publicly funded research suggest a 5X return on investment (Houghton, Rasmussen, & Sheehan, 2010).

It is gratifying to see the degree to which public access policies have spurred innovation and creativity in what appeared a decade ago to be a very conservative academic publishing industry that was not responding rapidly to the opportunities afforded by the world wide web. Examples of innovation that have been driven by the NIH deposit policy abound. For example, it seems doubtful that PLoS One, with its innovative model of peer review and journal funding, would have been successful absent funding agency mandates. And PLoS One has certainly been successful – it is expected to publish more than 14,500 peer reviewed articles during 2011, making it the largest peer-reviewed journal in the world.

One may also point to Google Scholar as a successful commercial (advertising-funded) service that depends on the existence of a large corpus of open access journal articles including PubMed Central.

More directly, the existence of a corpus of freely accessible publications is beginning to result in new tools where “readers” are computers rather than individuals. One typical example from the technology sector is recommendation engines such as those used by Google or Amazon (to offer improved search results based on mining an individual user’s previous search patterns and data on global popularity and click-thrus from particular searches). Another example is publishing tools for detecting plagiarism or duplicate publication such as iThenticate (iParadigms, 2011), almost all of which mine open access databases including PubMed as one of their resources. Applications that mine publication databases and produce meta-analyses or visualization of the pattern of scientific results from multiple studies have been in existence for substantially more than a decade (Kostoff & DeMarco, 2001), and indeed the field even has its own OA journals (e.g., *BMC Bioinformatics*), but the tools are maturing with the availability of more open publication data. Some well known examples include BioCreative, CoPub, and PubGene (Krallinger, Valencia, & Hirschman, 2008). Another promising new example is the new openSNP system (Greshake, Zimmer, Rausch, & Bayer, 2011), which in addition to collecting open data on genotypes correlates that data with relevant references in open access publications.

Concrete examples of additional steps that could be taken to improve access and analysis of peer reviewed publications include:

- Expanding the corpus of open access articles available in central and standardized repositories to a much wider range of disciplines and research areas, by collecting and making available with standardized interfaces copies of journal articles that would otherwise not be open access and extending NIH-style mandates to additional funding agencies;
- Clarifying that public access is based on a limited and nonexclusive rights transfer via prospective license from funded authors as part of federal research contracts, prior to any transfer of additional rights that an author may agree to as part of a publication agreement and hence not based on any theory of imminent domain or taking of publisher rights;

- Providing consistent programmatic interfaces and metadata that foster text mining of multiple simultaneous text collections;
- Providing more convenient mechanisms that as a supplement to existing mechanisms allow authors to deposit articles in institutional repositories, with automated harvesting using SWORD, OAI or similar protocols into centralized, federally maintained, archives (Harnad, 2008);
- Assuring widespread public and scholarly access (to preprints, to peer reviewed and accepted manuscripts, and to formal – often publisher maintained – copies of record) to maximize serendipitous and interdisciplinary discovery;
- Encouraging the development of “gold” open access publishing through appropriate funding, both via direct and automatic grant funding to extramural researchers to pay article processing charges, and via grant programs to stimulate the creation of new open access peer reviewed publications or to explore technologies, new business models, and new models of peer review that could further reduce the overall cost of publication and dissemination;
- Providing incentives to researchers to review manuscripts submitted to venues that make articles available open access;
- Investing in research into software and tools for text mining and visualization; collaborations between other agencies and those such as NSF and DARPA that have been traditional funders of machine learning and computer science research seem particularly fruitful;
- Investing as part of the federal grant funding and evaluation process in the internal use of text mining tools for identifying promising areas for future research, for example by identifying newly “hot” topics and research areas that cross disciplinary boundaries;
- Developing and promulgating standards for metadata that facilitate discovery and broad reuse, plus standards for programmatic access to publication data both for text mining and for mirroring of collections.

Specific costs and monetary benefits of providing access to publications depend heavily on the details of the implementation. We assume a mixed strategy where publishers may maintain a copy of record and where in many cases universities provide additional copies as part of their institutional repositories with tools that allow mirroring of articles and archives for preservation and specialized needs, but where the federal government, with its interest in guaranteeing access, provides access to additional copies and is able to assure consistent accountability, metadata, and access standards across multiple federal agencies. This NIH-style model has proven quite cost effective: NIH reports costs of \$3.5 to \$4.0 million per year (about

1/10,000 of the total NIH budget) to provide access (Lipman, 2011). Generalizing it to other agencies would leverage existing infrastructure and minimize costs and provide consistency in user interfaces. Note, though, that centralized archiving in the physics and mathematics communities (arXiv) is perhaps a factor of 3 even less expensive, so there is room for improvement. On the benefit side, calculating only the direct benefits of a FRPAA-style policy on improved U.S. R&D, Houghton et al. calculated that net present value of benefits over a 30 year period would be in the range of \$1.6B to \$1.75B, a factor of 4 to as high as 24 benefit-cost ratio depending on assumptions for the efficiency of providing the centralized access (Houghton, et al., 2010).

It is important to note here (for further discussion see Comment 2) that at a minimum the benefits of public access require unrestricted read access to a comprehensive collection of the texts of at least the author final versions of articles, plus the right of other researchers to use such articles in ways consistent with typical academic practice such as quoting from them. To the extent that the corpus is incomplete or embargo-limited it introduces serious problems with possible bias as some studies are ignored. To the extent that the access is limited by subscriptions or usage constraints it decreases the probability that new entrepreneurs will enter the market and may make it impossible to mirror archives for preservation or experimentation with new access approaches. Almost certainly further improvements in the use of federally funded articles will require standardization on a generally-accepted form of license that allows greater usage rights than the current NIH-mandated minimum. In the 3 years since the NIH mandate was established, Creative Commons licenses have emerged as the clearly preferred standard throughout the world, with hundreds of millions of works now released under Creative Commons licenses (Creative Commons, 2011).

It is not yet completely clear whether adequate commercialization opportunities will exist if works are made available under a license such as the minimalist Creative Commons CC-BY-ND (which does not even appear to grant any right to quote portions of an article beyond what is allowed by fair use); there are growing and compelling arguments that rich data mining and visualization require that authors also allow creation of derivative works. Certainly an application that used the corpus of Pub Med Central articles to provide improved machine translation of articles (perhaps from Mandarin to English, a tool that would be extremely valuable given the rapid growth of important scientific literature published only in Chinese) would require such rights. This suggests that although standardization on a CC-BY-ND license in 2012 might be politic, it is likely to be important in the future to make works available under a CC-BY license or equivalent to maximize the likelihood both of new applications and of the successful commercialization of derivative products.

References

[Please observe that works cited in this response are except as explicitly noted to works released as open access.]

Creative Commons. (2011). Creative Commons [web site] Retrieved 6 Dec 2011, from <http://creativecommons.org>

Greshake, B., Zimmer, F., Rausch, H., & Bayer, P. (2011). openSNP Retrieved 5 Dec 2011, from <http://opensnp.org/>

- Harnad, S. (2008). *Optimize the NIH Mandate Now: Deposit Institutionally, Harvest Centrally*. Technical Report. Electronics and Computer Science. University of Southampton. Retrieved from <http://eprints.ecs.soton.ac.uk/15002/>
- 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. Retrieved from <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>
- Houghton, J., & Sheehan, P. (2006). *The Economic Impact of Enhanced Access to Research Findings*. CSES Working Paper No. 23. Centre for Strategic Economic Studies. Victoria University. Melbourne, AU. Retrieved from <http://www.cfses.com/documents/wp23.pdf>
- iParadigms. (2011). Plagiarism Checker | Plagiarism Detection Software from iThenticate Retrieved 7 Dec 2011, from <http://www.ithenticate.com/>
- Kostoff, R., & DeMarco, R. (2001). Extracting information from the literature by text mining. *Analytical Chemistry*, 73(13), 371-379. doi: 10.1021/ac012472h
- Krallinger, M., Valencia, A., & Hirschman, L. (2008). Linking genes to literature: text mining, information extraction, and retrieval applications for biology. *Genome Biology*, 9. doi: Artn S8. DOI: 10.1186/Gb-2008-9-S2-S8
- Laakso, M., Welling, P., Bukvova, H., Nyman, L., Bjork, B. C., & Hedlund, T. (2011). The Development of Open Access Journal Publishing from 1993 to 2009. *Plos One*, 6(6). doi: ARTN e20961. DOI: 10.1371/journal.pone.0020961
- Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. Retrieved from <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>
- Wagner, A. B. (2010). Open access citation advantage: An annotated bibliography. *Issues in Science and Technology Librarianship*(60).

Comment 2

- (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 existing copyright and IP licensing framework provides adequate tools for protecting IP interests. However, since the IP interests of various stakeholders may not always be aligned it is necessary to balance competing interests. In addition, some of the defaults embodied in copyright law and traditional practice need to be avoided to maximize benefits.

One occasionally still sees arguments of the form that the government has no business mandating public access, arguing that such publications are funded by the private sector rather than by the federal government. Such arguments not only ignore the legislative mandate of COMPETES but miss the point on at least three grounds: the vast majority of the costs associated with publication are in the research that it reports and in the nominally-free services such as reviewing that are provided by the academic community. Remaining costs are indeed funded by a variety of sources most notably libraries that pay for subscriptions, but many of those libraries are not usually considered part of the “private sector.” Secondly, it ignores the recent growth in robust gold open access publishing models such as that of PLoS and BioMed Central. Thirdly, such consideration is irrelevant to an IP analysis. Equally arguably the only constitutionally mandated stakeholders with an IP interest are authors and inventors, though the constitutional mandate for copyright is justified by the goal of promoting “the Progress of Science and useful Arts,” implicating the public (and public access) as an important stakeholder. The primary IP interest for present purposes is that of the authors who produce new peer reviewed publications, and their interests are primarily in assuring that others make widespread and immediate use of their work within a framework that discourages misuse such as misquotation or use without attribution. The interests of this group of stakeholders largely align with the interests of consumers of federally funded research results, including other scientists, federal agencies and Congress, university tenure committees, library archives, entrepreneurs interested in commercializing reported discoveries, and the general public. In general, these stakeholders need at a minimum the right to read and copy works both individually and in bulk, and the right to make use of the knowledge contained in such articles to produce new discoveries and new articles.

For these groups of stakeholders the various Creative Commons licenses (Creative Commons, 2011) such as CC-BY are consistent with current U.S. copyright law and provide an excellent basis for effectively meeting the needs of authors and the public. The CC family of licenses have numerous advantages for use with academic works, for example careful handling of moral rights through a “no endorsement” clause applied

to derivative works, and of digital rights management that might otherwise circumvent the author's intention to make the work publicly available. There is clear community agreement that the CC family of licenses are well crafted and of small enough number to provide standardization in license terms, a feature highly beneficial to effective public access by non-lawyers.

Our own reading of typical copyright transfer agreements that authors routinely sign when publishing on traditional journals suggests they often are not well crafted to preserve the rights authors want. For example, they do not adequately preserve moral rights, and often include by reference external publisher websites that are subject to change and sometimes internally inconsistent. It is in the public interest to assist authors in clarifying the rights they actually want to transfer to publishers, but it is worth reiterating that the NIH mandate and related policies are fundamentally rooted in the IP interests that exist before such transfers.

There is currently some discussion in the academic community (McLennan & Malenfant, 2011) about whether a very restrictive license such as CC-BY-ND provides adequate reuse rights to allow robust text mining and computation on the corpus of publications while further protecting author needs for attribution and integrity or whether a license such as CC-BY is needed for effective academic use of archived works; conversely, it may be that an even more open (though in some senses more restrictive) license such as CC-BY-SA is needed for effective widespread use of derivative works created from publications. Given that our understanding of the actual needs of our users is evolving along with the types of uses that are technically realistic, a reasonable compromise strategy would be for federal mandates in 2012 to specify a minimum access corresponding to, or preferably explicitly referencing, CC-BY-ND, but also (a) provide metadata as part of any dissemination that describes the particular license adequately to allow readers and text mining applications to determine their rights, (b) encourage authors to license their work under the most non-restrictive terms they feel comfortable with (either using CC-BY or dual licensing under CC-BY-ND and some other orthogonal license, the way many open source software projects do), and (c) reevaluate the need for more open licenses as the community obtains greater experience with data mining and computational applications.

Apropos of the need for a bare minimum of CC-BY-ND, we observe that academic readers at a minimum need rights unambiguously sufficient to allow uses consistent with current academic practices by researchers at rich universities who can afford subscriptions to journals. Routine practice, for instance, is often to make copies of licensed articles to distribute to students in a graduate seminar. Similarly, needs by the general public are not yet well understood but likely extend beyond minimal access. The first author on this comment is currently undergoing chemotherapy, and observes information usage patterns in his oncologist's office; in that office the doctors do not subscribe to most journals and do not appear even to access relevant articles on PubMed Central even if such access would affect their treatment decisions. Failure to read PubMed Central articles seems largely an issue of convenience; they would be much more likely to do so if an assistant found relevant current research and distributed electronic copies to the doctors in the practice rather than suggest that they go beyond their computing abilities to view PubMed Central directly.

Apropos of CC-BY-ND versus CC-BY, the University of Oregon is an example of a typical academic publisher, with the University Libraries publishing 3 open access scholarly peer reviewed journals. We initially required that our authors license all articles using a CC-BY-ND (no derivative works) license, but are in the process of a transition to consistent use of CC-BY licenses. We have concluded as publishers that the public interest and our own interest as a publisher is only served by granting full reuse rights to readers, and that that license matches the desires of most of our potential authors, thereby improving our chances of publishing high quality work. Commercial publishers whose business model charges subscription fees for access to scholarly works may have a somewhat different set of needs, particularly for works where their business model still requires that they acquire copyright from the authors. In order to ensure the continued value of the publisher's copy, it may be desirable that full reuse rights apply at least to the author's final manuscript version, but that publishers have available the right to impose somewhat more restrictive rules governing the final published version. IP policies must also require that any use of or reference to works include a citation that references the publisher version as the copy of record. Some publishers have argued the need for an embargo period where public access to any version of the author's work is restricted. There is very little data to support the claim that embargoes actually strengthen the financial position of the publisher or influence the willingness of libraries to subscribe to a journal, and clear data that they decrease both visibility of the articles and the benefits of open access in fostering R&D (Houghton, Rasmussen, & Sheehan, 2010). However, it may be that short embargoes can be shown to be needed. If so, a reasonable way to balance interests for new peer reviewed works is to follow the apparently-working NIH PubMed Central model and allow for a short (less than one year) author specified embargo period, with full reuse rights (e.g., CC-BY licenses) applying after the end of the embargo period. Please see Comment 8 below for further discussion. Under the current implementation of the NIH public access rules, legal compliance responsibility falls primarily on the institutions that contract with NIH. It would be very desirable to provide standardized mechanisms to clarify and shift that responsibility to individual authors. One mechanism for doing so would be a small change to the standard contract terms that would require institutions to obtain on the government's behalf from all potentially affected employees a prospective (perpetual but nonexclusive) grant of license to works affected by the policy. Such a license would make explicit that the authors not only commit to depositing a copy of their work but grant to the federal government the right to public display or, as suggested here, the right to sublicense the work to the general public under the appropriate CC license. Such license terms would also be easy for institutions to implement as part of their standard employment contracts.

Additional specific steps that can be taken to protect copyright interests include:

- Enforcement of the standard federal requirements (National Science Foundation, 2011) for including in grant funded publications the detailed acknowledgement of the granting agency, including the award number (which is useful data in text mining and meta-analyses), plus extension of these rules to require that research publications produced by U.S. Government employees

be comparably labeled to make clear that such articles are in the public domain, and that additional metadata be provided;

- Careful documentation and systematic publication as part of publicly accessible article metadata of the IP rights retained by each party in articles;
- Education for authors about their rights and copyright responsibilities;
- A requirement that citations to works deposited in federal archives always include a full citation to the publisher copy of record, including DOI if available, perhaps with the limitation that publishers may waive this requirement;
- Prohibition on the transfer to federal repositories of copies of scholarly articles that contain technological protection mechanisms – under current copyright law defeating a technological protection mechanism is illegal, but clearly imposing such mechanisms would undermine the goals of policies such as the NIH deposit mandate.

A copyright-related policy that should not be implemented is the current PubMed Central prohibition on systematic downloading, which needlessly restricts large classes of use and imposes restrictions beyond those mandated by the copyright or license terms embodied in individual articles.

Another important class of IP interests derives from patent law. In many cases the innovations flowing from federally funded research will result in patents held by researchers, their institutions, or commercial entities. In most cases such rights do not conflict with the goals of a public access policy, though in some they may imply an additional need for a temporary embargo during filing. However, such rights do potentially limit the rights of the public to use results reported in peer reviewed publications. For example, in a case currently being considered by the U.S. Supreme Court (*Mayo v. Prometheus*) it is argued that a reported correlation between two variables and its usefulness in creating a cancer diagnostic limits the right of doctors to even discuss the correlation with patients (Anderson, 2011; Barnes, 2011; Lee, 2011). At a minimum this suggests that authors should be required to be transparent about any patent claims they or their institutions and assignees make associated with research they report, since such claims have an impact on how readers can use articles.

Currently, PubMed Central is in a somewhat awkward position with respect to licensing, since it contains works subject to a wide variety of usage restrictions. Some articles are embargoed and the license granted by the copyright owner to NIH does not even permit public viewing; many articles may be viewed but not copied despite the fact that typical PC viewing software often requires that the web browser make a non-transient downloaded copy of the file in order to display it to the user; other works are released under licenses that allow some copying but no further use, and still others allow a variety of uses. In most cases it is unclear to users what rights they have to individual articles. PubMed Central contributes to this confusion by pointing to the copyright statements included in individual articles,

which often do not note the additional rights that authors and publishers have granted to PMC or the public as part of the deposit process, or to external publisher websites that often provide confusing or contradictory information. Greater standardization of minimal licenses is needed, at the very least to the point where a user can reuse copies of individual articles and where the default absent explicit terms to the contrary is an attribution requirement that meets traditional academic standards for avoiding plagiarism.

References

- Anderson, J. (2011, Dec 08, 2011). Summary of Mayo v. Prometheus Oral Argument. Retrieved from <http://www.patentlyo.com/patent/2011/12/summary-of-mayo-v-prometheus-oral-argument.html>
- Barnes, R. (2011, Dec 7, 2011). Supreme Court has hard time finding an easy test for patents on medical processes, *Washington Post* [online; not available open access]. Retrieved from http://www.washingtonpost.com/politics/supreme-court-has-hard-time-finding-an-easy-test-for-patents-on-medical-processes/2011/12/07/gIQAnEuldO_story.html
- Creative Commons. (2011). Creative Commons [web site] Retrieved 6 Dec 2011, from <http://creativecommons.org>
- 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. Retrieved from <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>
- Lee, T. (2011, Dec. 15, 2011). Medical Mind Control. *Slate Magazine* [online].
- McLennan, J., & Malenfant, K. (2011, December 2, 2011). Getting the rights right: Next SPARC-ACRL forum at ALA announced Retrieved 11 Dec 2011, from <http://www.arl.org/sparc/media/11-1202.shtml>
- National Science Foundation. (2011). *Research Terms and Conditions, June 2011*. National Science Foundation Retrieved from http://www.nsf.gov/pubs/policydocs/rtc/termsidebyside_june11.pdf.

Comment 3

- (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 deposit of works in which the federal government has an interest has a long and successful tradition, starting with the creation of the Library of Congress and the requirement in 17 USC 407 that copyright holders deposit copies of a printed work as part of the copyright registration process. Note that as of 2010 the Library of Congress also has begun an on-demand requirement for deposit of online-only serials (Peters & Billington, 2010). Such deposit demonstrates the benefits of federally maintained centralized approaches, and addresses some issues of long term archival, but does not by itself meet the needs of public access.

It is important to distinguish among a variety of uses for corpora of peer reviewed manuscripts and to develop approaches that support the variety of uses, where different uses may require multiple deposit copies and where only some uses require centralized custody associated with the funding agency.

For long-term stewardship of federally funded research publications, the federal government is the appropriate long-term custodian. We believe that the preferred approach to maintaining public access is either a single managed repository – perhaps a somewhat expanded version of PubMed Central -- serving multiple federal agencies or is a distributed repository where multiple agencies each maintain copies of publications funded by that agency but do so in a software context that appears to the user as much as possible as being a single federated repository with uniform access and user interface characteristics.

The federal government has a continuing interest in making such works permanently available, and is the only current player with an interest in making the full corpus of federally funded works publicly available, in providing tools for using the collection as a whole, and in assuring that new services and products can be built from publicly funded information. Other private players including universities with their institutional repositories, disciplinary societies with their discipline-specific archives, and commercial and noncommercial publishers with their journal-specific collections have not in general demonstrated an ability to scale at low cost or to provide broad public access the way, for example, PubMed Central has.

Commercial entities and non-profit volunteer efforts in particular are at risk of being unable or unwilling to provide long-term (multiple-decade) access as their business models change or organizations go out of business. One need look no further than recent news stories such as the November 2011 changes in Amazon's Penguin e-book lending program (Van Camp, 2011) to realize that commercial entities have short term business incentives that may conflict with long term access. Similarly, there are lingering concerns about quality control in commercial journals (Grant, 2009).

Commercial entities also may have a conflict of interest when it comes to providing access that could allow competitors to develop new services based on an archive. Entities that have an interest in only a centralized “dark archive” in particular are not in a position to provide a viable solution since such an archive neither meets archival needs (regular access and use are vital to maintaining archival veracity) nor the public interest in access.

Centralized repositories make it easy to find materials, easy to curate, and easy to provide standard appearance, reliability, and quality. The benefit of centralization is particularly notable as we consider use of the repository by ordinary citizens who need simple access in order to find the publications that contain the information they need, and hence a repository that does not require the mediation of a librarian to use it effectively. At a minimum, the end user needs to see a single web site as the point of initial access and consistent procedures for actually viewing works of interest. Although some of the benefits of centralization can be achieved by federated search, such search is much easier to implement if multiple repositories all share the same management. In any case, federated search is a finding aid rather than an access tool, and does not solve the potential problem of multiple user interfaces and policies that can easily frustrate access even when the user has a link to the article desired.

It should be noted that centralized management and policy setting does not preclude outsourcing or cloud-sourcing if such outsourcing makes sense financially and suitable contractual arrangements can be established. For example, the federal government might contract for storage and compute resources from a commercial provider such as Amazon or Google or a consortium of publishers. However, any such outsourcing must clearly establish a set of stringent requirements for public access and a standardized interface; by far the easiest way to accomplish this would be to outsource only the back end with the expectation that the applications being run and the ingest policies were those currently in use for Pub Med Central. Even for outsourced data storage contracts must be clear that the corpus of copies is the property of the federal government, that the contractor is an agent of the federal government and is required to meet any legislatively mandated regulations that apply to the agency, and that adequate availability, security, and termination guarantees are in place. Under the last issue, for example, there must be a clear migration path for recovery of the data in the event that a provider is no longer able to meet its obligations. Outsourcing to multiple providers (for example, to individual publishers) would likely incur dramatically increased costs in providing standardized guaranteed access. Overall, our expectation is that it is unlikely that such outsourcing would prove cost effective.

However, it is also vitally important that other players be able to mirror data and that copies of publications be available across multiple public and private sources. As research in library archival has demonstrated, one of the most effective strategies for long-term preservation and access is based on principles like LOCKSS (“lots of copies keeps stuff safe”) (Reich, 2008) that distribute risk across multiple servers, organizational entities, and archival approaches [it is notable, however, that this is not an alternative to a federally maintained repository; current market attempts to implement LOCKSS itself have not yet been fully successful (LOCKSS Assessment Team, 2011)]. In addition, we anticipate collections that meet specialized needs both

for access (e.g. very high bandwidth access) or content (e.g. collections that include both publications and associated data sets or derivative works). Just as with populations of biological individuals, some genetic variation within the population of archives makes it more robust in the face of environmental stresses and new demands.

For example, we do not anticipate that federal agencies should maintain custody of research products that are not federally funded, but in many cases such research will be released for open access and will be part of the ecosystem that is viewed (and manipulated as a whole) by future researchers and the public. The situation is analogous for a university institutional repository, which may have a mandate to collect and make publicly accessible all peer reviewed publications generated by its own faculty (Brody, 2011) but not related research; in other cases however, such an entity would want to collect all or a subset of the articles in a federally maintained archive (for example, all articles related to a particular research area that the university is investing in, or all articles produced with funding from a federal agency that political scientists at the university are studying). Similarly, we anticipate that publishers will usually wish to maintain the copy of record for publications that have appeared in their journals, and that it will be important that other archives (some of which may store preprints, derivative works, or copies that contain formatting changes relevant to future scholars) provide clear links to the copy of record. In cases where publishers have adopted the DOI standard, this notion of a copy of record is easily mapped to the particular version that the DOI resolves to, and is analogous to the “best edition” as used for copyright deposit purposes and defined in 17 USC 101 and detailed rulemakings (Peters & Billington, 2010, pp. 3868-3869).

Mirroring requires appropriate access licenses to the texts in source repositories, clearly specified conditions for public accessibility and long term preservation, and technical solutions such as the OAI-PMH (Lagoze & Van de Sompel, 2011) that allow bulk harvesting of content. In addition to providing a centralized repository for federally funded articles, federal agencies need to consider standards for archive interoperability and need to invest in research and development of tools for effective mirroring and archive description. One important modification that needs to be made to present policies is relaxation of the gratuitous limit in PubMed Central on bulk copying. Such copying may be restricted by licenses to specific articles, but should be freely permitted if the article’s usage license permits it.

References

- Brody, T. (2011). ROARMAP: Registry of Open Access Repositories Mandatory Archiving Policies Retrieved 7 Dec 2011, from <http://roarmap.eprints.org/>
- Grant, B. (2009, 30 April 2009). Merck published fake journal. *The Scientist: Magazine of the life sciences*.
- Lagoze, C., & Van de Sompel, H. (2011). Open Archives Initiative Retrieved 7 Dec 2011, from <http://www.openarchives.org/>
- LOCKSS Assessment Team. (2011, October, 2011). Final Report of the 2CUL LOCKSS Assessment Team, Cornell University Library & Columbia University Library Retrieved 7 Dec 2011, from <http://2cul.org/sites/default/files/2CULLOCKSSFinalReport.pdf>

Peters, M., & Billington, J. H. (2010, January 25, 2010). Mandatory Deposit of Published Electronic Works Available Only Online [revises 37 CFR 202]. *Federal Register*, 75 (15), 3863-3870.

Reich, V. (2008). LOCKSS Retrieved 7 Dec 2011, from <http://www.lockss.org/lockss/Home>

Van Camp, J. (2011, November 22, 2011). Penguin halts Kindle library lending: Will more publishers disable the feature? *Digital Trends*.

Comment 4

- (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 class of partnerships is with research universities, particularly university libraries. Such libraries have extensive experience in preservation and archive infrastructure, and tend to have very long time horizons; the oldest private university libraries in the U.S. predate the federal government by more than a century. Some of the best non-Federal examples of successful archives, e.g. ArXiv. represent public-private (specifically Federal-university) partnerships.

Universities are also increasingly becoming active publishers of open access journals, and have a role to play in that regard. For example, the University of Oregon is fairly typical of major university libraries in having established a small program to publish very specialized open access journals, with a current focus mostly on humanities and social sciences. Such journals fill an important niche that is currently largely ignored by traditional commercial publishers. They, along with commercial open access journals such as those from BioMed Central, provide a model that avoids many of the intellectual property issues created by traditional publishers, since the license that the journal requires does not conflict with federal requirements for deposit in an archive such as PubMed Central. However, partnership is needed to develop better tools for automating such deposit along the lines of NIH submission “method B” (National Institutes of Health, 2011) that are specific to the (often open source, e.g. OJS, Drupal, Annotum, etc.) publishing software that university publishers typically use.

Numerous other examples of successful public/private partnerships also provide models. For example, consider the NSF International Children’s Digital Library, one of many resources provided by federal agency / private collaborations as part of the FREE website (Federal Resources for Educational Excellence, 2011). In many cases such resources can be greatly enriched by references to the peer reviewed literature and examples of the science that informs more popular presentations of information. References

Federal Resources for Educational Excellence. (2011). International Children's Digital Library Retrieved 18 Dec 2011, from http://free.ed.gov/resource.cfm?resource_id=2187

National Institutes of Health. (2011). Submission Methods Retrieved 10 Dec 2011, from http://publicaccess.nih.gov/submit_process.htm

Comment 5

- (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?

In this comment we primarily address the second and third questions posed. We believe that rich and well documented metadata, combined with implementation of current federated search protocols and with research on such protocols, will contribute to interoperability.

Many of the most important core metadata fields are already captured in standards such as the NLM-XML archiving tag set (NCBI, 2010), though are not always mandatory. Examples include standard citation information such as author names, article title, journal name, and so on. The archiving tag set provides a good tool for describing and ingesting arbitrarily structured journal articles, but fails to distinguish well between descriptive information contained within the article and metadata associated with it, and does not mandate a set of minimum core metadata. We recommend that the [PubMed Journal Article DTD](#) (NCBI, 2011) be adopted and extended as the target for metadata associated with new journal articles.

Although the archiving tag set is a good starting point that balances consistency with pragmatics, neither it nor the Journal Article DTD prescribes a sufficient minimal set of metadata elements to guarantee that a user can create a reference, but being able to create a reference to an article seems to be a minimal bar that defines metadata that most users are likely to need. Since citations are not always in NLM format, the required metadata needs to be rich enough to allow generation of minimal NLM, APA, Chicago, and MLA, references, suggesting that any datum that is common to and required in at least 3 of these should be considered mandatory metadata.

In addition, several article-level metadata fields are not widely standardized or encoded but are very important for effective use and should be included in a minimal mandatory set.

One specific metadata element that is needed for individual peer-reviewed publications is a precise statement of the license that a document is released under and who it applies to in a format that can be automatically processed by text mining software. The archiving tag set <permissions> entity is a good start in this direction and more flexible and detailed than the Journal Article <copyright> entity, but is oriented towards documenting restrictions and so for instance would not be used to affirmatively assert that a work was in the public domain because the author was a federal employee.

Another vital field is versioning information that describes the relationship of a particular document to the published version. At a minimum a simple controlled vocabulary that specifies whether a version is an author's preprint, the author's final

manuscript version, a copy of the publisher's copy of record (and if so, as of what date and in what ways it has been modified, since even formatting variations may prove relevant to future researchers), a derivative work, etc. We assume that in most cases publishers will wish to maintain the copy of record for publications that have appeared in their journals, and will associate with that copy a revision history (for example, tracking status if by mischance the article needs to be withdrawn). It is important that other archives provide clear links to the copy of record, for example by providing a DOI if available. Some but not all of this information can be encoded in the Journal Article <pub.status> and <eLocation> entities.

Another article-level metadata element that is needed is information as to what data sets, samples, case studies, and supplementary documents the conclusions rely on. Knowledge of the data sets used is particularly important in minimizing duplicate publication and in facilitating meta-analyses. The APA Publication Manual in Psychology describes numerous reasons to eschew duplicate publication, and states that: "As multiple reports from large-scale or longitudinal studies are created, authors are obligated to cite prior reports on the project to help the reader understand the work accurately... It is also important to make clear the degree of sample overlap in multiple reports from large studies" (American Psychological Association, 2010, pp. 13, 15). Reporting in systematic fashion the particular grant(s) that funded the research is a step in the right direction, but since many grants generate multiple studies and data sets the particular data set used needs to be reported in a fashion that allows machine analysis.

In addition to enhancing the core metadata fields, federal agencies need to improve their archives by clearly distinguishing metadata (what NLM-XML generally considers front matter) from the publication itself, making metadata available through both the repository web interface and through an API interface. One step that would make it more likely that metadata is consistently provided would be to move towards consistent upload of articles in NLM-XML or as an OAI-ORE. We suggest recommending submission of the "best format" (Peters & Billington, 2010, pp. 3869-3870) as now required for Library of Congress deposit of online-only journals. In addition, we recommend that all submissions be required to include a minimal set of metadata, either mechanically derivable from the article itself (as for instance, with NLM-XML), or separately provided by the submitter, and that the ingest software first attempt to fill out all fields based on the article itself, then ask the submitter for additional information as needed.(NCBI, 2010)

References

- American Psychological Association. (2010). *Publication Manual of the American Psychological Association* (6th ed.). Washington, D.C.: American Psychological Association.
- NCBI. (2010, June 6, 2010). Archiving and Interchange Tag Set Retrieved 14 Dec 2011, from <http://dtd.nlm.nih.gov/archiving/>
- NCBI. (2011, 2011-10-11). PubMed Journal Article DTD Version 2.6 Retrieved 14 Dec 2011, from <http://www.ncbi.nlm.nih.gov/entrez/query/static/PubMed.dtd>
- Peters, M., & Billington, J. H. (2010, January 25, 2010). Mandatory Deposit of Published Electronic Works Available Only Online [revises 37 CFR 202]. *Federal Register*, 75 (15), 3863-3870.

Comment 6

(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?

Five strategies seem particularly appropriate for maximizing benefit while minimizing costs:

1. Standardize licenses to include the maximum possible reuse rights (such as those embodied in Creative Commons CC-BY licenses) for the public;
2. Ensure consistency (and to the extent possible, temporal stability) in requirements and processes, particularly across multiple agencies;
3. Take advantage of current software technologies to minimize cost, both to agencies and to depositors;
4. Provide simplified interfaces to make mandate compliance as easy as possible.
5. Encourage commercialization, but only in domains where competition exists to avoid monopoly profits; in all cases of commercialization weigh the benefits of standardization and the overall cost efficiencies of providing services centrally or of contracting for services with non-profit entities such as universities.

Under “standardization,” please see this submission’s Comment 2, noting that standardization of license terms also contributes to consistency and simplicity. Under the “consistency” umbrella, it is particularly desirable that all federal agencies to which a mandate applies use a standardized set of criteria to determine what materials need to be deposited and how deposit should occur. This is obviously relevant to researchers who may have funding from multiple federal agencies, but is also important because it reduces support costs, e.g. for documentation and for training of support staff (both at federal agencies and within institutions that assist their researchers with compliance). For example, at the University of Oregon our total post-award sponsored project administration staff comprises 8 FTE despite a moderately large number of extramural projects. Minimizing the complexity involved in monitoring regulatory mandates and advising/assisting PIs on compliance is critical to ensuring not only that cost of administration is contained and compliance reviewed quickly, but that mandates are actually followed.

Under “current technology,” one observation is that agencies have an interest in encouraging the development of new tools to make compliance simpler. One specific example that relates to consistency is that authors may need to deposit in multiple repositories based on multiple simultaneous mandates. Different components of a University Of Oregon research project might be funded by Wellcome Trust and NIH, plus some departments at the University of Oregon have departmental deposit mandates into our institutional repository. Why should the author have to go through 3 processes to meet three sets of deposit requirements? It should be

possible using protocols such as SWORD for the author to automatically deposit an author's final draft in multiple repositories simultaneously. It should be possible as part of the deposit process to provide automated notification to interested parties such as an institutional grants management system or institutional faculty profile databases. It should be very easy to extract in a standardized output-neutral format such as Citation Style Language (CSL) all citation data for a single researcher or an entire institution. The goal should be an easy process for moving citation data to free and commercial bibliography management and publishing tools such as bibtex, endnote, or mendeley. Standardization on a powerful and open citation format would encourage standardization within the bibliographic software industry. Under "simplicity" one very useful approach would be for an agency to employ a human interface design consultant to conduct tests and identify issues that make it difficult for real users to comply with the mandates.

Comment 7

(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?

Although it is desirable for all peer reviewed publications resulting from federally funded research to be widely available, the policies under which they are made available may need to differ depending on the type of material. For example, the interests of authors in peer reviewed journal articles where they generally receive no financial remuneration may differ from those of authors of commercial textbooks. Another plausible next step in expanding types of publications would be to peer-reviewed conference presentations. One issue here is that such presentations show a much wider range of formats than peer reviewed journal articles, and so present more technical challenges. If a conference presentation is a multimedia presentation using Mathematica (rather than a text with an incidental use of a demonstration), must deposit consist of both the Mathematica workbook and a purchased copy of the version of Mathematica required to display it? As conferences increasingly become hands-on the problems of adequately capturing the “presentation” get even greater; what about that presentation that allows audience members to use a Microsoft Kinect to explore a virtual human?

Given the complexities involved, we at the University of Oregon believe that efforts at this time to extend public access policies to include book chapters or conference proceedings within a single public access policy would increase complexity and make it harder to achieve the more important goal of widespread access to peer reviewed journal articles.

However, one class of materials that are particularly important to make available are research instruments, survey forms, and protocols. Journals are currently quite inconsistent in their requirements for publication of such supplementary materials and tend to be driven by no-longer-relevant concerns about costs associated with hardcopy distribution, but in a number of disciplines the relevant professional organization makes clear that it mandates public access to such materials. For example, in Psychology the APA Ethics Code requires that researcher retain and make available to other researchers not just data but also supplementary materials including “[o]ther information related to the research (e.g., instructions, treatment manuals, details of procedures, code for mathematical models reported in journal articles) ...; such information is necessary if others are to attempt replication.” (American Psychological Association, 2010, p. 12). Given the central role of replication in the progress of science, mechanisms should be established to routinely and systematically collect such materials and make them available online as supplements to the text of the peer reviewed article.

Although not peer reviewed publications per se nor adequate substitutes for such, research progress reports and final reports are potentially very useful in contextualizing grant-funded scholarly articles. Such reports are not a substitute for the peer reviewed articles that document results, and indeed one notes that it is routine for reports to reference the publications. However, the grant reports do provide useful auxiliary information and in particular may be helpful in determining

relationships between multiple peer reviewed publications flowing from a single grant. Federal agencies should take steps to ensure not only that such reports are freely available to the public in timely fashion, but that tools are available to cross-walk between those reports and corresponding articles.

References

American Psychological Association. (2010). *Publication Manual of the American Psychological Association* (6th ed.). Washington, D.C.: American Psychological Association [not available open access].

Comment 8

(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?

Embargoes of any length impose 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. Houghton et al estimate (Houghton, Rasmussen, & Sheehan, 2010, p. 8) that "a six-month embargo reduces the returns [in benefits from increased R&D] by around \$120 million (NPV)." Similarly, increasing evidence (Wagner, 2010) indicates that open access in general increases citation rates for peer reviewed publications, suggesting that embargoes may have a negative effect on readership of the embargoed journal. For example, data from a sample of Chinese journals (Cheng & Ren, 2008) indicates that journals in their sample that had embargoes (delayed vs immediate open access) experienced a 20% citation disadvantage (1-1.26/1.57).

The argument in support of embargoes is of course that such embargoes putatively encourage academic libraries to continue to subscribe to the journals. We know of no studies that directly examine this hypothesis or of documented examples of journals whose financial viability has been significantly damaged by public access policies such as the NIH public access mandate. It would be very useful to be able to consider empirical data, preferably in peer-reviewed economics journals, that facilitated measurement of this possible relationship and the more general question of the economic effect of publication embargoes.

Consensus in the library community seems to be that cancellation decisions are budget related, not access related. Anecdotal evidence from our University of Oregon journal cancellation projects does not show that embargoes themselves (including open access moving walls and embargoes on availability of full text within licensed databases) have had any substantial influence discouraging journal cancellation decisions, particularly given the constraints of publisher bundling of multiple journals in a single subscription. One UO library subject specialist and department head, however, goes further and reports that "I do take embargoes into account. I tend to view an embargo such as a 1 year embargo on access to full text in a licensed database negatively" (Frantz, 2011).

It is important to note as well that the benefit of embargoes accrues only to those publishers who use particular economic models. The rapid growth of the open access journal market, where subscription fees are replaced by author fees (or possibly by APCs supplemented by institutional subsidies, advertising, and other revenue streams) have very different needs and no clear benefit from embargoes. It may be in publishers' interests to move to APC funding models, which would

completely vitiate any need for embargoes. On the other hand, if journals move to advertising-driven open access models, then their publishers may continue to have an interest in embargoes.

In examining the length of embargo periods, it is important to note that a maximum embargo period of six months is becoming the norm among biomedical research funders, with NIH an outlier at allowing 12 months. (Carlson, 2011).

The prompt specifically addresses the question of whether delay should be different in different disciplines. Although different disciplines and even narrow but nearby subdisciplines clearly show different patterns of article usage – there is a much stronger premium on early access to preprints and published articles in rapidly changing subdisciplines and in those STEM fields where research results often lead directly to commercialization – it is not clear whether such differences ought to be considered even if they imply differential economic effects. For one thing, the effects on subscription-based publishers should be balanced by the economic benefits of widespread access to the economy as a whole. For another, there are substantial costs in increased administration complexity and user confusion as soon as one allows differential embargoes.

Our overall impression at the University of Oregon is that the PubMed Central model, with variable embargoes from 0 to 1 year, is working adequately, and does not need to be changed at this time. However, we also believe that there is reasonably strong evidence that a standard maximum embargo period of 6 months would be preferable. It is also important that embargo periods be established and approved by the individual author within those parameters, since it is the author who is granting to the federal government the rights allowing public access, and the author who best understands the negative impact of an embargo on rapid readership for his or her work.

References

- Carlson, D. (2011). *Open Letter to Dr. Francis Collins, Director, National Institutes of Health*. SPARC. Washington, DC. Retrieved from <http://www.arl.org/sparc/bm~doc/nih-sparc-final-11-0406.pdf>
- Cheng, W. H., & Ren, S. L. (2008). Evolution of open access publishing in Chinese scientific journals. *Learned Publishing*, 21(2), 140-152. doi: 10.1087/095315108x288884 [not available open access]
- Frantz, P. (2011, 12 December 2011). [personal communication].
- 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. Retrieved from <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>
- Wagner, A. B. (2010). Open access citation advantage: An annotated bibliography. *Issues in Science and Technology Librarianship*(60).

Subject: public access of research

Date: December 21, 2011 5:02:04 AM EST

Dear Sirs,

Research achievements of public-funded research must be available to society free of restrictions. The scientific ethics dictate that benefits from science is the right of the public and not the self-oriented privileges of few researchers who were educated , trained and supported by tax-payers money.

As an independent researcher I find it often difficult to retrieve information that is necessary for progress of science

Björn Hammarskjöld
M.D., Ph.D:
Independent researcher

Subject: Public access to results of funded research

Date: December 21, 2011 9:49:22 AM EST

Under the current system research a significant part of the results of research funded by the US government is not accessible to the public due to the cost barrier imposed by publishers. While this may have made sense at a certain point in time when publishers did make a significant contribution to the work by helping in its dissemination, that is no longer the case in the current world where the cost of distributing information has dropped to near zero. Under these conditions, the overwhelming value of the published research resides in the research itself, payed by our tax dollars, with very little value added by the publishers.

**Response to the White House Office of Science and Technology Policy's
*Request for Information: Public Access to Peer-Reviewed Scholarly
Publications Resulting From Federally Funded Research* submitted by
the Canadian Association of Research Libraries.**

20th of December, 2011

CARL is the leadership organization for the Canadian research library community. The Association's members include Canada's major academic research libraries.



The Canadian Association of Research Libraries (CARL) welcomes the opportunity to submit a response to the White House Office for Science and Technology Policy's Request for Information: Public Access to Peer-reviewed Scholarly Publications Resulting from Federally Funded Research. CARL is the leadership organization for the Canadian research library community. The Association's members are the 29 major academic research libraries across Canada together with Library and Archives Canada, the Canada Institute for Scientific and Technical Information (CISTI), and the Library of Parliament.

Across the globe, demands are growing to make publicly funded research more readily available to academia, civil society and industry. Any further steps the White House takes to ensure public access to the results of publicly funded research will lead to more efficient linkages between science and innovation and are highly commendable. While there are already open access policies in place at funding agencies in Canada and across the world, a US policy requiring public access to federally funded research results would likely have a significant global impact and pave the way other countries to follow. CARL's response to the Request for Information is as follows:

(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?

A key step that agencies can take to stimulate new markets related to peer-reviewed publications is to implement policies requiring free and unrestricted online access to the results of federally funded research. Such policies would enable the development of numerous value-added services to be built on top of this content. While it is impossible to envision all of the types of services that may be created, we can anticipate a variety of tools that will evolve as a result-such as search engines, disciplinary portals, data mining, as well as repository services.

In addition, small and medium sized businesses represent a significant portion of library users in Canada and elsewhere, reflecting the importance of scholarly literature to business development. Access to the full corpus of scholarly articles would contribute to the development of new products and services based on the results of 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?

Open access policies will result in greater usage of the research results and will contribute to maximizing the government's investment in research. This statement is validated through a large number of studies that have found that open access increases usage and citation of peer-reviewed articles by researchers.^{1 2} Statistics from PubMed Central, the open access repository for biomedical literature, shows impressively high levels of usage of its open access papers.

¹ "The effect of open access and downloads ('hits') on citation impact: a bibliography of studies". (2011) <http://opcit.eprints.org/oacitation-biblio.html>

² Philip M Davis, Bruce V Lewenstein, Daniel H Simon, James G Booth and Mathew J L Connolly. (2008) Open access publishing, article downloads, and citations: randomised controlled trial. BMJ 2008;337: a568. <http://www.bmj.com/content/337/bmj.a568>

Between March 2008 and March 2010, for example, the monthly number of articles retrieved from PubMed Central doubled from 10 million to 20 million.³

The benefits of open access policies are not limited to the research community; industry, public policy and the general public all stand to benefit and use the information contained within the peer-reviewed literature. The majority of users of PubMed Central are from outside the research community. According to David J. Lipman, Director, National Center for Biotechnology Information the users of PubMed Central “represent a mix of people from the education and business sectors, as well as private citizens.” He estimates that “approximately 25% of our users are from universities, 40% are private citizens or those using personal Internet accounts, and 17% are from companies (the remainder consists of government users or others). These kinds of numbers support the notion that PubMed Central has become a broad-based repository for researchers, students, clinicians, entrepreneurs, patients and their families.”⁴

In addition, economic analyses have also shown that the open access model will cost less at the national level than the existing subscription-based system, leading to a more efficient and productive system for disseminating and exploiting the results of publicly funded research.^{5 6} In particular, these studies have found that the benefits of open access via repository self-archiving would result in the greatest cost savings and accrue significant benefits.

What are the relative costs and benefits of such policies?

There are start-up and ongoing costs associated with open access policies, which include repository development and management, monitoring compliance, and so on. However, these costs represent a relatively small percentage of the large public investment already being made in research. The start-up costs for PubMed Central “were about \$500,000. Annual operating costs for the system, including ingest of articles, refinement of the submission system and search tools, staffing of a help desk and a central coordinating office for NIH, are approximately \$3.5-\$4.0 million per year. This represents a small fraction of NIH’s budget authority of more than \$30 billion per year.”⁷

³ Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>

⁴ Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>

⁵ John Houghton, Bruce Rasmussen and Peter Sheehan. (2010) “Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs”. <http://www.arl.org/sparc/publications/papers/vuFRPAA/index.shtml>

⁶ John Houghton. (2009) “Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark”. <http://www.knowledge-exchange.info/Default.aspx?ID=316>

⁷ Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives.

In addition, many universities in the US (as elsewhere) already have institutional repositories.⁸ These repositories are usually managed by the university library and have dedicated staff to support the deposit of research articles.

As mentioned earlier, economic analyses have found that there are significant overall savings in an open access model. Conservative estimates suggest such policies would result in an estimated five-fold increase on the return on public funds invested in research. Specifically, the benefits of an open access policy, such as has been implemented by the NIH, are estimated to be at least eight times larger than the costs. Extending this type of policy to all US scientific agencies has been estimated to result in a net value of \$1.5 billion, of which an estimated 60% would accrue directly to the U.S. economy.⁹

Open access policies for research articles are directly aligned with the Open Government initiative which is seeking to redress a culture “ where information is locked up, taxpayer dollars disappear without a trace”¹⁰. Other key benefits to these policies are that they will contribute to a more transparent government, the development of scientifically sound policy decision-making, as well as serving to increase federal agency accountability.

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

Full and open access to the results of research as published in peer-reviewed scholarly/scientific journals is the ideal type of access. Any restrictions that place limits on the usage of the results of scientific research will dilute the impact of the policy. In order to derive additional benefits via computational analyses and value added tools, open access articles should be free to be crawled and indexed by search engines.

(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?

It is possible to implement public access policies for federally funded research while fully respecting and working within current copyright frameworks. Creative Commons licenses could

United States House of Representatives. <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>

⁸ 2 Clifford A. Lynch and Joan K. Lippincott, "Institutional Repository Deployment in the United States as of Early 2005," D-Lib Magazine 11, 9 (2005), <http://www.dlib.org/dlib/september05/lynch/09lynch.html>

⁹ Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>

¹⁰ About Open Government. (2011). <http://www.whitehouse.gov/open/about>

be used in this context which would allow authors and publishers to retain their copyright but mark their work with the re-use rights they feel are most appropriate, while still receiving credit for their work.

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?

CARL strongly advocates for the broadest possible access to knowledge in all fields, and would recommend against the establishment of policies that place unnecessary restrictions on access or reuse of content, such as policies which include lengthy embargo times, or allow “read only” access to scholarly 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?

The most effective approach (centralized or decentralized) will vary depending on current practices and infrastructure in any given sector. For physics and biomedical literature, for example, which already have established traditions for depositing into discipline-based repositories, the centralized approach may be most appropriate. For other fields, there are significant benefits to adopting a decentralized approach, making use of the existing institutional repository network based at the universities. University-based repositories already have the capacity and mandate to undertake the long-term preservation and management of scholarly knowledge. The most efficient approach would be to extend the current distributed network environment, which is composed of both institution and discipline based repositories. While it may be easier to monitor compliance via centralized repositories, there are ways to accommodate funders’ need to monitor compliance via metadata for specific research grant numbers.

Regardless of approach, it is critical that open access repositories adhere to common standards and metadata that will ensure article level interoperability across repository and across national boundaries. The research enterprise is truly global and it will be critical that US-funded papers are interoperable with the global corpus of research knowledge. The Confederation of Open Access Repositories (COAR), a large international organization representing repository organizations (of which CARL is a member) states, “The real value of repositories lies in the potential to interconnect them to create a network of repositories, a network that can provide unified access to research outputs and be (re-) used by machines and researchers.”¹¹

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?

¹¹ Confederation of Open Access Repositories. (2011). “The Case for Interoperability for Open Access Repositories”. http://www.coar-repositories.org/files/COAR_Interoperability_Briefing.pdf

The federal government has an important role to play in ensuring that the nation's research outputs are preserved and available into the future. Private industry, which is motivated by profits, has no incentive to preserve content which it can no longer "sell" for access. In addition, there is a risk that commercial enterprises engaged in preserving published content would change business models and go out of business.

A distributed approach that includes a broad range of players including research libraries and other long-term stable institutions may be most likely to succeed. In addition, some redundancy is desired from a preservation point of view – especially with born digital content – by way central, subject and institutional open digital repositories as to where content is archived and accessed.

(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 terms of storage, organizations such as Portico, LOCKSS, and DuraCloud are working with publishers and libraries to assist with the preservation of digital content. Indexing and abstracting services already provide value added services as do journal hosting and repository development services. Indeed, there are numerous models for public private partnerships that should be encouraged with the proviso that these services do not inhibit the easy and open access to research articles.

(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?

Interoperability at the article level is crucial for the discovery and development of value added services, such as harvesters, which are used to aggregate scholarly papers. Metadata should facilitate both reuse and analysis of published research including machine-readable and machine-interoperable analyses. While it is likely that metadata standards will evolve, we recommend for now the adoption of Open Archives Institute Protocol for Metadata Harvesting (OAI-PMH), which is the international standard currently being applied across repositories. OAI-PMH will ensure a minimum level of interoperability across the global corpus of research literature.

Repository interoperability could also be further enabled and enhanced by the adoption of a few additional protocols such as the Open Archives Initiative Object Reuse and Exchange (OAI-ORE), which serves to facilitate content sharing among systems and the Simple Web-service Offering Repository Deposit (SWORD) that permits researchers to deposit an article through a single interface while also routing the same item to multiple repositories. As well, the Confederation of Open Access Repositories (COAR) recommends the use of the Resource Description Framework (RDF) that expresses digital objects relationships rendered in a machine-understandable way "allowing machines to create sophisticated services over the global representation of knowledge

distributed across repositories and other systems, to make cross-discipline connections, and to combine disparate findings to arrive at new insights.”¹²

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 in the interest of all stakeholders to ensure that there is widespread use of standards and metadata. Agencies should work with institutions, universities and, publishers and repository communities to ensure that the appropriate metadata standards are adhered to.

(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 system as it is now is simply not sustainable. Journal prices continue to increase at rates far above the standard of living, to the point that even the most well-endowed research library cannot afford to provide access to all of the content requested by its faculty and students. The implementation of an open access mandate would significantly reduce the systemic costs of access (as discussed earlier in response to question 1).

One way of minimizing the burden for all would be to ensure that there is consistency in the policy language across all agencies. Uniformity in article deposit requirements and procedures across all funding agencies will reduce costs and confusion and ensure a high rate of compliance with open access mandates. In addition, there are a growing number of tools being developed internationally to minimize efforts for depositing articles into repositories, such as the SWORD protocol (discussed above). Agencies can further encourage the development of additional tools and services for depositing and act as facilitators for publishers and repository organizations to work together.

(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 results of publicly funded research in all forms including monographs, conference proceedings and research data should be covered by public access policies. Making all of these products of research available would contribute greatly to maximizing the many billions of dollars that the United States invests each year in research. However, given the different infrastructure and requirements necessary for collecting and providing access to these different types of research output, it may not be appropriate to include all of these under a single policy.

¹² Confederation of Open Access Repositories. (2011). “The Case for Interoperability for Open Access Repositories”. http://www.coar-repositories.org/files/COAR_Interoperability_Briefing.pdf

(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?

In order to maximize the benefit of open access policies, CARL supports the shortest possible embargo periods from 0 to 12 months, maximum. Publishers call for embargo periods as a way of protecting their subscription base. However, there is no evidence that archiving in open access repositories has a negative effect on journal subscription rates. For example, arXiv, the online repository for papers in physics and other fields, has been providing open access to a very high percentage of papers in those fields, but has not resulted in journal cancellations. In fact, the established journals in those fields have continued to thrive.

Canadian Association of Research Libraries (CARL)
600-350 Albert Street
Ottawa, Ontario, Canada K1R 1B1
Telephone: 613.482.9344
E-mail: info@carl-abrc.ca
Website: www.carl-abrc.ca

President
Thomas Hickerson

Executive Director
Brent Roe



University Library Member Institutions

University of Alberta
University of British Columbia
Brock University
University of Calgary
Carleton University
Concordia University
Dalhousie University
University of Guelph
Université Laval
University of Manitoba
McGill University
McMaster University
Memorial University of
Newfoundland
Université de Montréal

University of New Brunswick
University of Ottawa
Université du Québec à Montréal
Queen's University
University of Regina
Ryerson University
University of Saskatchewan
Université de Sherbrooke
Simon Fraser University
University of Toronto
University of Victoria
University of Waterloo
University of Western Ontario
University of Windsor
York University

Subject: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Date: December 21, 2011 10:51:56 AM EST

I am a federally funded biomedical researcher working at a research oriented medical school. Therefore my answers reflect my own experience publishing NIH 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: I believe that the results of all federally funded research should be freely and readily available to the public. The current NIH policy of making peer-reviewed articles available to the public after 1 year tries to strike a balance between providing accessibility to the public and maintaining an economic incentive for publishers. However I personally believe a full year delay for public access is excessive (although I understand the logic behind this delay). The problem is that a one year delay produces an uneven information playing field for the american public during the most critical period for acting on the results of federally funded research.

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: I believe that it is essential that the government maintain a centralized repository of all peer-reviewed publicly funded research. With the rise of digital-only journals there is a real risk that failure of any commercial publisher could result in the loss of important research results that had been generated with public funds.

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: It is my impression that commercial publishers are currently leveraging their archives to force libraries into financial arrangements that are one sided and unjustifiable. This may be good business for the publishers but it is predatory, especially since most of the content that libraries are being forced to buy (so they can continue to do their job) was paid by taxpayers. This is unacceptable. I suggest that any embargo period be made conditional to the release to the public domain of the publisher's archives and include a sunset clause for any

embargo period. agreed upon

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: Yes

Thank you for the opportunity to comment on this important issue.

Rodrigo Andrade, PhD
Professor
Wayne State University School of Medicine

TO: Office of Science and Technology Policy publicaccess@ostp.gov
FROM: American Library Association and the Association of College and Research Libraries
RE: Recommendations on Public Access to Peer-Reviewed Scholarly Publications Resulting
From Federally Funded Research
DATE: Wednesday, December 21, 2011

The American Library Association (ALA) and the Association of College and Research Libraries (ACRL) write in response to the request for information issued November 3, 2011, by the Office of Science and Technology Policy (OSTP) regarding 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. ALA is a nonprofit professional organization of more than 63,000 librarians, library trustees, and other friends of libraries dedicated to providing and improving library services and promoting the public interest in a free and open information society. ACRL, the largest division of ALA, is dedicated to enhancing the ability of academic library and information professionals to serve the information needs of the higher education community and to improving learning, teaching, and research. Both ALA and ACRL publish scholarly, peer-reviewed journals in the field of library and information science.

ALA and ACRL appreciate the opportunity to comment on increasing public access to scholarly publications resulting from federally funded research. Many of our individual members and their libraries will also submit detailed comments to OSTP. ALA and ACRL have long believed that ensuring public access to the fruits of federally funded research is a logical, feasible, and widely beneficial goal. Both ALA and ACRL have endorsed "The Federal Research Public Access Act of 2009" (S. 1373) noting, "It reflects ALA policy regarding access to federal government information by providing for the long-term preservation of, and no-fee public access to, government-sponsored, taxpayer funded, published research findings."

ALA and ACRL offers the following comment to the specific 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 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 complete collection of articles resulting from publicly funded research should be made immediately freely accessible, so that the public can fully use them – (i.e. text mine, data mine, compute on them, create derivative works) without commercial restriction.
- For any public-access policy to be successful, there must be consistency of requirements and mandates. Institutions often have researchers who hold grants from multiple agencies concurrently; therefore, uniform requirements and procedures regarding deposit of peer-reviewed literature should be established across all funding agencies covered. Uniformity of deposit requirements will reduce the complexity and cost while at the same time increasing the rate of compliance. Based on the initial experience of low

manuscript deposit rates under a voluntary National Institutes of Health (NIH) Public Access Policy, mandatory policies are needed to ensure compliance and routine uptake of such submissions.

(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 given by statute, under the authority granted by the United States Constitution, Article 1, clause 8, to the authors of works specifically in order to provide an incentive for further creation (“to promote the progress of science and useful art”). Public access to scholarly articles not only does not impair this incentive; it actively supports it. The rewards that scholars obtain for the journal articles (promotion, tenure, grants, etc.) all come from having a significant impact on their disciplines. Public access has been repeatedly shown to increase citation rates and, in general, improves the chances that a scientific article will find more appreciative readers, especially those who might not have had access to a particular subscription journal. Thus public access itself will enhance the value of authors’ intellectual property.

(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 made permanently accessible, and useable. However, we feel there should not be only one single point of access. Many academic libraries are ready to assist the management of these publications by having them deposited in their institutional repositories that they manage on behalf of researchers.
- The authorized repositories should provide support for converting the file to a standard mark-up language, such as the currently preferred XML, if the file is not submitted in that format. PDF, a document format in ubiquitous use, does not support robust searching, linking, text-mining, or reformatting over the long-term, nor does it provide full accessibility for the blind and reading impaired. Standardization of format across the board is a key element to long-term public access. The options for submission format should follow the conventions of the disciplines from which the papers come, and not create an undue burden for the authors or publishers.

(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 provided repositories meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles. Under no condition should any single site be the single point of access for these articles.
- ACRL is well-suited to assist our libraries and campuses again by sharing best practices, promulgating model publication agreements, and providing education about the rapidly changing scholarly publishing environment in which these policies fit. ALA will continue its mission to assist our patrons in accessing such critical federally funded information that would be made available.
- The NIH public access policy for federally research has proven to be a good model. However, further mandates should reduce the allowable embargo period for public access to research articles and take steps to enhance the ability of users to reuse the material they find in the repositories

(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?

- Making these works openly available and free for reuse is the first fundamental step to encouraging these goals. Metadata associated with these articles should be viewed as a means for enabling specific actions, rather than simply an item description and should facilitate use, reuse, and analysis of published works. Further, it should be machine-readable and machine-interoperable. Metadata should support the proper context for published resources, i.e., controlled vocabulary, attribution for funding organizations, grant IDs, and provide descriptions of relationships between entities (such as existing means such as RDF and OWL enable).

(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?

- Effective implementation strategies for public-access policies can help maximize returns to taxpayers by ensuring that complete results are widely available in a timely manner. Policies should take advantage of existing protocols to facilitate automatic deposit of manuscripts to multiple repositories (i.e., SWORD).
- Policies should encourage development of additional tools/ services, such as integrating articles with grants management systems (internal and external) or improving agency accountability.
- The NIH has implemented an agency-funded model for public access that has proven to be very cost-effective. The benefits in terms of faster research, improved innovation, greater government accountability and opportunities for cost-effective commercialization easily outweigh the smalls costs associated

with managing a repository

(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 core of a public access policy should clearly be peer-reviewed journal articles, since this is a medium for which there are clear standards and expectations, whose use is familiar to most people, and for which the intellectual property rights situation is well-understood. Authors of these papers do not typically depend on royalty payments as an incentive, so participation by the funded researchers will be less problematic in this area than those in which commercial expectations are paramount.
- Other types of materials will have different conditions, and the intellectual property (IP) rights can get very complicated in some cases. Educational objects – digital objects created to teach a particular concept or point – would be another type of material to consider, however the standards of peer-review are not yet clearly in place for those works, and the IP rights can be very unclear. So while these educational objects, if funded by federal research money, might well be the subject of a separate mandate crafted for the specific conditions, it is clear that journal articles should be the starting point. The complexity of other materials should not be allowed to delay the implementation of a mandate for public access to funded articles across federal agencies.

(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?

- Immediate access is the ideal time to optimize scientific and commercial utility of information contained in these articles. We urge a short embargo period and recommend a 6-month maximum to bring U.S. policy into alignment with policies already in place in Canada, the United Kingdom, and the European Union. However, to accommodate those journal publishers who continue to rely on subscription income, an author-determined embargo period of 0-12 months has proven effective across multiple disciplines.

ALA and ACRL would like to bring the following studies to the attention of the Office for Science and Technology Policy as it considers the most effective implementation of public access policies:

“Of Mice and Academics: Examining the Effect of Openness on Innovation.”
Fiona Murray, Philippe Aghion, Mathias Dewatripont, Julian Kolev, Scott Stern
NBER Working Paper No. 14819, issued in March 2009
<http://www.nber.org/papers/w14819>

“Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research.”

Jeffrey L. Furman, Scott Stern NBER Working Paper No. 12523, issued in September 2006
<http://www.nber.org/papers/w12523.pdf>

“Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs — Report to SPARC.”

John Houghton, Bruce Rasmussen, and Peter Sheehan
<http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>

“Economic implications of alternative scholarly publishing models: Exploring the costs and benefits.”

JISC EI-ASPM Project, a report to the Joint Information Systems Committee

John Houghton, Bruce Rasmussen and Peter Sheehan, Charles Oppenheim, Anne Morris, Claire Creaser, Helen Greenwood, Mark Summers and Adrian Gourlay

http://www.cfses.com/documents/2009_CSES_JISC_Final_%20Report.pdf

Statement by David J. Lipman, M.D. Director, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health

U.S. Department of Health and Human Services:

“On Public Access to Federally-Funded Research before Committee on Oversight and Governmental Reform Subcommittee on Information Policy, Census and National Archives United States House of Representatives”

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

Response to Request for Information Public Access to Peer Reviewed Scholarly Publications Resulting from Federally Funded Research

Richard E. Luce, Vice Provost and Director of University Libraries
Emory University
Atlanta, Georgia
[REDACTED]

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 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 the U.S. economic growth and improve the productivity of the American scientific enterprise?]

Because of the benefits both to scientific productivity and the economy, federal agencies should adopt policies to make articles resulting from federally funded research publicly accessible. Scientific research creates marketplace demands for labor, products, services, and further research. Private investment and commercial applications that result from cutting-edge research will spur economic growth, potentially creating jobs across broad sectors of the economy. With public access to previous articles, scientists can avoid repeating research, and can discover new avenues of inquiry. Finally, public access means that taxpayers have access to the results of research their dollars funded.

Science is, by its very nature, cumulative—all new knowledge builds on earlier findings. The ability to access and use previous research accelerates scientific discovery and innovation. Open access to research articles increases citations and follow-on research, diversifies pathways in research, and accelerates applications in industry and commercial settings. It is critical that works be free for access and re-use (with minimal embargo restrictions), to enable a full range of activities including text and data mining, computing, and the creation of derivative works. Enabling not only open access but also full reuse means additional value—experiments are not repeated, and research not re-done. This, essentially, enables us to “do more with less.”

Making research articles open access also enables the inclusion of machines as a new category of reader, opening up research pathways and making new connections possible. Scientists can not only work faster, they can work smarter, deploying new search, discovery and semantic tools. For example, the National Center for Biotechnology Information (NCBI)'s PubMed Central (PMC) provides research articles in a machine-readable format. This means not only that these articles are openly accessible, but also that articles in PMC can be a gateway to some 40 databases and 13 trillion bytes of NCBI data.

Such access can advance scientific discovery immeasurably. (<http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>).

Because of existing infrastructure, the relative costs of public access policies are manageable. The example of PubMed Central is instructive. After start-up costs of \$500,000, annual operating costs (including ingest, refinement of submission system and search tools, and national office and help desk staffing) run to \$3.5 - \$4.0 million per year. This is only a fraction of NIH's \$30 billion/year overall budget—about 1/100th of 1 percent. The benefits of an open access policy like the NIH policy are estimated to encompass approximately 8 times 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—with 60% of that accruing directly to the U.S. economy. Open access policies for other agencies can be implemented in a cost-effective manner by leveraging the PMC infrastructure that exists for making articles open access. Supplementing existing access points with additional manuscripts would represent only small incremental costs.

Ideally, the type of access enabled by such policies should be immediate and full—with embargoes either eliminated or minimized. Since investment in science should create jobs now—not twelve months from now—restrictions should be minimal.

Scientific progress leads to economic development, and limiting this progress by granting access only to a privileged few hinders it. Scientific research drives new discoveries and technologies, which affect the economy by creating new business needs, services, and jobs. When considering the costs of doing duplicate research and pursuing blind alleys, the potential cost saving in dollars and researcher hours would likely run high. By saving time and money, more research could be done for the same expenditure.

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?]

Adopting public access policies does not mean turning copyright law on its head. Agencies can work within the current copyright framework to provide public access to scientific research. Both scientists and the public, however, deserve access to and full use of these articles sooner than the current term of copyright allows. Access can be accomplished, while still protecting the intellectual property interests of all stakeholders, by implementing appropriate licenses (e.g. Creative Commons CC-BY licenses). Such licenses enforce, rather than subvert, current copyright law.

Policies resulting in “read-only” databases should be avoided. Mechanisms to enable full use of research—including distribution, reuse, text mining, computation, etc.—should be included in any government public access policy. In attempting to balance stakeholder

interests, a useful strategy to consider would be a tiered approach. First, articles could be provided for a period where current rights apply; after a period of time, articles could be accessible with full reuse rights under an appropriate open license (CC-BY or similar).

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 in a unique position to ensure that publicly funded articles are preserved and made accessible. Any public access policies that are developed must give the federal government adequate rights to archive and distribute publicly funded articles. PubMed Central illustrates the benefits of both a centralized repository and a government mandate to open access. In its first few years of existence, contribution was voluntary. However, only 5% of eligible articles were submitted by their authors. While journals also submitted a portion of articles, under this voluntary system, only 19% of articles eligible under the policy were included in PubMed Central. Now that submission is mandatory, NIH is able to increase content while refining their information systems for ease of use, navigation, and information finding.

A federal public access policy could involve multiple repositories, but these repositories must be able to ensure long-term preservation, access and usability. Interoperability and long-term preservation could be maintained by third parties, encouraging partnerships between the public and private sectors. Simply providing the government with a copy to put in a “dark archive” is not a viable solution; without regular use, archival veracity cannot be ensured. Library experiences have shown that regular access and use of digital materials is a crucial element in effective long-term preservation.

The drawback of a decentralized approach is the potential for failure. Current attempts at archiving scientific literature are inadequate. In a widely cited study, Cornell and Columbia report that only about 15% of their combined journal holdings are currently archived by the LOCKSS (Lots of Copies Keep Stuff Safe) initiative and Portico archive combined (<http://2cul.org/node/22>). Despite this dismaying figure, the Cornell/Columbia efforts are considered to be a relatively successful digital preservation initiative. LOCKSS and Portico have some degree of overlap, but both services preserve titles uniquely. There is much room for improvement, which could arise from federal stewardship of research products.

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 could be a vital component of a federal public access policy, provided repositories meet conditions for public accessibility, use rights, interoperability, and long-term preservation. Since no single site should be the sole point of access for research articles, partnerships between federal agencies and private entities could be the key to ensuring a viable system. Universities and libraries have extensive experience and existing archive infrastructure, and should be actively encouraged to partner with federal agencies.

One model for public-private partnership is PubMed Central. When PubMed Central was launched in 2000, a number of publishers collaborated, depositing their journals in the database. Using these submissions, NLM developed the structured digital format for representing digital journal articles—the NLM DTD. In turn, this format is being adopted by publishers and libraries (and is in the process of becoming a NISO standard).

There are also good examples of partnerships among academic research institutions, and companies, providing access and preservation. For example, the DRIVER initiative is a European network of digital repositories housing 3,500,000 scientific publications—a consortium of sorts of 295 open access repositories, from 38 countries (<http://www.driver-repository.eu/>). HathiTrust is a partnership of leading research institutions and libraries, with sixty partners worldwide (<http://www.hathitrust.org/>). Both initiatives seek to preserve and provide access to content, with both a central location and integration into partner libraries' local systems.

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 for 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 should facilitate use, reuse and analysis of published works, and must be machine-readable and machine interoperable. Metadata in its current form is a means for enabling actions, rather than simply embodying description of items. Existing standards (Dublin Core, OAI-PMH, Datacite Metadata Schema, Europeana Semantic Elements, ORCID, I-2, Counter/SUSHI) may help inform metadata requirements, and agencies currently working to improve metadata interoperability (e.g. NISO, LoC) will be essential.

Furthermore, one critical element of metadata is that it must be coupled with an API for standards-based data exchange. It is not enough to create metadata—descriptors must be easily shareable among institutions and repositories. The importance of shareable metadata indicates that libraries must be vital partners in this process.

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 of stakeholders, including awardee institutions, scientists, publishers, Federal agencies, and libraries?]

In order to maximize the benefit of public access policies to taxpayers, while minimizing burdens and costs to stakeholders, uniform requirements and procedures surrounding the deposit of articles should be established across all funding agencies. Institutions and researchers often hold grants from multiple agencies concurrently, and common requirements will reduce complexity and increase compliance. Policies should take advantage of existing protocols (for example, the SWORD standard, <http://swordapp.org/>) to facilitate automatic deposit of manuscripts to multiple repositories. Agencies can integrate articles with grants management systems, improving agency accountability and compliance while providing increased information to the public on research results.

The PubMed Central database currently contains over 2 million articles, and is used by more than 500,000 users per day, underscoring a deep demand for this information in the public sector. NLM statistics indicate that while 25% of users of PMC are from universities, 40% are private citizens or those using personal internet accounts, and 17% are from companies (the remainder being government or other users). PubMed Central is a broad-based repository used not only by researchers and students, but also clinicians, patients and their families, and entrepreneurs. Other federal agencies can take these statistics to indicate that the public values access to scientific research.

A major benefit of public access policies is that taxpayers can access the results of research their tax dollars fund—immediately and fully. Simply put, opening access to publicly funded research articles increases the public's return on their investment in that research. Since, by paying taxes, Americans are the investors in federally funded scientific research, public access policies maximize the value of investment in research.

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?]

Along with journal articles, a host of other materials are good candidates for public access. For example, educational materials (such as book chapters, texts, and conference proceedings) could be made readily accessible to the public. Making these materials open access will improve the quality of education, allowing cutting-edge research to be quickly incorporated into the teaching process. Additionally, much as in the case of journal articles, more access will mean better science. However, the policies under which these materials are made accessible will likely differ from those directed at journal articles. While authors are not paid for journal articles, they may make significant royalties from textbook

chapters, for example. Since scientists and scholars may rely on this income, policies will need to reflect these differences in the current scholarly publishing apparatus.

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?]

To optimize both scientific progress and economic growth, immediate access is best. However, journals do rely on subscription income. An embargo period of up to 12 months has proven to be an effective accommodation of journals' economic models, while still resulting in improved access to scientific research. It is important to note that no publisher has provided empirical data indicating that this embargo system (currently used by NIH and numerous funders worldwide) harms journal revenues or drastically alters their economic models. Embargoes of 12 months or less have been adopted by hundreds of journals (<http://highwire.stanford.edu/lists/freeart.dtl>) and research funder policies around the globe (<http://roarmap.eprints.org/>).

Public access to scientific research does not need to interfere with journal revenues. Even publishers who initially balked at opening access to back content have changed practices. In the name of scholarly exchange of knowledge, the Royal Society recently opened its entire historic journal archive—including the first ever peer-reviewed journal—dating back to 1665, adopting a 12-month embargo period. (<http://royalsociety.org/news/Royal-Society-journal-archive-made-permanently-free-to-access/>). They noted that this back file accounted for less than one half of one percent of their overall publishing revenue.

The claim that public access policies negatively affect publishers (presumably because of cancellations) often does not take into account other factors, including growth of directly competing journals, price and pricing history, revenue resulting from "long-tail" citations, impact of publisher-required bundles versus single journals, and library budget numbers. These and other market conditions affect journal cancellations, and must be accounted for in any analysis of lost revenue.

In fact, studies have shown that open access articles actually have the advantage in both visibility and citations. Articles whose authors supplemented subscription-based access by self-archiving their final drafts gained citations, in some cases doubling citation rates for articles that are not open access. Thus, if taking into account both public benefits and scientists' interests in their own research, open access articles have an advantage.

Subject: Response from Provost of The University of Kansas

Date: December 21, 2011 1:20:51 PM EST

Background: The University of Kansas has demonstrated over the past 10 years a strong commitment to the principle and practice of making scholarship publicly accessible. In 2009, for example, the faculty adopted an open access policy, and revised and strengthened it in 2010. In the summer of 2011, the University of Kansas became a founding member of the Coalition of Open Access Policy Institutions (COAPI) and most recently in the fall of this year the Chancellor of the University signed the Berlin Declaration of Open Access. The faculty, university administrators and librarians are committed to policy enactments that provide public access to the research that the public funds. We take active steps in sharing the work that we, at KU author, and strongly support a broadening of the federal government's funding agencies' mandate on public access to publicly funded research.

1. Question 1 in four parts:

- 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?

First and foremost, in order to ensure the widest exposure possible, peer reviewed scholarly articles resulting from research funded with public funds must be made permanently publicly available. When publicly available, such access encourages investigation, interdisciplinary and innovative thinking across commercial sectors which can lead to new products, as well as the creation and support of new, growing and even struggling markets. With unfettered and immediate access to such works, the opportunity for rapid commercialization quickens the research life cycle allowing the untapped potential of such research in various sectors (agriculture, technology, energy, publishing, medicine, public policy, education, digital humanities, to name a few) to be fully realized.

Additional uses of such freely accessible works include the public's ability (in various industrial, research and education environments) to text mine, data mine, compute on the data from such work, and create derivative works, all without the commercial restrictions that closed access and rights-limited research are subject to. Innovative small businesses, individuals, research and educational centers can use this information to create new services and products and encourage new and as yet not fully realized economies. Already there are companies that use public research data (like data provided by the National Weather Service) and build products and services like the, Google Scholar, goPubMed, local and regional news and weather providers.

Agencies can help ensure that publicly available research will be fully utilized by promoting standard-based resource description and data exchange. This includes the use of existing and emerging standards such as, for example, Dublin Core, the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), Resource Description Framework (RDF), and the Simple Web-service Offering Repository Deposit (SWORD). It will be crucial for this metadata to extend beyond simple resource description to rights and reuse statements, in order to facilitate the proper use, analysis, and reuse of publicly available research. Published APIs for the exchange of data in standard formats such as JSON and XML can enable machine-interoperable exchange of resource metadata and full-text, greatly expanding the potential use and re-use of publicly available research.

- 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?

Public access to research offers greater opportunities for serendipitous contributions in the scholarly and research endeavor and increases the opportunities for the cross-fertilization of ideas, innovative outcomes, as well as new products created by a wider and perhaps unexpected group of citizen scientists. Unexpected/unforeseen participants, including international ones, who can access work otherwise outside of their reach, can offer contributions downstream that can lead to advancements in new research, discovery, and commercialization of knowledge. This can have tangible economic and marketable effects.

An expanding range of applications are expected when more segments of society can access the results of research. Publicly accessible research is cited more often, providing the opportunity for greater reach and follow-on research. (See Fiona Murray paper - <http://www.nber.org/papers/w14819> and Furman & Stern paper <http://www.nber.org/papers/w12523.pdf>. Note the papers cited here are publicly accessible research papers and are thereby available to policy analysts and the general public without having to pay (again) for them.)

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

There are many benefits associated with such policies and all benefits bring some costs. Many such benefits have been reported in the Houghton report, (See Houghton, J., & Sheehan, P. (2006). *The Economic Impact of Enhanced Access to Research Findings*. CSES Working Paper No. 23. Centre for Strategic Economic Studies. Victoria University. Melbourne, AU. Retrieved from <http://www.cfses.com/documents/wp23.pdf>. Note that this document is also a publicly accessible research paper that citizens' tax dollars, somewhere, funded.) It has been demonstrated that there is an increase in the return on investment, maximizing productivity; that the benefits of public access are eight times larger than the costs; and that a public access policy is cost-effective—NIH's costs are 1/100th of 1 percent of its overall budget, with visitors from the public, private, commercial sectors (patients, doctors, researchers, inventors...downloading publicly accessible research.).

Public access increases citations and supports new and diverse research questions and investigation. (See Murray and Furman & Stern papers referenced above) and provides a means for technology to be used to identify, analyze, and mine information, metadata, and data from publicly accessible research reports.

Costs can continue to be kept to a minimum by utilizing existing infrastructure and investments made already in the NIH public policy, for example. Open sharing of federally funded research endeavors can also demonstrate the cost efficiency and accountability for the public funds spent to sponsor this research. It also provides transparency for the return on investment.

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?

The type of access needed is immediate, public access with reuse rights in order that future scholars and citizens may fully engage and build upon the work—leading to new uses and outputs for future works.

Access to scholarship that does not allow for robust reuse rights limits the kind of engagement that researchers, scholars, inventors, creators can have with the published works, as individual works or more broadly as bodies of work. Broad reuse rights permit researchers far into the future to revisit older works and engage them fully. Remember that copyrights and restrictive licenses on scholarly work generally last (and are transferred to others for) 70 years past the life of the author. Without liberal permissions to use the work now, fully, and into the future, scholars, the public sector, and inventors (etc.) are handicapped to extract the full potential of the work.

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?

Read-only access to the results of scholarship and/or the online publicly accessible databases in which they would reside is not sufficient to allow the full suite of public and private interests to exploit the work, the metadata and the data.

Rights belong originally to the authors (who are funded by the public and thereby are stewards of the public access). Publishers need only certain copyrights to distribute/publish the work. We recommend that non-exclusive rights be granted to publishers, with copyrights remaining with the author. A limited embargo on the public access to the papers published in journals would allow publishers to still attract subscribers (in a closed-access environment) for periods of time. The author, retaining rights, could then provide Creative Commons Licenses (or similar licenses) post-embargo to all readers, such as the CC-BY License (Creative Commons, Attribution license) that allows the work to be reused with attribution to the original author.

In the NIH model, the government makes no claim on the work itself, but must be granted the permissions (license) to make the work available (disseminate) to the public. The terms of the sharing of intellectual property between the author and publisher happen prior to the work being deposited in a central repository. It is highly recommended that an author retains his/her copyrights but grant the needed licenses to the publisher and the federal agency as well as future readers to disseminate and use/reuse the work as needed. It would be important to consider offering a single template for rights transfer, to ease the processes involved for authors, author's institutions, publishers, federal agencies and future readers/users of the work.

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 published results of scholarship/research are funded by the government on behalf of the public, which means that the results are a public good. It stands to reason that the government should maintain shared custody of that work on behalf of the public it serves. As recipients of public funds, universities are a primary locus of production of this public good. It follows that the universities as proxies for the government could be good candidates to assist in the stewardship of the work on behalf of the public.

Many universities around the U.S. have established digital and publicly accessible archives for just such scholarship. A network (and digitally interconnected) of such repositories could serve as excellent secondary holding spaces for such work. Such a network can be enabled through the adoption of standard-based approaches to creation and exchange of resource metadata, including rights and re-use statements. In the case of such networks, emerging standards such as Open Researcher and Contributor ID (ORCID) will help enable the seamless exchange of distributed metadata.

It is not appropriate for commercial entities to be the sole custodians of the public's research results. Commercial publishers merge, change policies, shift priorities, go out of business, and do not have at the center of their mission nor the incentive to provide the long-term essential public access needed. Partnerships with universities or publishers to house, store and provide access to scholarly papers are possible. However, at minimum the government should be required to maintain mirrored, public online versions of the content, perhaps via a

portal or e-hub mechanism. Federal stewardship (as mentioned earlier) is cost effective—see the case of NIH’s operating budget to run the PMC public access program—costs of less than 1/100th of one percent.

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 use of proprietary (and/or privately controlled) archives is a risky business and to be strongly discouraged. Currently none of the over 50 research funders with public access policies use propriety sites as the final archive. A good example would be funders partnering with academic and research institutions that have similar interests in the wide and public access to the work and whose primary stakeholder/constituency is NOT stockholders.

A carefully coordinated effort between universities with open archives for scholarship could be utilized to advise or assist in the federal efforts to house and guarantee access to the scholarship which its public funds.

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?

A fundamental reason to use an open access model, aside from the principle of “public funding = public good,” is that research is increasingly begun in open internet search engines (see Schonfeld, Roger C., and Ross Housewright. 2010. Faculty Survey 2009: Key Strategic Insights for Libraries, Publishers, and Societies. Ithaka S + R.) Closed-access represents at best “friction” or, at worst, an impediment to both discovery and access. This means that research practices are driven down the path of least resistance. It follows that openly accessible literature will be seen and thus considered by the largest number or researchers, while non-open access literature will remain “hidden” from researchers and evade the reach of at least a percentage of inquirers. Outsell, in its 2009 report (Open Access Primer (Public Version): a report by Outsell Vol. 3, December 14, 2009. Outsell, Inc. p.33) finds that approximately 9.4% of scholarly journal articles are currently publicly accessible (without subscription). Although 100% of such research is funded ultimately by the public, the public can only access about 10% of the work without a subscription, or affiliation with a subscribing institution. As open access becomes the norm, closed-access-only will increasingly become, in practice, “no-access.” This makes public access an imperative rather than merely a best-practice.

Published texts can be used as data objects making the metadata an important feature in discoverability. Metadata should not be viewed as solely descriptive of the work itself but as a means for enabling discoverability, use, reuse, and analysis and be machine readable and interoperable, utilizing appropriate standards.

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?

Consistent public access policies across federal agencies are critical to minimize the burden on authors, database submitters, publishers and the public and reduce the complexity of submission, discoverability and cost.

Such changes in policy would begin to mitigate a long-term burden on the public: lack of access to the research it has and will continue to fund. Such a policy would correct a moral wrong: without public access to this research, the public (including small business, K-12 and higher education) has limited access and bears the cost burdens without benefitting from access to the products it funds.

Currently the public pays four times before it is able to have access to the results of scholarship: it pays the salaries of the scholars/academics who do the research, and review and edit for journals; it pays taxes which seed billions of dollars of federal funding agencies; it pays when the institutions in which the authors work subscribe to those journals (at the University of Kansas, close to \$4 million dollars per year for journal subscriptions) and, if not a member of such a subscribing institution, a citizen must pay a retail charge for an article on the open market (prices that are currently far from nominal—on the order of tens of dollars for a single article). These citizens are small business owners, investors, policy makers, scholars, researchers, government officials from the smallest towns to large cities. They are high school and junior college students and teachers—they are patients, entrepreneurs, innovators, parents, PTA members, investors, researchers—the unforeseen and unexpected participants in the advancement of science, public policy, education and scholarly pursuits. The current structure of scholarly publishing is perverse and abusive: the public funds are diverted from the public interest and have been captured by private, commercial interests.

Mitigating burdens on publishers, researchers and universities would include adopting compliance methods that are as consistent and stream-lined as possible for all. NIH's model would be one to contemplate and learn from.

Public access required by such policies would also assist the scholarly publishing industry to become more diverse in its own approaches to access and business models. A growing number of publishers are experimenting with moving to all-open access models for their content. We know of no publishers that have supplied data which indicate that NIH's current policy has been a financial burden on their business models. Whether those are commercial, close-access journals/publishers or open access non-profit publishers on the other end of the spectrum, publishers have learned to work within the constraints of the authors' requirements to share their work.

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 possibility of including other materials in such a policy certainly does exist, although at this time we would recommend keeping the focus strongly on scholarly journals. Journal articles provide more current information and they are the preference of researchers to share their work quickly among peers. The length of time for embargos and the rights issues for other types of work could be significantly different and make what are clear cut ethical/philosophical issues and more complex issues around such a public policy even more complex.

Keep the focus on scholarly journal articles at this time. Scholarly, peer reviewed journal articles are the primary method that scholars communicate. There is a clear and unbroken association between the research conducted and the primary method that scholars communicate the results of their scholarship—no matter what combination of public agencies fund the work—and that is through the aegis of the peer reviewed publications. Research “reports” are not a substitute nor are they the currency used by scholars, citizen scientists,

public policy makers (the list goes on). There is one common currency and that is the scholarly peer reviewed journal article. It is our understanding that the first COMPETES act required that NSF make public research reports and it was uniformly ignored—as an irrelevant method to communicate the results of research. Besides being un-reviewed by referees in the field, the work is not commonly used, circulated or searched for by others scholars.

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?

Immediate public access is ideal in order to maximize the scientific, commercial and public/private use of the information.

That said, in order to minimize the burden on publishers that still rely on closed-access subscription income, a modest embargo of 6 months could be considered. Embargos of 12 months or less are the norm for researcher funders with public access mandates around the world. See ROAR MAP's site which indexes such mandates, <http://roarmap.eprints.org/>. Hundreds of journals already have accepted 12 month or less embargos on their content. See Stanford's Highwire publisher's site, <http://highwire.stanford.edu/lists/freart.dtl>

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.

Another year passes and additional comments are requested, when virtually all stakeholders in the system of scholarly journal publishing except the publishers consider public access to the publically funded research results self-evident. Although a very powerful lobbying machine, can commercial interests be allowed to continue to commercialize and commodify what is, in all minds, a public good—and which the public funds? Borne of public funds, public needs, public curiosity and efforts, are these results not a public good? A public access policy like NIHs would allow publishers to make a return on their efforts while the public and the nations' research enterprise can see a return on its investment.

It is time for a broader public access policy to be completed and implemented. Commercial interests harm all other stakeholders in the system of scholarly communication by preventing them from accessing, building upon and utilizing the research it funds and completes.

In short, such a policy must include federal stewardship of publicly accessible scholarship funded with public monies; the works made available must allow for robust reuse; must be available immediately or at maximum under brief, 6-12 month embargoes; must build on the successes and lessons learned during the NIH public access policy implementation; must move toward efficient and simple submissions, cross-agency consistency in policy and implementation and again, guarantee the long-term public access and preservation of the work.

Response to this RFI is voluntary. Responders are free to address any or all the above items, as well as provide additional information that they think is relevant to developing policies consistent with increased public access to peer-reviewed scholarly publications resulting from federally funded research. Please note that the U.S. Government will not pay for response preparation or for the use of any information contained in the response.

How To Submit a Response: All comments must be submitted electronically to: <mailto:publicaccess@ostp.gov>

Responses to this RFI will be accepted through January 2, 2012. You will receive an electronic confirmation acknowledging receipt of your response, but will not receive individualized feedback on any suggestions. No basis for claims against the U.S. Government shall arise as a result of a response to this request for information or from the Government's use of such information.

Jeffrey S. Vitter Provost and Executive Vice Chancellor Roy A. Roberts Distinguished Professor The University of Kansas | Web www.provost.ku.edu

Subject: Request for Information Response

Date: December 21, 2011 2:42:26 PM EST

On behalf of the American Society for Laser Medicine and Surgery, and the Editorial Board of Lasers in Surgery and Medicine (ISSN0196-8092) please find our response to the RFI. We do not have an ID number

December 21, 2011

Beth Mallen Watson
Managing Editor Lasers in Surgery and Medicine
American Society for Laser Medicine and Surgery

Comment 1

The NIH public access policy is well intentioned but seriously flawed. Surely, the public has a right to access meaningful information on health-related research funded by federal dollars.

But, we would argue that simply making our publications available to the public fails to accomplish this while at the same time imposing a significant administrative burden on authors, journals, and publishers. The key point that seems to have been lost to the policy makers is this - the vast majority of the public cannot read or interpret scientific/medical research articles, which are written for a highly technical audience of specialists.

Comment 2

If we really wanted to make our federally funded research findings available to the public, it would be necessary for some entity to prepare frequent review articles on a rather vast range of subjects written for a lay audience. A more productive approach may be to require federal grant Principle Investigator (PIs) to write short summaries of their research, written for non-specialists, at the end of each 3-5 year funding cycle. These summaries could be posted somewhere and made freely available to the public.

Comment 3

There is indeed an implication by government that "public" access will result in more rapid access to research content. Yes, some groups do provide immediate and free access to their content, particularly in electronic forms. This is a business decision on their part and there is a cost in doing so, which must be borne by members of the group, patrons or other sponsors. There is a cost involved in producing the material including review, editing, formatting and cataloging the material, among other details that cannot and do not happen without the involvement of several individuals. While reviewers are uncompensated, it is unreasonable to assume that the entire scientific publishing enterprise should be uncompensated.

The mandates do infringe upon copyright laws and processes and it is unreasonable to single out science and scientific literature.

Comment 4

There is a sense that the implementation of the regulations as currently promulgated would actually serve to undermine and promote the collapse of the publishing enterprise, as we understand it today. This would result in less material being brought to scientific peer review and publication and decrease availability.

Arguably, the public does have access to the content free or pay-per-view. Most if not all publishers will sell reprints, access or license content they own, and content is indexed and available through several sources including PubMed and others.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Boston College University Libraries

Response to:

OFFICE OF SCIENCE AND TECHNOLOGY POLICY

Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research, November 3, 2011

From:

Thomas B. Wall, University Librarian

The Scholarly Communication Committee

Boston College University Libraries

Chestnut Hill, MA

(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?

Key points

- Full and immediate open access to products of taxpayer-funded research has the greatest potential for:
 - Rapid spread of ideas
 - Increased commercial benefits
 - Acceleration of scientific discovery
 - Increased interdisciplinary discovery
 - Enabling the unintended or disadvantaged reader
 - Enabling computer analysis
- Existing business models can adapt, with help from grant-funders during the transition.
- Additional costs would be small, for great benefit.

Full immediate open access to the products of taxpayer funded research holds the most promise for rapid spread of ideas. This strategy has the potential to create the greatest return on investment for taxpayer money and to increase commercial benefits. Access should be online, without economic barriers, free of technical barriers and coupled with the right to reuse and build upon the knowledge accessed. It is important that users have the ability to reuse, redistribute, make derivative works, datamine and reconfigure content as necessary to maximize potential benefit and leverage data to innovate.

The economic well-being of the nation benefits from open access as the government's return on investment in research is augmented. Access to previously hard to obtain research encourages innovation, stimulates new ideas, services, products, and creates new opportunities for job creation. Research is exploited

through new and diverse pathways that lead to greater modernization and commercialization of many aspects of the US economy. Open access to the results of federally funded research is beneficial to other nations too, particularly less developed ones. Institutions and people in developing countries find it exceedingly difficult to afford the scholarly journals published in the developed world. Having open access to such publications provides greater opportunities to researchers in developing nations to engage in better scientific, medical, technological, business and other research. The American public has the opportunity to become better educated so that it can more meaningfully understand and support scientific, industrial and related enterprises. Open access to the results of federally funded research provides them with previously unavailable materials relating to health, energy, environment, and other areas important to day to day living. Similar concerns and goals were expressed recently in the UK government report, *Innovation and Research Strategy for Growth*, which stated:

The Government, in line with our overarching commitment to transparency and open data, is committed to ensuring that publicly-funded research should be accessible free of charge. Free and open access to taxpayer-funded research offers significant social and economic benefits by spreading knowledge, raising the prestige of UK research and encouraging technology transfer. At the moment, such research is often difficult to find and expensive to access. This can defeat the original purpose of taxpayer-funded academic research and limits understanding and innovation... .[W]e need to go much further if, as a nation, we are to gain the full potential benefits of publicly-funded research.

Innovation and Research Strategy for Growth, Presented to Parliament by the Secretary of State for Business, Innovation and Skills, December 2011. <http://www.bis.gov.uk/assets/biscore/innovation/docs/i/11-1387-innovation-and-research-strategy-for-growth.pdf>

Currently, university graduates who have become accustomed to full access to the peer-reviewed literature in their field find themselves, as practitioners in the field, with no access to the literature. Their access to new developments reported in the broad range of scholarly journals will be frozen at graduation. Requiring that taxpayer-funded literature be made openly accessible to all will allow current practitioners in all fields to have equal access to the literature. Making the literature open to all in a free database will give the high school student, the inventor in the garage, the small businessman and the clinician in a non-profit clinic the same access to new knowledge as the scientist with a large R&D budget at his disposal. Discovery should be made easy by allowing indexing and searching by common search engines as well as by the database's own interface.

When publications are produced in the traditional manner, much time can elapse between the end of the research phase and the final publication. Many disciplines have found the length of this cycle prohibitive to their advancement. In the field of physics for example, the arXiv preprint server fills a perceived need to disseminate research more quickly. This wider, faster distribution can accelerate the process of scientific discovery and innovation.

The increasing interdisciplinarity of research and approaches to complex scientific problems would also argue for broader access. Even an enterprise with a large R&D budget will likely put its money only into the publications most central to its field. With the costs of publication rising at a rate many times that of the CPI, subscription cuts have been necessary for many. (See *Monograph and Serial Expenditures in ARL Libraries, 1986-2004*. <http://www.arl.org/bm~doc/monser06.pdf>)

It is not hard to imagine that answers to climate change issues might be found in research in biology, physics, material science or mathematics, yet many scientists now have access to a narrow range of publications central only to their field. Open access to taxpayer funded research would eliminate economic barriers to knowledge and maximize exposure of the research to foster scientific productivity in any quarter. Just as wider access to the literature associated with the Human Genome Project brought unprecedented advances in scientific knowledge, unfettered access could again result in major strides toward addressing this looming global threat.

Open access also provides an opportunity for a new kind of reader with new capabilities. If the contents of the articles (and the underlying data) are openly accessible in compliance with metadata standards, they can be “read”, digested, analyzed, and compared by computer programs that recognize new patterns and derive new knowledge based on massive amounts of data. If the articles are closed and inaccessible, this is not possible. This process can help scientists identify relevant data more efficiently and accurately. It opens avenues for IT innovators to create tools and processes to help with data analysis, strengthening and providing new areas of growth for the IT industry.

By far the greatest cost of funding research and making it openly accessible is the cost of the research itself. Taking the NIH experience as a model, the cost of making the research available in an open access database is only .01 of the overall NIH budget. That incremental increase in cost makes 2.2 million articles available for download by 500,000 persons per day. (See testimony of David J. Lipman, M.D., Director, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services (HHS) on Public Access to Federally-Funded Research before Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives, United States House of Representatives.

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

Many of these articles are published traditionally as well. The 12 month embargo period that NIH allows provides an economic balance – allowing the publishers to make their initial profits (recouping their costs during the time) but making the information accessible fairly soon. This embargo period might be shortened or eliminated by providing other financial incentives to those publishers who publish articles produced with federal funding. The article publishing fees that many publishers charge could be funded through the author’s grant, providing a substitute revenue stream for publishers.

An important analysis of the potential benefits to result from making federally funded research open access is the July 2010 report to SPARC [The Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs](#) (John Houghton, Bruce Rasmussen and Peter Sheehan). The primary focus of this study was the FRPAA bill. The report states:

Over a transitional period of 30 years from implementation, the potential *incremental* benefits of an open access archiving mandate for all FRPAA agencies’ funded R&D[2] might be worth around \$1.6 billion (Net Present Value), around 4 times the estimated cost using the higher end lifecycle costing, 8 times the cost using NIH costing and more than 24 times the cost using arXiv costing. Perhaps some \$1 billion of these benefits would accrue within the US, with the remainder spilling over to other countries. Hence, the US national benefits might be around 5 times the costs.[3] The overall impacts of openly archiving all FRPAA agencies’ funded R&D article outputs would be greater than these incremental impacts, with likely US national benefits of around 8 times the costs (Table 4). 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.

This data indicates that immediate open access would produce the greatest returns.

(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?

Key points

- Publication services and compensation can evolve.
- Lack of taxpayer access is the primary imbalance to address.
- Open access will be beneficial to authors' careers.
- Open access and standard licensing would eliminate chilling risk calculations by users of content.

Publishers who are not creators have no inherent intellectual property rights. They obtain them from author/creators as compensation for their publishing services. Exploration of new business models can find new ways to compensate publishers for a new array of useful services. It may be that we no longer need print content with high cost and limited distribution. We may still want the editorial, design, and distribution services and networks that publishers provide. Currently, embargo periods serve to allow publishers to recoup costs, but we may be able to pay those costs directly as a very small part of the research funding. Beyond the transitional period, it is possible that traditional publishers will change into providers of a different type of service, one that provides content freely but, as an example, also provides analysis and data mining tools and services for a fee.

As needs change, publishers will evolve to continue to meet those needs, but this should not be done at the expense of public access to publicly funded knowledge, or at the risk of slowing the progress of science to a pace that meets nineteenth century distribution methods. Publishers are already proving that they can innovate in this area, by providing deposit services to authors after embargo periods and by creating hybrid journal models where traditional closed articles coexist with openly readable articles.

Currently, in cases where open access to research is not required, the public is accorded no rights in the research it funds. This is the primary imbalance in the system that should be addressed.

Scientists typically receive little or no monetary compensation for the intellectual property created by writing and peer-reviewing articles. Their compensation comes in the form of enhanced career prospects, recognition and advancement within their fields. Making their articles openly accessible would not adversely affect them and in fact may advance their careers even further as new applications are built on their discoveries. Open access allows greater readership and the citation counts on which their careers are currently are often evaluated. (See Swan, A. (2010) The Open Access citation advantage: Studies and results to date. Technical Report, School of Electronics & Computer Science, University of Southampton.

<http://eprints.ecs.soton.ac.uk/18516/>)

By making all articles openly accessible and by allowing adoption of standard licenses defining reuse we can eliminate ambiguity that currently plagues the intellectual property environment. Merely extending the existing NIH mandate to other agencies would not accomplish this. The current NIH mandate is deficient in this respect. Rights of reuse should be defined and stated upfront. The current chilling effect of the need to evaluate risk would be eliminated and no one would need to fear infringement in their legitimate reuse of scholarly work.

(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?

Key points

- A centralized system could apply one set of standards for applications, metadata, access and interoperability.
- Federal operation would ensure longevity and sustainability and facilitate international cooperation.
- Government/academic collaboration would take advantage of collective expertise.
- Geographic distribution of mirrored content would be vital for digital preservation.

A centralized approach would employ one set of standards so that access methods and applications, methods of analysis and data mining would be made easier and more universal. We recommend a collaborative partnership between government and universities for the stewardship of these digital materials. This arrangement would benefit from the collective expertise of its stakeholder-participants, minimize any duplication of effort and maximize interoperability. Many universities have already joined efforts to promote discovery, access, and preservation for their extensive digital collections. And, many of the lessons they have learned in doing so would be relevant to ensuring long-term curation of published articles.

If the centralized system were federally operated, that would ensure longevity and sustainability, without the uncertainty of continued operation of private entities. A national archive would also facilitate international cooperative efforts, which will be extremely important for global scientific advancement. Openly accessible content would, by definition, be better preserved because it could be copied and distributed in geographically diverse preservation archives without running afoul of copyright restrictions. Greater regular use would ensure that any degradation in quality of the digital data would come to light more quickly. Geographic distribution of identical content would be vital for digital preservation. Several preservation initiatives serving libraries, such as PORTICO, LOCKSS and the MetaArchive Cooperative may provide some model components of a preservation program.

Some attributes that would promote trustworthy stewardship include:

- Conformance with the Open Archival Information System reference model
- Multiple copies –widely dispersed geographically and with fixity checking on a regular schedule
- The use of non-commercial open source software
- The use of non-proprietary file formats and a plan for future migration
- Adherence to best practices and standards, especially for metadata creation
- Trustworthy Repositories Audit & Certification for all membership repositories (Currently draft international standard ISO/DIS 16363. For more details, see: http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf)
- Mirrored and centralized online searching

(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?

Three existing models come to mind:

- The current relationship between taxpayer-funded research grants awarded by NIH, often to researchers at private institutions, resulting in publication by private publishers and then deposited and made openly accessible (although not necessarily free of copyright constraints) in the public database, PubMed Central.
- The collaborative efforts of the HathiTrust members, consisting of some private and some public institutions, engaged in a cooperative enterprise to build a database of published literature, making it accessible to all. In addition to the nearly 3 million public domain volumes available to all, this project has resulted in nearly 10,000,000 volumes being available and accessible to visually impaired readers.
- MetaArchive Cooperative, a group consisting of public and private university partners, engaged in a cooperative enterprise to provide digital preservation for the member institutions' digital scholarly publications. The diversity of membership allows for shared, duplicated, geographically distributed preservation of content.

(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 avoid excessive costs, clear guidelines should be developed to define a minimum core and as much metadata as possible should be populated from dropdown lists and other assistive methods. A schema which is granular enough to support services such as open url generation, creation of derivative reports and products, and inclusion in search engines is important. MODS comes to mind as a schema that would work.

Note that the Google Scholar inclusion guidelines for webmasters indicate that the appropriate metadata schema for journal articles must have unambiguous fields for journal title, volume, issue, and page numbers. Some popular schema, such as Dublin Core, are deficient in this regard.

<http://scholar.google.com/intl/en/scholar/inclusion.html>

(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?

Key points

- Representatives of all stakeholder groups should be involved in setting policy.
- The policy should be simple and consistent across all Federal agencies.
- All agencies should allow use of grant funds to pay article processing fees.
- Publishers should handle deposits to public databases.
- Databases should allow easy discovery and interoperability.

In order to make sure that concerns of all stakeholders are recognized and addressed, representatives of all groups should be involved in making policy.

It will be important to craft a policy that is simple and consistent across all funding agencies. Currently the differing requirements can be frustrating for grant applicants. Grant application requirements should be uniform, except to the extent required by disciplinary differences. If embargo periods are established before public access is allowed, these should be as uniform as possible. Again, differences in disciplines will be important and should take into account legitimate intellectual property interests of scientists and institutions in development of patents.

All agencies should allow use of grant funds to pay article processing fees so that publishers can recoup direct costs. Publishers should be required to handle public deposit, relieving researchers of the need to keep track of embargoes and deposit mechanisms. Deposit track records should be a factor in grant awards, both for researchers and for publishers – and these should be publically available information.

Central deposit would have the benefit of the ability to generate institutional and individual pages tracking content and contributions which could be used in institutional assessments. Standardized deposit mechanisms and metadata would be key to this kind of data reuse. Librarians and IT professionals should be involved in database design to make sure that material in the database is easily discoverable and interoperable with current and developing systems.

(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?

Key points

- All types of publications resulting from funded research should be made available open access.
- Books and conference proceedings may require different business models than journal articles.

Most types of peer-reviewed publications that result from federal funding should be made immediately available as open access. This would certainly include book chapters, articles in conference proceedings and technical reports. Different economic interests may be at stake, however, requiring a somewhat different analysis.

Academic authors are not usually financially compensated for journal articles. Textbook authors may, however, receive substantial royalties. If the foundation of the work is taxpayer funded research, the product of the research should be publicly available, but some portion of the grant award might be used to compensate both publishers and authors, at least in a transitional period. Online publishing, while not free, requires a significantly smaller investment from the publisher. One model for books/monographs that result from federal funding might be for a portion of the grant to be set aside as a subvention for a) publishers to produce the resultant publication and b) as appropriate royalties for the author. Another strategy is to publish with a publisher of open access books. A growing number of important presses are engaged in publishing OA books. See [http://oad.simmons.edu/oadwiki/Publishers of OA books](http://oad.simmons.edu/oadwiki/Publishers_of_OA_books)

Open standards should be developed for resulting e-books. Currently the e-book landscape is highly diverse, unstandardized and difficult for individuals to navigate. Publishers might be provided with incentives for innovative and universal solutions to these issues. The chaotic landscape provides publishers of vision with an opportunity for innovation.

Conference proceedings should also be made available if the research is publicly funded. Consideration should be given to impacts on scholarly societies. Revenue for associations with dwindling memberships is going down. Receipt of conference proceeding volumes is one of the incentives for payment of membership dues. Like publishers, societies may need to investigate providing different types of value and services to their members to maintain their income. If grant funding is allowed to compensate for journal processing fees, it might also subsidize open online conference proceedings.

(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?

Key points

- Access should be immediate if possible.
- The publishers' business models are evolving.
- Disciplinary differences will need consideration.

In terms of maximizing benefits of taxpayer funded research, the current traditional publishing system has faults that are ingrained by its origins in a print culture. Compared to the speed and global dissemination of an internet-based distribution system, it is antiquated and slow.

For maximum benefit to the public, free and immediate open access would deliver the most benefit.

Ideally there should be no embargo period – the results of federally funded research should be available immediately. Publishers should permit author open access self-archiving. However, for the near future it is likely that many publishers which allow some aspect of open access will demand a delayed access period in order to protect their sales. This model assumes that scholars want the most up-to-date articles immediately upon publication, and that educational institutions will continue to pay the subscription fees to access them. Scholars do not want to wait for the end of the embargo period to read journal articles. Thus embargoes, though beneficial for publishers, are problematic for scholars, students, and the public who have often paid for much of the research. They do damage to the principle of genuine open access and delay the benefits of publicly funded research.

The Houghton report stated:

Hence, a six-month embargo reduces the returns by around \$120 million (NPV). Of course, the impact of an embargo delaying open accessibility will vary significantly between fields of research and disciplines, having greater impact in faster moving fields of research and practice than in those where the progress of knowledge, application and practice is slower.

[The Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs](#)

In short, if embargoes have to be implemented, it is desired that the delayed access period be *as brief as possible*. An embargo period of twelve months or less has been voluntarily adopted by many commercial journal publishers. This indicates that they do not think a longer period is necessary for their business model. (See <http://highwire.stanford.edu/lists/freeart.dtl> and

<http://romeo.jiscinvolve.org/wp/2011/11/24/60-of-journals-allow-immediate-archiving-of-peer-reviewed-articles-but-it-gets-much-much-better/>)

Disciplinary differences may need to be adopted. The creation of preprint servers in some fields indicates that the value of information in some disciplines has a shorter half-life. That appears to be the case in physics and computer science (shown by extensive use of arXiv), in economics (RePEc contributions are high) and in law (extensive use of SSRN is common). The development of these repositories appears to be driven by inherent discipline-based need among scholars. It may be that public need, particularly in scientific disciplines where the stakes are high, necessitates consideration of shorter embargo periods.

If publishers are compensated differently for their contribution to the publishing enterprise than by the current subscription system, they would not need to delay access to allow for subscription. To maximize availability of publicly funded research, alternatives to the traditional subscription system must be found. In addition, it should be noted that library budgets have been severely impacted by rising subscription costs. Not only is this trend unsustainable but it is symptomatic of market failure. The users of the information (scholars) are largely insulated from the increased costs. An interesting analysis of this economic imbalance is available in an interview with Prof. Stuart Shieber, Director of the Office for Scholarly Communication at Harvard University.

<http://scientificdatasharing.com/general/interview-with-stuart-shieber/>

Given the conclusions in the Houghton Report, cited above, about the adverse effect of even a six-month embargo, any embargo should be as short as possible. Embargoes should be seen as temporary measures intended to help the publishing industry during a transitional period. It would be helpful to define the end date of that transitional period, by which embargoes would be phased out.

RFI: Beach Center on Disability, The University of Kansas

Background:

The Beach Center on Disability is a federally funded research and training center committed to making significant and sustained enhancements of the quality of life of individuals and families affected by intellectual, developmental, emotional-behavioral, and sensory disabilities through research and its application to the policies and service-delivery systems affecting them.

The Center's investigators receive or have received peer-reviewed competitively awarded federal funds from the National Institutes of Health (both Child Health and Human Development and Human Genome Project), Department of Education (Institute for Education Science, National Institute on Disability and Rehabilitation Research, and Office of Special Education and Rehabilitation Services), Department of Defense (Office of Community Services for Individuals with Special Needs), and Department of Health and Human Services (Administration on Developmental Disabilities). Its investigators consist or have consisted of individuals with terminal or next-to-terminal degrees in special education, developmental psychology, law, social welfare, and public health.

We are the co-founders and co-directors of the Beach Center and have our terminal degrees (Ann Turnbull, Ed.D; H. Rutherford Turnbull, LL.M.). We established the Center in 1988 with funding from the Department of Education, supplemented then and later by federal, state, and private funding.

We incorporate by reference the responses of the Provost, The University of Kansas, and we particularize our responses by linking them directly to our Center.

Question 1

Part a. We concur with the University of Kansas Provost's statements that (a) research articles must be made permanently and publicly available and (b) availability of research benefits several interested groups of research stakeholders (beneficiaries). These beneficiaries are (a) members of the research community themselves; (b) individuals with disabilities and their families; (c) professionals who provide services to those individuals and their families; and (d) federal and state policy-makers. We concur with the Provost for two reasons.

First, given our Center's mission, we are particularly determined to assure that our research is available, accessible, and appropriate to the four groups of research-stakeholders.

Second, our field entails a consistent and often unbroken chain between invention, innovation, intervention, and integration (the 4-I cycle). Basic research in disabilities constitutes invention; it creates data that are grounded on theories/hypotheses about human development and public services. Invention spurs innovation by stimulating new understandings and potential approaches to enhancing the quality of life of individuals affected by disability. These

understandings and approaches in turn lead to interventions through service-delivery that is authorized and funded by evidence-based policy. Finally, the interventions enhance the legal and actual opportunities of individuals with disabilities to participate in and contribute to the mainstreams of their communities of choice.

Part b. We concur with the University of Kansas' Provost's statement that public access to research creates greater opportunities for serendipitous contributions in research. But we add that such access also creates deliberate (as distinguished from serendipitous) opportunities related to policy and practice (the 4-I cycle, mentioned above).

When the 4-I cycle exists, research feeds on itself. Researchers learn from service-delivery professionals and individuals/families how relevant their research is to quality of life. That reciprocal research-to-practice model cannot exist unless researchers and practitioners, together with individuals and families, have easy access to new research results. In turn, policy makers in turn learn from individuals, their families, service-delivery professionals, and researchers alike "what works" and, thus fortified by evidence, are able to craft policies that authorize and appropriate funds for services and further research.

This interactivity solves or tends to solve some questions addressed by the research community, even as it unveils other questions for research. More than that, however, the interconnectivity creates new services within the nation's economy. For example, special education is a human service that has grown significantly over the last 30 years as a direct result of research that became the basis for policy and practice. There are numerous examples of the power of research to affect policy and practice; among them are research on educability, language acquisition, assistive technology, behavior modification and especially non-aversive, positive behavior intervention and support, and integration in schools and communities. Similar effects of research on policy and practice obtain with respect to rehabilitation, housing, transportation, employment, and homeland security.

Part c. Part b, above, describes the benefits of open-access policies. The costs of open access are relatively minor and consist largely of those that the research community meet when converting research findings into language that lay persons and policy leaders can readily use.

Part d. We concur with the University of Kansas' Provost's statement that immediate public access with re-use rights is essential to contribute to economic growth, augment the research enterprise, and enhance policy and practice. The longer the delay in the 4-I cycle (mentioned above), the lower the likelihood that economic growth and enhanced quality of life will obtain; instead, they may be impeded by the continuation of policies and practices that are deliberately segregating and demeaning, and often cost more than research-based practices.

When too-costly and too-segregating practices continue in spite of the existence of contrary data, scarce human-resources funding is wasted, economic growth in cost-effective services is delayed, researchers seeking to bolster their peers' more cost-effective and beneficial research are blunted in their efforts, and policy improvement is thwarted.

The continued use of institutional models – involuntary placement of individuals with disabilities into large, expensive, and generally ineffective institutions – proves the validity of these assertions. Likewise, the continued use of seclusion, restraint, and aversive interventions to shape the behavior of individuals with challenging behaviors further evidences waste of fiscal resources when, instead, less costly and more intervention-effective practices, grounded solidly in research, exist but may not be as accessible as they should be.

Question 2

We concur with the University of Kansas' Provost's answer to Question 2. We add that the data we and our colleagues develop and publish have a single purpose. It is to make significant and sustainable enhancements in the quality of life of individuals and families affected by disability.

Although our research undoubtedly adds to our individual standing (and our university's standing) in the research community, our research's real benefit is far beyond reputational. Our research asks questions and proposes answers; these answers in turn generate potential solutions (the 4-I cycle). Consistent with the Beach Center's mission, the greater interest to our intellectual property lies with individuals, families, policy makers, and practitioners, not with us and our publishers.

Question 3

We lack sufficient knowledge to answer Question 3 and therefore defer to the University of Kansas Provost's answer.

Question 4

We repeat our answer to Question 3.

Question 5

We adopt as our own the answer of the University of Kansas Provost.

Question 6

We adopt as our own the answer of the University of Kansas' Provost.

We add, however, that the end-beneficiaries of research – individuals and families, and, before them, service-delivery professionals and policy makers – are challenged by information-overload. When they seek evidence-based answers to real-world, practice-related inquiries, they may as well submit themselves to a fire-hydrant flow of data when in fact all they want is a carefully modulated, highly targeted drip of data.

Knowledge translation requires the research community to invent/create, then synthesize, and finally assist in developing research-based policies, procedures, practices, and personnel to enhance people's quality of life. Granted that knowledge translation usually is not the principal concern of the research community; it is, nevertheless, an element of the ethical obligation of that community to answer the "so what" question: So, now that there are data suggesting solutions in policy and practice, what may be done to make those data available to individuals, families, practitioners, and policy makers?

It seems utterly logical for federal agencies that sponsor research to require and provide funds for knowledge translation.

Question 7

It is essential to make available to the public book chapters, conference proceedings, technical reports, "white papers," and expert-witness testimony before Congress. Often these publications collect, analyze, synthesize, and thereby make more accessible to the beneficiaries of research the many peer-reviewed published articles that the research community produces. They are as useful, and sometimes they are more useful, to research beneficiaries than peer-reviewed articles because they present in a nutshell what researchers present in a forest of peer-reviewed journals.

Question 8

We agree with and adopt the University of Kansas' Provost's answer to this question.

Other Items for the Task Force

There are two other items for the Task Force to consider.

The first relates to knowledge translation, which we discussed above.

The second relates to accessibility of published research results and data to individuals who have visual impairments. Given that the data with which the Task Force is concerned are "published" in journals or other formats, these data regularly are inaccessible to visually limited individuals. They are, then, not truly published – not truly made public. Instead, they are quasi-published – made public in a way that only some people can access them. Visually limited individuals are researchers themselves; more often, they are individuals with disabilities or their family members, service providers, and policy makers.

The Task Force should address both knowledge translation and research accessibility.

Respectfully submitted,

Ann P. Turnbull, Distinguished Professor in Special Education and Life Span Studies

H. Rutherford Turnbull, III, Distinguished Professor in Special Education and Life Span Studies

*U:Open Access RFI Whit House Office of Science and Technology Practice
HRT III
19 December 2011*

Subject: Response to the OSTP RFI on Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research

Date: December 21, 2011 5:14:57 PM EST

December 21, 2011

White House Office of Science and Technology Policy
Request for Information on Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research.

Response from Arizona State University Libraries

(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?

For the purpose of this request for information, the primary step that agencies could take to grow existing and new markets related to access and analysis of peer-reviewed publications would be to develop policies that require free, unrestricted access to all articles resulting from federally funded research. Public access to this research will encourage and enhance scientific innovation and creativity by allowing researchers access to the most current research discoveries and enable them to leverage new information discovery technology and tools to make unforeseen connections, promote serendipitous discovery, and increase productivity. These policies can be used to grow the economy and improve the productivity of the scientific enterprise in a variety of ways. For example:

- Allowing broad and full reuse allows researchers to build on the results of others, rather than repeating research that has already been done. Researchers can also utilize new technological tools, such as applications that mine publication databases and produce meta-analyses of scientific results from multiple studies and datasets to make new connections and discoveries that would not be possible without computer assistance. Full open access to peer-reviewed publications encourages diversity, innovation, and efficiency, which improves the productivity of the scientific enterprise.
- Public access increases the pool of possible participants by taking advantage of unexpected readers who can innovate in ways not anticipated by the original researchers or their funders. These unanticipated users, who include commercial innovators, small companies, under-resourced researchers and potential collaborators, can hasten scientific discovery and innovation through engaging a larger pool of participants. For example, a 2006 study showed that sharing information about unsolved scientific problems to a large group of outside users was an effective and efficient means of finding solutions to problems that had not be solved internally (<http://www.hbs.edu/research/pdf/07-050.pdf>).
- Public access increases the global reach and impact of federally-funded research, thus enhancing the prestige of the U.S. scientific community.
- It can grow the economy through providing the opportunity for new information-based business models, as well as driving software development to facilitate new types of information discovery and tools for research.
- Faster commercialization spurs economic growth in broad sectors of the economy, such as

biotechnology, agriculture, and renewable energy, and encourages private investment in information technology to capitalize on government resources.

Some relative costs and benefits of such policies are as follows:

- Public access policies will increase the public's return on investment in research. It can be argued that the public currently pays multiple times during the research life cycle and may still not be able to access to the results. A public access policy could limit that to two (taxes that pay for the salaries and infrastructure for scholars and taxes that are used for large federal funding agencies), meaning that the public could immediately access the work once published without having to pay again. A relative cost will be reduced by a broader public access policy.
- Public access policies can reduce education costs by providing free access to authoritative scientific publications, while also engaging students and encouraging teachers to participate in the scientific community.
- Public access policies will increase government agency accountability and support informed, transparent, and evidence-based budget and policy decision-making by providing government agencies with an improved accounting of the outcomes of research they have sponsored, congressional budget drafters, appropriators, and authorizers with more reliable information to assess the value of existing expenditures and target funding priorities, and will promote more informed policy debate and enhanced access to the information underpinning policy decisions.
- An NIH-style public-access policy has proven to be cost effective: the NIH reports it costs \$3.5 - \$4.6 million annually (on a \$30 billion budget) to provide access to results of all their funded research. Therefore, an investment of about 1/100th of 1 percent of NIH's overall budget results in access to 2.2 million articles. As access and reuse is increased beyond the NIH model, that return on investment will rise even higher.

In conclusion, free, digital, online and immediate access is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise. Access to information without the ability to reuse or build upon it limits the value of that information.

(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 most important consideration regarding intellectual property protection lies with the researchers and authors, who are the initial rights holders and remain so until or unless they transfer their copyright to another, such as through publishing agreements. However, since most researchers are not paid for the production of scholarly publications, their primary concern with regards to copyright is its effect on their desire to have the greatest possible impact on their research field through attribution and citation; meaning they benefit most from the widest dissemination of their work. For this reason, public access does not threaten the incentives for the creation of scholarship at all - it enhances them because it supports faster research, more innovation and greater impact on the field.

Providing public access to federally-funded research requires no change in copyright law, but merely requires that the licenses that authors give to the agencies as a condition of their funding and the licenses that the agencies grant to the public facilitate both access and maximum reuse

of research information. A Creative Commons attribution license is an example of a license that would fulfill those purposes.

Agency policies should be mandatory for authors with a requirement to deposit their final (post-peer-review) author manuscripts, as this has proven to be successful for the NIH policy. Consequently, publisher agreements for federally-funded research articles should not be structured in ways that conflict with the licenses that researchers grant to the agencies. However, in order to ensure the continued value of the publisher's copy, we recommend that intellectual property policies require that the use of published scholarly works include a citation that refers to the publisher's version as the copy of record. This policy framework would balance the needs and interests of research authors, agencies, publishers, and the general public.

A copyright-related policy that should not be implemented is the current PubMed Central prohibition on systematic downloading, which needlessly restricts use, imposes unnecessary restrictions, and hinders scientific advancement.

(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 to federally funded research has many pros and very few cons. A centralized approach, managed by the federal government, would ensure that uniform standards for discovery, access, and persistence are maintained. Additionally, a centralized approach is more cost effective, both with regard to policy development and adherence, and in infrastructure development and maintenance. A standardized set of submission requirements facilitate policy compliance by making it simple for researchers to understand and follow a single set of workflows. As mentioned above, the NIH has proven that maintaining a central repository is very economical. This success could be expanded to include other federal agencies.

However, the most important aspect of public access to federally-funded research is the ability to search across a broad and interdisciplinary set of databases and have full and unrestricted access to the results. With this in mind, a decentralized approach involving multiple repositories might be a viable solution. In order to facilitate access and use of federally-funded research, federally-managed standards for 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 would be required. It is important to consider that, even with carefully crafted regulatory requirements, it is more difficult to establish and maintain strict standards under a decentralized framework that includes partners outside the federal government, which is a considerable downside to a decentralized approach. Even if a decentralized approach proved feasible, it would be desirable for the federal government to maintain mirrored and accessible versions of any participating repositories in order to protect the public's investment and ensure accountability.

The federal government has a long-term interest in making such works permanently available, and is the only current entity with the interest in and capability for making federally funded works publicly available, in providing tools for use, and in guaranteeing that new services and

products can be built from publicly funded information. It is the only entity with the ability to establish and enforce standards of interoperability. The federal government is therefore the most appropriate body to maintain custody of publicly funded articles.

(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?

Academic and research institutions have already partnered with many research funders to provide permanent archives and have proven their expertise and focus on long-term preservation and access. The DRIVER project (<http://www.driver-repository.eu/>), funded by the European Commission, is an excellent examples of a decentralized, federated repository structure involving cooperation from universities and research institutes in several European countries.

Publishers could be encouraged to participate in public-private partnerships by voluntarily providing the final published versions of articles. However, given their focus on immediate income, commercial publishing firms in particular should not be relied upon solely for digital archiving. To ensure long-term preservation, under no condition should a publisher site - or any other stakeholder site - be the single point of access for publicly funded articles.

As mentioned above, regardless of where partners are found, standards for access, preservation and interoperability must be maintained. Since a primary goal is to encourage innovation in accessibility and interoperability, it is important that creativity and experimentation be allowed in the tools that will be used in searching the publicly accessible articles. The ability to create new kinds of search tools, and to compute on the contents of these repositories (text-mining, etc.), must be afforded to users in order to exploit the full range of opportunities for innovative research and economic 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?

Federal agencies, publishers, and scholarly and professional societies can encourage interoperable search, discovery and analysis capacity across disciplines and archives by involving agencies that have experience in the development of metadata schema, such as NISO and the Library of Congress, in the creation of appropriate standards. Using previously developed models means that researchers and librarians, as well as automated systems, already have search techniques that work. At the minimum, core metadata standards for scholarly publications should be designed to facilitate the functions of use, reuse, and analysis described in previous responses.

Federal agencies can 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 by creating specifications that support multiple metadata standards. Some of the primary goals of the specifications (along with examples of related standards) would be to:

- Provide institutional information for published sources (grant IDs, funding organization, I2 – Institutional Identifier, etc.),
- Provide descriptive information for both the repository and published versions (Dublin Core, ORCID),
- Support searching through both key words and controlled vocabulary schema appropriate to disciplines,
- Incorporate abstracts,
- Facilitate full text searching and web crawling,
- Support metadata harvesting (OAI-PMH),
- Establish relationships through semantic web standards (RDF),
- Support usage tracking (COUNTER),
- Support description of related data (DataCite Metadata Schema),
- Support data exchange standards (JSON),
- Document IP rights.

(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 that fund science can maximize the benefit of effective public-access policies to taxpayers by ensuring that complete results are widely available in a timely manner. To this end, any successful public-access policy must be consistent in its requirements and mandates. Researchers often hold grants from multiple agencies concurrently, so uniform requirements and procedures regarding deposit of peer-reviewed literature should be established across all funding agencies involved. In particular, a standardized set of criteria to determine what materials need to be deposited and how deposit should occur will reduce the complexity and cost and increase the rate of compliance. Such uniformity will also ease the burden on institutional and university repositories who often offer assistance to their researchers in meeting the deposit requirements for granting agencies such as NIH.

In order to minimize the burden and costs for stakeholders, federal policies should take advantage of existing protocols, such as SWORD, to facilitate automatic deposit of manuscripts to multiple repositories and should encourage development of additional tools and services. Additionally, licenses should be standardized to include the maximum possible reuse rights, such as those embodied in the Creative Commons CC-BY licenses, for the public.

Policies that integrate articles with grants management systems, both internal and external, will improve agency accountability, providing the public with more information about the results of their investment in scientific research. Policies can create opportunities to create enhanced productivity management tools for federal and internal reporting, such as enhanced bibliographies and primary investigator (PI) profiles. These tools would offer new opportunities for universities to better measure research output and promote branding of research.

It is important to note that these policy changes would address a burden that has been on the public for a long time - the lack of access to the research it funds. Without public access to federally-funded research, the public, small businesses, K-12 and higher education has borne extra cost burdens without the benefit of access to the products it funds.

(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 majority of scientific knowledge resulting from federal funding appears in form of peer-reviewed journal articles, the primary mechanism for scientific communication. As noted above, dissemination of the results of federally-funded research is 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 concentrate their public access policies on journal articles as well as related supporting materials that document the research process, such as data, protocols, survey instruments, etc., and facilitate the replication of results.

Educational materials, such as book chapters, texts, and conference proceedings, resulting from publicly funded research should also be made accessible to the public. However, different conditions may apply to different types of materials. For example, authors are generally not paid for journals articles, but may be paid for textbook chapters. Any policies should reflect these differences.

In any case, the immediate effort should stay focused on the primary mechanisms that scientists use to communicate their results, namely peer-reviewed scholarly journal articles, and the current limitations on access to such publications. The complexity of providing access to other materials should not be allowed to delay the implementation of a mandate for public access to funded articles across federal agencies.

(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 mentioned in comments above, immediate access to the full content of peer-reviewed scholarly publications resulting from federally funded research is strongly recommended. Any embargo period diminishes the possibilities for innovation and growth that a mandate is trying to realize. In economic terms, each day of an embargo period represents lost opportunity costs that decrease the return on investment for the policy. For this reason, an embargo period is a compromise and should be as short as possible.

Empirical evidence supporting the value of public access is plentiful and growing all the time. Therefore, the burden of proof for the need for embargoes should be provided by those who desire them. For example, there is no evidence that libraries have cancelled journal subscriptions because of publicly accessible research articles. Before this and similar assertions are allowed to impose limitations on public access, evidence should be required to demonstrate actual harm. Where no such evidence can be provided, immediate public access should be the granted as the best way to foster innovation, competition, economic growth and scientific progress.

If it can be demonstrated that an embargo is necessary, a uniform embargo period of six months or less should apply across all funding agencies.

These comments are submitted on behalf of Arizona State University Libraries by:

Anali Maughan Perry
Collections and Scholarly Communications Librarian

Sherrie Schmidt
University Librarian



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

About IFT

For more than 70 years, IFT has existed to advance the science of food. Our scientific society—more than 17,000 members from more than 100 countries—brings together food scientists and technologists from academia, government, and industry.

By advocating for the science of food, we educate the media and policy makers, and serve as a catalyst for new ideas that benefit the consuming public. Our community's shared commitment to our mission helps to ensure a safe and abundant food supply contributing to healthier people everywhere.

Office of Science and Technology Policy
725 17th Street
Washington, DC 20502

Re: Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research

Submitted by email to: publicaccess@ostp.gov

The Institute of Food Technologists (IFT) appreciates the opportunity to submit comments to the Office of Science and Technology Policy (OSTP) to offer its perspective and practical insights on public access to federally-funded research appearing in peer-reviewed journals. IFT exists to advance the science of food and we are committed to the free flow of scientific information. Furthermore, our peer-reviewed publishing efforts are critical to our nonprofit organization's success in fostering innovations in the food science and technology field.

IFT serves more than 17,000 members, affiliated with academia, industry, and government, and all those interested in food science and technology by publishing three internationally renowned peer-reviewed journals and a technical magazine. As a publisher, IFT assembles more than 1,000 preeminent food scientists, technologists, and engineers among our comprehensive pool of peer reviewers. Such an extensive resource of peer reviewers ensures that the research made available to the scientific community is important, comprehensive, and of high quality and integrity.

Two of IFT's peer-reviewed e-publications (*Journal of Food Science Education and Comprehensive Reviews in Food Science and Food Safety*) are currently freely available online. In addition, the *Journal of Food Science* is available to IFT members at a discounted subscription rate and to others on a per article download charge or through subscription. *Food Technology* magazine, published monthly, is available to IFT members in print and online. The online version of *Food Technology* is initially available free to the public for about one month, and then becomes accessible only to IFT members. To ensure that the public is informed of research findings, IFT, along with publishing partner Wiley-Blackwell, has a well-organized system to promote particularly newsworthy research through popular media and news releases.

Based on our 70-plus year history of scientific publishing, IFT strongly believes that the current system for handling and releasing peer-reviewed research—federally-funded or otherwise—is not broken and does not need a federally-mandated open access policy.



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

Much of the discussion around the issue of public access was fueled by an NIH decision in February, 2005. In its "Policy on Enhancing Public Access to Archived Publications Resulting from NIH-funded Research," NIH-funded researchers were requested to submit electronic versions of the author's final manuscript upon acceptance for publication to PubMed Central to be released to the public as soon as possible and within 12 months of the publisher's official date of final publication. IFT is not aware of any studies on the effectiveness of this voluntary open access model.

Perhaps the most important quality that can be attributed to a research publication is whether or not it has been subjected to peer review. Through the peer-review system, there is an opportunity for scholars to collectively make a very good scientific manuscript become a very great work through a thoughtful editorial process. The important exchange of ideas between editor and author ultimately benefits the public with a more comprehensive scientific work.

Peer review is carried out through the auspices of the publisher, and its associated costs are borne primarily through subscriptions. If the publisher is a professional scientific society, as IFT is, however, some expenses may be borne through member dues as a member service. In either case the costs associated with handling the manuscript including peer review and copy editing are real costs. Without some grace period prior to free access, the publisher will be forced to charge the authors page charges (which some publishers now do). Such a move is likely to delay publication of research results, an unintended consequence of a public access policy.

In addition, IFT offers the responses to these points below for OSTP consideration.

(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?



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

Most if not all publishers are moving to digital access, which reflects a concerted response to meet the information demands from the scientific community. Given the rapid evolution of digital communications, there have been dramatic improvements in access to scientific information, reprints and authors due to the ongoing improvements in search engine technology. Therefore, from a publishing and a scientific perspective, it is important that there are not mandates establishing a single approach to public access since it has the potential to stifle innovation in what is now a rapidly changing environment. Not only does an open access system have the potential to decrease the amount that publishers are able to invest, it also may inadvertently reduce incentives to develop new tools, delivery vehicles and functionality—all of which are critical to the free flow of scientific information.

The assumption that there are scientists who are restricted in their access to research may be erroneous. To date, IFT is not aware of any definitive studies that validate this hypothesis. Surely implementation of a policy that would have the unintended consequence of irreparable harm to the peer review system would not be undertaken without sound scientific data to justify the action. The peer review system functions through scientific journals, many of which are supported through scientific societies by subscriptions from individuals, libraries, and corporate entities. Mandatory release of research at the time a manuscript is accepted for publication would undermine subscriptions and result in loss of revenue for carrying out the peer review activity.

(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?

There is a valuable opportunity to forge a more cohesive public-private partnership between the government, publishers and scientific organizations. For example, new opportunities could be created to enhance access to published articles. In this regard, the government could make additional funding available for authors to publish open access articles. Content licensing arrangements with publishers and scientific societies could make more information available to specific audiences. Make the funder-collected and maintained outputs of taxpayer-funded research, including grant reports



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

or research progress reports, freely available to the public. Work with private sector publishers to make that content findable and link it to the journal literature. While new public access approaches offer interesting opportunities for the future, these must be weighed with the current realities of patent law and the history of a very effective process that authors and publishers have forged to unveil the first public announcement of a scientific work. From a patent law standpoint, public access could generate considerable legal challenges. For example, there are a host of intellectual property issues associated with making progress reports for government funded research available to the public and even more confusion could reign over turning government-funded science into intellectual property. All these issues must be viewed through the lens of our current patent system.

(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 advent of the internet has made it possible to deliver journal content in new ways, and many publishers have made considerable investment in new publishing technologies, methods to archive and backup up articles, and the creation of new technologies to deliver published material. In regard to the government's role to serve as the steward, it may be costly to maintain and manage a centralized government system since it may be cost prohibitive to keep pace and maintain compatibility with rapidly evolving technologies. In this regard, new investments by publishers and societies may allow for increased access through new media.

Based on question #3 above, there needs to be greater clarity and definition on government's role in "stewardship" and in maintaining "custody." This implies the potential for government's role to potentially override copyrighted material that has been developed by publishers and scientific societies. A contrary view is that government's role is not to be steward and manager of a centralized system. Instead, it is in the public's best interest if the government plays a role as a contributor to the scientific dialogue—to complement the research article contributions of industry and academia—by encouraging broad distribution of federally-funded research through publishers and societies that develop peer reviewed journals. This increases



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

the reach of federally-funded research, and it also generates opportunities for broader interpretation of research through the peer-review process. Instead of mandating that manuscripts are freely available within an archive, it may benefit the scientific community and the public to provide public access to research reports and other information that it already controls.

(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?

IFT is working with our publishing partner to better standardize and track funding agency information and support initiatives such as ORCID and semantic content mining for increased clarity, discoverability, and efficiency in scientific publishing. If it is a high priority for the government to raise awareness of federally-funded research, it would be valuable for government agencies to make more funds available to research authors to publish their articles open access in the journal of their choice. For example, the *Journal of Food Science* has a program for authors who choose to pay for open access for their articles within the journal (OnlineOpen). Many journals currently offer this “hybrid model” option, and it also supports the comprehensive peer-review processes that are already in place.

(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 the model that is described in Section #4 above.

(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?



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

Publishers and learned societies, like IFT, are committed to dissemination of research, which is precisely why we publish journals. In that regard, the current model has proven highly effective since it provides broad, comprehensive access to peer-reviewed journal literature. Therefore, utilization of the research by taxpayers already occurs in the current model. Broadly, the key issue is how to make a system that is economically sustainable, maintains a permanent record, and is totally accessible. Previously, IFT used a model whereby the author, usually through the funding agency, would pay a page charge for having their research published. After careful consideration, this system was considered unsustainable and so IFT adopted the "no page charge" model for IFT members. Included in this model is the option for the author and the funding agency to immediately subscribe to open access through a charge system. IFT is committed to the proposition that authors are free to have their research peer-reviewed in the journal of their choice so as to maximize the impact of their research. To make the funded open access model viable, funding agencies—both federal and non-federal—would require funding to cover publishing fees. Under this model, there is likely to be less funding available to actually do the 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?

This notion undercuts efforts by nonprofits, such as IFT, and thwarts efforts designed to produce new content in new media. It would be cost prohibitive for nonprofits to continue publishing or conducting conferences with an added Federal mandate to include book proceedings, scientific sessions, proceedings—especially when considering the rapid development of new communications technologies that currently are successful in making this information available to foster a scientific discussion.

(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



Institute of Food Technologists
1025 Connecticut Ave., N.W., Suite 503
Washington, D.C. 20036-5422 USA

+1.202.466.5980
+1.202.466.5988 Fax
ift.org

period should be different for specific disciplines or types of publications?

To the best of our knowledge, there has not been extensive review of the effectiveness of embargo systems and whether there is a single monthly timeframe that would be appropriate across a multitude of research disciplines. In other words, what may be an appropriate embargo time frame for the physics field may not be suitable for food science field.

However, if the government deems it necessary to establish such embargo parameters, it may be feasible that a 12-month embargo time frame could be implemented so that publishers and scientific societies, such as IFT, do not experience financial hardship as content is made available freely. A key issue for the government to consider is that accepted-but-unedited content may contain mistakes and cause confusion since it is not the official version of record. Scientific societies and publishers, with their expert peer reviewers, provide a valuable service to their fields by editing, polishing, publishing, and maintaining the version of record, and the government must consider the impact of embargo periods and the resulting impact on scientific accuracy of articles.

Food scientists and technologists have been leading innovations for generations and the peer-review system has served the profession and the public very well. In regard to public access policies, it is important that we do not impose federal mandates on publishers and scientific societies since our current system is not broken.

We appreciate the opportunity to share our insights and we look forward to continuing to contribute to this important dialogue.

Daryl Lund
Editor-in-Chief, IFT Scientific Journals

Subject: OSTP response

Date: December 21, 2011 10:26:30 PM EST

Dennis Vance, PhD
Distinguished University Professor
Scientist of the Alberta Innovates - Health Solutions

Department of Biochemistry and Group on the Molecular and Cell Biology
of Lipids
School of Translational Medicine
University of Alberta

1.

The current system for publishing works very well. Scientists have the opportunity to publish and pay for free access by readers. Or they can use commercial publishers where readers/libraries pay the bill. Someone has to pay the bill.

3.

The governments have no business in trying to manage publications. The free market can and does take care of this very well.

6.

Federal agencies don't need to be involved. The system works very well. If it ain't broken, don't fix it.

8.

My understanding the embargo period is 6 months and that seems appropriate.

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

Date: December 22, 2011 7:59:33 AM EST

Sent to: publicaccess@ostp.gov.

Phil Daly, Senior Publisher, Wolters Kluwer Health Ltd (UK)
Response to the Office of Science and Technology Policy 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 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 expertise for delivering content already exists and agencies need to work with publishers to facilitate access without undermining the economy and productivity resulting from the access of research papers through the publishing industry. The value of access is diminished without the process of peer-review and moulding of content that the scientific/medical community have adopted and publishers facilitate and support. As such we help facilitate access to what is important. As a publisher of learned medical information we strive to serve the healthcare professionals in academia and industry by disseminating content that helps them explore, fuel discovery, or enhance patient care. We are committed to the wide dissemination of content and the delivery of electronic research information for medical and academic institutions, and corporations globally. As a leading publisher we continue to invest in the development of services and technological advances that makes dissemination and interlinking of content possible. The current access to important content

<http://www.publishingresearch.net/projects.htm> (Access vs. Importance) is fuelled by the publishing industry, an existing economy, supporting review and dissemination of scientific publication for subsequent scientific and medical progress. Undermining these existing services potentially undermine the development of expertise and incentive of publishers to continue to invest in a rapidly changing digital world that seems essential for the stimulation of competitive future economic growth.

(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?

There are a number of existing publishing models for dissemination of research papers: from subscription, to author pays open access, licensing of content, and deposit to repositories. These will continue to evolve, but they must be sustainable and financially viable forums, and one size will not fit all scientific/medical disciplines. Publishers have invested in the development of services and technological advances that makes review, dissemination and

interlinking of content possible, and will need to continue to do so in a rapidly changing digital world. The value added by publishers does not begin and end with the finished published version of record of a paper, but begins with the ease by which articles are submitted online for consideration to an indeterminate end (as articles are increasingly supplemented with additional material including reference to data that a paper is based on), and of course they are contained within databases interrogated by systems that allows efficient identification of important research. Appropriating published articles derived from data and research reports would undermine the value of originality that is important to publishers in a competitive market, and subsequently the income derived that is required to support peer-review, dissemination and investment.

(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?

Aggregation and linking of content already exists through the services provided by individual publishers and aggregators of content from multiple publishers, and cooperation between publishers. Preservation and correction of the publication record defines the role publishers have played from their formation, and initiatives such as Portico and CLOCKSS provide preservation insurance. Publishers have invested in digitalising archives of content back through decades and even centuries, and it is part of their *raison d'être* to maintain the publication record now and in the future. Cross reference of published material is through an organization founded by publishers, Crossref (www.crossref.org), and establishment of digital object identifiers provide unique labels for papers in a burgeoning internet. Decentralisation of content provides incentives for competition between publishers as systems are developed for competitive advantage to drive accessibility and usage; a centralised system would not have that competition and there would be less incentive to be innovative. Innovation has catapulted a number of technological companies, and in a global market that needs to be preserved if one country is to compete against another.

(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 expertise for delivering content already exists and agencies need to work with publishers to facilitate access without undermining the economy and productivity resulting from the access of research papers through the publishing industry. Publishers already make that investment in peer-reviewed literature as we work with the scientific/medical communities to establish mechanisms, processes, and financially sustainable publication models to publish important research. Why do the same? And what would be the incentive to develop those services when innovation and progress is already being driven in an external, competitive, and evolving market that already includes a number of publication models from subscription to authors pays. Making original research reports required by Federal agencies freely available may be a simple way of achieving "public access", as distinct from the published research paper that is derivative of the work.

(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. As publishers we invest heavily in all types of content we produce, whether they are journals or books, and the same principles of financial sustainability of a product would apply. The availability of some content that would be freely available for “public access” would undermine viability, the investment and value added by the publisher to a product, and value of the content to the scientific community we would aim to serve.

(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. An embargo effectively shortens the life and value of an article to a publisher who has invested in its peer-review and delivery, and as such reduces the return on investment and incentive to continue to invest for the future. Publishers should be able to determine these embargo periods themselves based on the return on investment and taking into consideration the differing half-lives for various markets and disciplines.

Yours sincerely

Phil Daly

Phil J. Daly, PhD
Senior Publisher
Wolters Kluwer Health Medical Research
Lippincott Williams & Wilkins
Ovid Technologies



22 December 2011

Re: FR Doc. 2011–28623

Dear Sir or Madam,

I am writing to respond to OSTP's November 3, 2011, "Request for Information" (RFI) regarding "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research." The American Geophysical Union (AGU) is a 501(c)(3) corporation with more than 60,000 members whose research interests take in a broad range of earth and space sciences. AGU publishes 19 peer-reviewed journals across these scientific disciplines. AGU employs 180 staff including 120 engaged in publishing.

As a scholarly society publisher, our goal is to expand access to the latest breakthroughs in the scientific research and developments in academic thought. The main purpose of a scholarly journal is to publish articles that analyze and interpret original research or experimentation in order to make such information available to the global research community. We are uniquely positioned to help the federal government in expanding public access to publications that describe and interpret federally funded research, ensuring the long-term stewardship of these publications, and supporting innovation and economic development derived from scholarly discovery.

AGU has contributed significantly in support of the scientific community by expanding accessibility to content, improving interoperability, and fueling innovation. We have invested in a digital platform with evolving Web capabilities to provide students, faculty, researchers, and other interested users with faster and more robust delivery of earth and space science information, including dynamic ways to present data and scientific articles. Working with others in the publishing community, we have improved interoperability through new metadata standards and pilot projects, which are driving innovation and providing for better information discovery and expanded use of research results.

The core publisher activities of supporting peer review, ensuring the continued integrity and reliability of the scholarly record, formatting this content to make it accessible to users worldwide, and preserving the scholarly record for future generations do not come without costs and ongoing investment. These activities are threatened by access policies that do not take these costs into account. In considering policies that could potentially expand public access to research results, it is critically important that any new policy does not damage the private institutions on which the federal government and its scientific enterprise depend.

AGU is not unique among scientific societies in that a major portion of the funds supporting the activities of the members through meetings, programs, and outreach is provided by net revenue earned from publishing operations. Beyond the obvious cost of loss of jobs if revenues are

impacted through federal government mandates with an adverse impact on publishing, the entire scientific program of AGU is at risk. Multiplied across the hundreds of society publishers similarly dependent on publishing revenues, a major portion of activities in the scientific community is at risk.

A federal agency public access policy that is sustainable in the long term and maximizes benefits to researchers and the public at large will function as a balanced public-private partnership to enhance access and interoperability; adequately protect fundamental intellectual property rights; and respect proprietary contributions of added value to ensure sustained private investment in innovation. This approach meets the needs of the scientific community by relying on evidence-based assessments and providing access to taxpayer-funded research results through both public and private channels.

The America COMPETES Act, which established a public access policy for research funded by the National Science Foundation (NSF), provides a constructive model that can be replicated in a timely manner at other federal agencies. This is to be contrasted with the NIH policy, which has the potential to significantly damage a well-functioning system of communication for scientific discovery and innovation, reduce economic benefits and employment, and undermine intellectual property. Each of these models should be carefully analyzed to be sure its long-term impact on all stakeholders is fully understood.

The America COMPETES Reauthorization Act of 2010, signed by President Obama earlier this year, calls upon OSTP to coordinate agency policies related to the dissemination and long-term stewardship of the results of federally funded unclassified research. We strongly support this goal, as well as the guiding principles of transparency, participation, and collaboration that President Obama and OMB Director Orszag articulated at the onset of this administration.

Scholarly publishers strongly support the view that the federal government should be guided by “principles of transparency, participation and collaboration” as noted in the Transparency and Open Government Memorandum and Open Government Directive. We stand ready to work in collaboration with all partners to ensure the continued success, vibrancy, and innovation of the U.S. scientific community.

In addition to the general suggestions above, we would like to comment specifically on some of the questions outlined in OSTP’s November 3 Federal Register Notice requesting public 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?

AGU provides a robust, innovative market for peer-reviewed publications around the world that has led to multiple channels of access to and analysis of research across a broad array of scientific disciplines. AGU has worked actively with library consortia, both in the United States and globally, to disseminate our content to libraries that on their own would not be likely to subscribe to our content, thereby making content available to users who would not otherwise have access. These and other publisher initiatives have helped accelerate and broaden electronic access to the peer-reviewed literature.

AGU is willing and able to work with all stakeholders to address existing or future gaps in access. 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. Researchers, the general public, funders, students, teaching faculty, and others each have different information requirements. AGU is committed to identifying and addressing these access gaps.

Some options to broaden access to materials that analyze and interpret research for scientists and the public:

- Work to develop standards for data and metadata to make research more readily searchable and discoverable. AGU is already working with other publishers to develop standardized information and collections through initiatives like CrossRef.
- Work with researchers and other stakeholders to create appropriate policies that recognize costs to make the federal agency–collected and -maintained outputs of taxpayer-funded research, such as grant reports and research progress reports, freely available to the public.
- Make funds available to support payment for services publishers provide to authors so authors may elect to provide unfettered public access to their articles. Publishers provide valuable services in managing the peer review process, disseminating research, archiving content, and collaborating to make discovery of science easy and complete. Publishers are currently compensated for those services principally through subscription revenue; a shift to compensation from authors for these services has some inherent faults but is an alternative that needs to be explored. Several research funders already do this (e.g., Howard Hughes Medical Institute, The Wellcome Trust, Max Planck Institutes).
- License content from publishers to make it available to specific audiences. Publishers license content to customers of many kinds and can generally customize those licenses to meet specific or specialized user needs, including those of government agencies, and have the ability to ensure the availability of their content with existing infrastructure.

Public access government mandates have significant costs to the U.S. economy and the scientific enterprise:

- The data from PubMedCentral at NIH indicate that two-thirds of users are from overseas, undermining critical export opportunities for an \$8 billion publishing industry that employs 50,000 Americans. Those users are being supported by U.S. taxpayers; current technology capable of limiting access to U.S. users only would be expensive to implement and would be easily circumvented.
- Significant value added by the publishing industry could be eliminated if revenue channels necessary for publishers to reinvest in their businesses and innovations continue to be threatened by government-mandated access policies that provide free access to publishers' works and enable piracy and unauthorized reuse.
- Even as the U.S. government is considering policies that would have significant long-range costs to the government by creating repositories that must be maintained and migrated as technology changes, there have been reports of library closing and other budget cuts. Publishers provide a robust archive of articles through their own databases, through shared archiving in projects such as CLOCKSS (<http://www.clockss.org/clockss/Home>), and through commercial archives such as Portico (<http://www.portico.org/digital-preservation/>) as one of the many services we provide.
- We do not yet know all of the impacts of the NIH policy, but there are significant concerns that it may undermine U.S. competitiveness and negatively impact U.S. jobs.

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 best way for the federal government to protect the intellectual property interests of all stakeholders in the scholarly communication process is to work in collaboration with researchers to create appropriate policies for making public the research progress reports that it receives as a condition of grantmaking. This approach would include free access to all final project reports and citations of published research documents publicly available via the Internet. It would ensure that public access policies do not undermine the peer-reviewed scientific journals in which scientists publish.

Any regulations should be constructed 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. We have seen that the NIH policy undermines intellectual property rights and promotes pirating. Not only is PMC undermining an important U.S. export market, but PMC copies of copyrighted material appear on other unauthorized sites, contributing to millions of dollars in annual losses to U.S. publishers.

The federal government should:

- Provide open access to a final research report, rather than asserting a type of eminent domain over the peer-reviewed journal article. This solution would allow standardization of information reported, rapid and broad dissemination of the government-funded materials even before publication of a peer-reviewed article, and the preservation of intellectual property. The reports, which are already mandated and filed as a part of grants, should be open to the public and would constitute the resolution of the requirement that results of publicly funded research be made publicly available. The further value that publishers add is outside of this mandate and would work in conjunction with this resolution, but AGU and other publishers are willing and eager to cross-link between the reports and final published articles. The scientific record would be enhanced by this linkage, and users would have an option if they do not have access to the published article.
- Support the continued operation of various models of publishing to ensure access to innovation and the ability of researchers to publish in the venue of their choice. The traditional subscription model successfully serves many researchers, users, and libraries both in the United States and internationally. A different model, often titled Open Access, combined with the strategy of authors compensating publishers for valuable services has gained some traction in recent years. There may be a different model that evolves or the market may settle on one of these, but the government should not mandate a specific model, especially with imperfect information about the consequences of such an action; AGU is already experimenting with alternative models and will experiment with new models as they arise.

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 many opinions about whether centralized government control of scientific publications and the tools by which they can be searched and analyzed is consistent with democratic and free market principles. Centralized, government-controlled custody of publication carries significant risks and few benefits.

Long-term stewardship of content carries significant costs that are already being borne by publishers. In an era of dwindling federal resources, central federal repositories are duplicative and an unnecessary expense and recurring burden that may not be viable for long-term stewardship. As noted above, federal agencies are already being forced to close libraries; are the agencies involved able to commit funds far into the future to maintain and evolve depositories of electronic articles? What impact will shrinking budgets have on various aspects of repositories? How will these repositories be integrated with each other? With nongovernmental repositories?

With current publishers' repositories? In the latter case, even if all current content funded by the federal government were located in government repositories, previous content still has significant intellectual value to researchers and must be integrated to be effectively used. With multiple sources of scholarly publications, many of which are not based on government-funded research, partnerships among stakeholders are essential for achieving effective access to literature that represents the latest scholarly discoveries.

A centralized governmental approach will deter private sector innovation by establishing unnecessary levels of oversight and bureaucracy that stifle creativity. It will also minimize scientific and commercial opportunities by reducing potential traffic to innovative new applications that facilitate the work of researchers. The government must be prepared to invest in technology to provide features, functionality, services, and innovations that match those provided by AGU and other publishers.

The publishing industry is already doing a good job of promoting interoperability, search, development of analytic tools, and other scientific and commercial opportunities. There is no reason to doubt that publishers will continue to provide innovative products and services, unless their financial livelihood is undermined by harmful policies.

A competitive publishing environment of not-for-profit and for-profit organizations — all of whom must receive a return on investment to survive — has led to robust technology development in scholarly publishing during the past 20 years. This sector of the publishing community, which includes professional associations, commercial publishers, and university presses, moved quickly and decisively to introduce new technologies that meet researchers' demands for faster and more user-friendly delivery of scholarly information. That innovation has not stopped; new ideas are constantly being evaluated and field tested. Many of these ideas add value to the scientific enterprise by enhancing the reach of scholarly communication, making data more easily discovered and analyzed, and by expediting the dissemination of results of research to an ever-expanding pool of scientists. Publishers over the past decade have developed the Digital Object Identifier (DOI), a unique identifier for each piece of content in a scholarly publication. In partnership with stakeholders, we are continuing to innovate in the creation and standardization of metadata to make it easier for researchers and the public to find and use scientific research information.

A collaboration of publishers, librarians, and database providers established 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. This enables libraries to better understand how the digital collections are being used, and it allows publishers to better understand the usage patterns of their digital content. COUNTER standards continue to be refined and extended to provide ever more information about usage of published content.

Internet search engines, abstracting services, and other tools do an excellent job of ensuring the discoverability of research, and innovations in this area occur every day without government interference.

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?

AGU is engaged in a Microsoft research project to create a new search and discovery service (Microsoft Academic Search) that will add another service for searching articles and mining text and will provide linkages among authors. The advantages of such a service to AGU constituents include the following:

- Enhancing discoverability of articles.
- Enabling editors to identify reviewers in areas that are underserved or emerging.
- Enabling scientists to identify possible collaborators for research.
- Enabling scientists to identify emerging areas of research and those who are working in those areas.

AGU staff has participated in workshops with NSF and NAS to explore what options fit the mandate of the America COMPETES Act to create the greatest benefit to science without harming publishers, especially society publishers. AGU seeks out opportunities such as this request for participation and will continue to work formally and informally with interested stakeholders. AGU regularly consults with our library community on the needs of researchers and other constituents, and to understand the changing needs of librarians in this dynamic economic environment.

Journal publishers are actively working with federal research agencies to develop and implement multiple collaborative projects that will enhance the public access, utility, and preservation of materials that report on, analyze, and interpret federally funded research, including progress reports, scholarly publications, and data for use by both the research community and the general public.

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 dedicated to the widest possible dissemination and discoverability of publications that analyze and interpret research. Partnerships with industry are already underway to determine, develop, and include appropriate metadata in publications. Examples of such metadata initiatives include the following:

1) Standards and persistent identifiers to enhance the discoverability of publications that analyze and interpret government-funded research and to promote interoperability among the funding agencies, publishers, and any third-party databases and platforms:

- a. Standardizing and facilitating funding agency information. Publishers are collaborating with agencies to create a pilot initiative to clearly indicate the funding agency responsible for research described and analyzed in a scholarly publication or an associated data set, giving the research community and public easy links to a variety of access options on the publishers' site. Working with the publishing community to gather and link this information will save agencies considerable effort and expense compared with producing or maintaining such information or services on agency websites.
- b. DOIs for data sets and article supplementary material. There is considerable opportunity for strengthening the multiple organizational partnerships that already exist to promote the identification, discoverability, and archiving of data, including Datacite (www.datacite.org) and the NISO/NFAIS Working Group on Supplementary Journal Information (www.niso.org).
- c. Author and institution disambiguation. Name ambiguity and attribution are persistent, critical problems embedded in the scholarly research system. The Open Researcher & Contributor ID (ORCID) project (www.orcid.org) is a successful public-private partnership with 275 participating organizations, funded by \$2M in loans from publishing partners and building on successful investments by publishers in the past. A pilot demonstration began earlier this year and is on schedule, and institutional IDs will be addressed in a second stage.

2) Discovery tools to facilitate journal content mining, access dark archives, and improve data management:

- a. Content mining. Content mining projects could be developed as collaborations between publishers and federal funders. Publishers are already working on projects to mine journal and book content, and it might be helpful for federal funders to develop a content mining demonstrator to illustrate the value of content mining to the broader scientific community.
- b. Author-driven data management. For many years, publishers have produced and archived data-specific journals, and they are maintaining and updating such data sets with DOIs and semantic tagging. There are also many examples of pilot projects on data management.

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.

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 should maximize public access to the federally funded research in which taxpayers invest. However, it is publishers that invest in the systems and processes that result in

peer-reviewed literature. The federal investment is in research, not publication, with significant value added by the publisher. U.S. federally funded research agencies should neither take publishers' products nor require publishers to provide their products for public access.

An excellent mechanism to ensure public access to materials that analyze and interpret research funded by the taxpayer is already partially implemented. By law, every federally funded research project is required to provide a detailed final report. Some science funding agencies make these reports freely available via the Web, but others do not. Making them all available would solve the access problem. Other agencies, such as NSF, are exploring alternate means of providing public access to researcher-supplied reports.

Agencies should seek productive and mutually beneficial projects and partnerships that ensure greater availability of both taxpayer-funded research directly from the government and peer-reviewed, value-added publisher content. For example, publishers are ready to partner with federal agencies to provide easy links between progress reports detailing research results, perhaps including lay summaries, and the peer-reviewed version of record, including complete access to the abstract or summary. Such projects would result in interoperability between funding agencies and publisher content, ensuring more timely and complete availability of scientific communication related to federally funded research, as well as better reporting on the results of taxpayer funding for 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?

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. As with any kind of content published by a nongovernmental entity at its own initiative, government-mandated access to books, proceedings, or other such materials is an expropriation of private property. The same model used for journal publishing, in which the federal agency publishes the detailed final report from a grant and the publisher provides valuable peer review, editing, and curation services should be followed for books publishing.

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

The federal government should not mandate any such embargo period because 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, publishers should determine the business models on which their publications operate, and this should include the time, if any, at which the final peer-reviewed manuscript or final published article is made publicly available. AGU and several other publishers permit limited distribution of articles by the author as a matter of policy, but that policy should be at the publisher's discretion or at the option of the author. Many publishers that derive revenue from subscriptions have options to allow authors to allow anyone to have unfettered access to the authors' article; at AGU this is a program called Author Choice. For a fee to compensate AGU for the valuable publishing services provided to the author, the author may make an article open, subject to copyright restrictions on the end user.

Peer-reviewed articles are not the direct result of the expenditure of taxpayer funds; on the contrary, they result from a significant publisher investment. The ability to recoup that investment enables innovation, allows infrastructure to be developed (including archives and metadata), and provides incentives to try new approaches.

There are, therefore, no "appropriate" embargo periods, and the research in different disciplines (and even subfields) has different life spans. For example, articles published in AGU's 12 journals have a long citation half-life of 8 years and lifetime usage that exceeds 20 years. Recent articles are important to any field of research, but at AGU and other publishers, articles retain important value many years into the future. Any embargo period is a dramatic shortening of the period of copyright protection afforded all publishers and is likely to significantly impact publishers' ability to add value and innovate.

We appreciate the opportunity to respond to this RFI and hope that AGU and other stakeholders will be involved in the continuing discussion of this topic.

Sincerely,

A handwritten signature in black ink that reads "Christine W. McEntee". The signature is fluid and cursive, with the first name "Christine" written in a larger, more prominent script than the last name "McEntee".

Christine W. McEntee
Executive Director/CEO

Subject: Comments regarding 'Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research'

Date: December 22, 2011 11:30:57 AM EST

I am the Editor in Chief of The Electricity Journal, a peer-reviewed scholarly publication. We sometimes receive papers from authors who have conducted research that was federally funded. I feel compelled to respond to your initiative and offer the following:

Any government attempt to establish peer review research reporting policies should be accomplished with full participation of all stakeholders to insure the integrity of the scientific research. Our goal is simply to have a fair system that allows access to research articles to advance knowledge. There is required a considerable investment to maintain the highest standards for peer-reviewed scientific publications by publishers. A sustainable mechanism is required to enable these systems to work without bias. Access policies must take these costs into account, and it is critically important that any new policies do not damage the publishing institutions on which the Federal Government and science depend.

Respectfully submitted,

Richard Cohen
Editor
The Electricity Journal



THE AMERICAN ASSOCIATION OF
IMMUNOLOGISTS

December 22, 2011

by email to publicaccess@ostp.gov

Re: FR Doc. 2011-28623

Dear Sir or Madam,

The American Association of Immunologists (AAI), the largest professional association of immunologists in the world, represents more than 7,400 basic and clinical immunologists. AAI appreciates having this opportunity to submit comments on the November 3, 2011, "Request for Information" (RFI) by the Office of Science and Technology Policy, Executive Office of the President, regarding "**Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research.**"

AAI is the owner and publisher of *The Journal of Immunology* (*The JI*), the largest and most highly cited journal in the field. (See <http://www.jimmunol.org/>). *The JI* was first published in 1916 and has been continuously published since then; thus it has been the repository of our profession's scientific advancements and achievements for nearly a century. The publication and stewardship of *The JI* is a core component of our organization's mission. Ensuring its future is as essential to our profession as preserving its past, and we take this responsibility seriously.

The JI publishes novel, peer-reviewed articles describing original research findings in all areas of experimental immunology, including innate and adaptive immunity, inflammation, host defense, clinical immunology, autoimmunity and more. The journal's Editorial Board is comprised of 184 editors: 10 Deputy Editors, 52 Section Editors, and 122 Associate Editors. Together with more than 3,600 other practicing scientists from all around the country, they donate enormous time and extraordinary expertise to serving as peer reviewers and/or editors of the nearly 4,000 manuscripts submitted annually for publication in *The JI*.¹ A trained in-house staff ensures that *The JI*, which published 1,572 articles in 2010, solicits high quality peer reviewers for each submitted manuscript; follows up with all reviewers to ensure that each manuscript has received a comprehensive review; works with vendors to ensure correct copy-editing of all articles and with authors on necessary changes; ensures that the text of the articles and all figures and tables are presented in print and online in a clear and accurate manner, and that any supplemental material is published accurately online; works with vendors to shepherd the compiled journal through the printing process and mail the journal to all print subscribers; and ensures that it is posted online – and all of this, twice each month. In addition, the Editor-in-Chief, Executive Editor, Publication Director and others ensure that any allegations of fraud, fabrication or plagiarism are promptly and fully addressed; protect the journal's copyright (often violated through premature posting of manuscripts on NIH's PubMed Central); and identify and stop the systematic downloading of copyrighted material from the journal's website for any unauthorized purpose.

¹ In 2010 alone, these scientists provided to *The JI* more than 11,000 reviews of newly submitted and/or revised manuscripts.

Publication of *The JI* is a time consuming, costly process,² to which AAI members and the immunology profession are devoted. To ensure *The JI*'s future, AAI must - and does - consider carefully the economics of publication: we know we must continue to publish and disseminate cutting edge research to our readers as affordably and quickly as we can. And we do this, year after year, and intend to do this in perpetuity.

Like many scholarly publishers,³ AAI seeks to provide prompt access to the latest scientific breakthroughs in our field and to provide high quality, reliable services - including peer review (as described above) and commentary - to our authors and readers. *The JI* includes special sections designed to further advance our readers' knowledge: "Cutting Edge",⁴ "Brief Reviews",⁵ "Pillars of Immunology",⁶ "***In This Issue***",⁷ "***Translating Immunology***",⁸ and "***Letters to the Editor***".⁹ ***To foster the rapid dissemination of its contents, The JI makes abstracts available online - at no cost to any reader - immediately upon publication, provides a "publish ahead of print" version ("Next in The JI")¹⁰ available to subscribers almost three weeks ahead of the print version, and provides free public access to the full text of all articles published in The JI twelve months after publication (before twelve months, the full text is under subscription control). In addition, "Immunocasts," podcasts of research highlights from each issue of The JI, are available to anyone without charge at the time the issue is published.***

AAI believes that our commitment to our members, authors, and readers; our scientific expertise; and our long publication experience enable us to both comment on, and assist with, federal efforts to expand public access to scholarly publications that report federally funded research. ***In our view, working in partnership with professional societies and other scholarly publishers offers the federal government the most cost-effective and efficient way of ensuring that private sector, scholarly journals survive, preserving their crucially important service of providing independent, expert peer review (accomplished at publishers' expense) of government-funded scientific research.***

AAI has invested considerable time and resources to expand public access, improve interoperability, and foster publishing innovations. Our investments have created digital platforms with the latest and continually evolving Web capabilities, thus providing researchers with faster and more robust delivery of

² Like any top tier journal, *The JI* is expensive to publish. In 2010, publication costs reached nearly \$5 million.

³ Scholarly publishers include professional associations/societies, commercial publishers and university presses.

⁴ Cutting Edge is the rapid publication section of *The JI*, presenting short reports describing significant advances in an area of immunology. Manuscripts in this section present scientifically sound and novel research in a clear and concise fashion, and contain conclusions of unusual interest to immunologists that are justified from the data presented. Chief criteria for acceptance are scientific novelty and quality, originality, clarity, and conciseness.

⁵ *The JI* publishes a small number of short, invited reviews ("Brief Reviews") on a regular basis. They cover a focused area on the advancing edge of immunology and provide a balanced view of current research that can be understood by researchers outside of that specialty.

⁶ In the first issue of each month, *The JI* reprints papers that have come to be regarded as classics in the field ("Pillars of Immunology"). An invited commentary accompanies each reprinted article.

⁷ A small number of papers regarded by reviewers and editors as the top 10% in their field are highlighted in the "***In This Issue***" section.

⁸ The "Translating Immunology" section documents discoveries from the field of immunology that have become effective treatments, drugs or diagnostic devices used in healthcare.

⁹ *The JI* publishes letters of general interest which comment on recent work published in *The JI*. The authors of the original work are invited to respond, and both the original letter and the authors' response are published together.

¹⁰ In "Next in *The JI*," *The JI* publishes the final, copyedited, author-approved version of the article online ~2½-3 weeks ahead of the print issue.

scholarly information. We have improved interoperability through new metadata standards, which provide for better information discovery and expanded use of research results. And we have launched a mobile website that is optimized for easy navigation and viewing on small screens to allow web-browsing on mobile devices.

Publishing *The JI*, and providing all of the associated services (including but not limited to providing independent, expert peer review of taxpayer funded research; editing; formatting; and ensuring the integrity of the research reported) is costly and time-consuming, and requires considerable expertise. It is essential, therefore, that any public access policies mandated by the government consider both the value of these activities and their immutable costs. Such policies must also protect publishers' intellectual property rights, to ensure both sustained investment by the private sector and the viability of the publishers' products; and to preserve the confidence of readers who rely on the integrity of the journals' published works.¹¹

Finally, before addressing the specific questions raised in the RFI, AAI wishes to express our continued willingness to work with the government on the issues raised in the RFI, and our appreciation for this opportunity to describe the value and importance of *The JI* to our members and to the field of immunology.

AAI does, however, want to express our clear opposition to government mandates which require private sector publishers to make their legally-owned property (i.e., journal manuscripts, published articles and associated data) available online on sites other than our own, or to comply with a government-determined embargo period. These mandates allow the government to take private property without owner authorization or compensation, and threaten the sustainability of our nation's premier peer-review publishing system.

Below are the responses of AAI to the questions raised 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?**

AAI is one of thousands of U.S. publishers in myriad disciplines which publish peer-reviewed publications and disseminate them internationally.¹² Subscribers to *The JI* are primarily scientists and the institutions which employ them. Other subscribers include libraries of hospitals, for-profit and non-profit businesses, and government agencies. Because the nature of what is reported in *The JI* is highly technical and discipline specific (as described on page 1, above: "novel, peer-reviewed articles describing original research findings in all areas of experimental immunology, including innate and adaptive immunity,

¹¹ The recently enacted America COMPETES Act established a public access policy for research funded by the National Science Foundation (NSF). This policy, as well as the policy currently in place at NIH, should be carefully analyzed to determine which (if either) best accomplishes the goal of ensuring the greatest possible public access without harming scholarly publishers or the scientific enterprise.

¹² The American publishing industry is valued at \$8 billion and employs 50,000 Americans. See "The STM Report: An overview of scientific and scholarly journal publishing," *International Association of Scientific, Technical & Medical Publishers*, September 2009, and an Association PowerPoint presentation noting that of the 110,000 people employed globally, 50,000 are employed in the U.S.

inflammation, host defense, clinical immunology, autoimmunity and more...”), it is unlikely that our “market” could grow beyond those who are already engaged in immunological or related research. The vast majority of these scientists already have access to the full content of *The JI*, either as a benefit of AAI membership or as an employee of an academic institution or independent research institute.

For most members of the general public, print copies of *The JI* are available for purchase, abstracts are available online immediately and at no charge, and the full journal content is available online at no charge after a twelve month embargo period. Also, during the twelve month embargo period, members of the public may purchase a PDF of any article online.¹³

Regarding this question’s implication that increased access could “maximize U.S. economic growth and improve the productivity of the American scientific enterprise,” AAI would point out that increased “free” access is unlikely to benefit the scientific enterprise in the U.S., where virtually every scientist who needs or wants access to *The JI* (and similar scholarly journals) already has it. Rather, increased “free” access is likely to benefit scientists in other nations,¹⁴ whether allies or enemies. In some instances, this will enhance international cooperation in the sciences, but it is not necessarily beneficial to the U.S. economy as even our friendly competitors will gladly take our research findings for free.¹⁵

Consequently, what federal agencies can do to help publishers grow new markets is to ***not*** mandate public access to peer-reviewed publications; foreign markets represent our best and probably only new market to tap. Yet there is no need for foreign scientists or institutions to subscribe to our publications if the U.S. government makes the peer-reviewed articles that they seek available at no charge.¹⁶ Neither publishers, nor the U.S. scientific enterprise, nor the U.S. taxpayer benefits from the “giving away” of our peer-reviewed publications.

To maximize U.S. economic growth and improve the productivity of the American scientific enterprise, federal agencies should support the vigorous enforcement of publishers’ intellectual property rights, including copyright, and assist publishers in preventing online piracy and unauthorized reuse.

¹³ Members of the public may purchase a copy of any article from *The JI* for \$10, or purchase access to the entire online archive for two weeks for \$40. See our longstanding policy at: <http://www.jimmunol.org/help/features/major-features/pay-for-access>

¹⁴ According to NIH’s PubMedCentral, two-thirds of its users are from other countries. Letter from Susan Cornell, J.D., NIH Freedom of Information Act (FOIA) Officer, to Allan Adler, Vice President for Legal and Government Affairs, Association of American Publishers (May 19, 2011).

¹⁵ House Energy and Commerce Committee Subcommittee on Health Chairman Joe Pitts, in a letter to NIH Director Francis Collins, M.D., Ph.D., expressed concern that “[th]e NIH Public Access Policy may undermine the competitiveness of the STM publishing industry.” <http://www.pspcentral.org/documents/LettertoNIH111011.pdf> In that letter, he sought additional information on the NIH Public Access Policy, PubMedCentral, and “its impact on the science, technology and medical publishing fields,” including any evidence that NIH has “to support [its] claim... that PMC is ‘accelerating scientific discovery in the biomedical sciences.’”

¹⁶ If public access is mandated and publishers have to replace lost subscription revenue, they will be forced to impose additional charges on authors in order to cover the cost of peer review and publishing.

Finally, the federal government should adopt policies which support the wide variety of publishing models; this will foster competition between - and innovation by - publishers, likely resulting in greater access to a larger audience.

(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 adopt mandates which take or compromise the private sector's intellectual property. Rather, the government should set the example of respecting the rights of copyright holders, and, if necessary, providing fair compensation. Scholarly publishers expect the federal government to protect and abide by publishers' constitutional rights: the federal government's unauthorized taking of property is surpassed only by its enabling of third parties to engage in piracy and the unauthorized dissemination of copyrighted works.

As was recognized by our founding fathers, who gave Congress 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,"¹⁷ copyright is a fundamental tool to foster innovation and creativity. Laws or policies which undermine this right are particularly troubling when they have the imprimatur of the federal government. The NIH Public Access Policy requires that "all investigators funded by the NIH submit or have submitted for them to the NLM'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." Although AAI appreciates that the NIH embargo period is consistent with the policy of *The JI*, and that the law recognizes the importance of respecting copyright, the mandate itself allows the government to take copyrighted works and make them available for its own purposes, without the permission of the copyright holder. Further, the policy duplicates what the private sector is already doing quite effectively; reroutes traffic away from journal websites; allows anyone in the world (not just U.S. taxpayers, who funded the research presented in the final published article) to "own" copyrighted materials without the permission of the copyright holder; and enables the posting of copyrighted materials on rogue sites. Although PMC could play a valuable role in facilitating publisher efforts to disseminate scholarly publications (for example, through linking to journals on publisher websites),¹⁸ the policy as currently implemented actually undermines U.S. private sector publishing, contributing to a loss of subscribers and significant revenue.

¹⁷ U.S. Constitution, [Article I Section 8](#)

¹⁸ In 2005, AAI and 56 other not-for-profit scientific publishers offered to NIH a "Linking Proposal:" to provide seamless links on PubMed Central to the journals' websites, enabling readers to access the full text of any article funded by NIH (and in many instances, the full text of all articles published in the journal, irrespective of funding source). This proposal would provide the public with free access to all published articles funded by the NIH; provide access to the final, copy-edited articles; address publishers' copyright concerns; satisfy current law; and comply with copyright law by ensuring that an article cannot be posted before the journals' embargo period is over. It would also be cost effective, since the NIH would not have to maintain a repository, educate grantees about compliance and copyright, or monitor for compliance. Despite its many advantages, and despite subsequent offers from publishers to consider ways to satisfy NIH's need for a repository of all NIH funded works (i.e. to help NIH populate a "dark archive" for internal NIH use only), NIH rejected the publishers' Proposal.

(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?

AAI is deeply concerned about the increasingly aggressive efforts of federal agencies to seek control and custody of privately owned, scholarly, peer-reviewed publications. Although the concept is superficially appealing - that a government agency would maintain a database of all the research which it has fully or partially funded - AAI has deep concerns about the ultimate result of this effort, which is to control not only the research findings, but also the final product reporting that research (which product was made possible through essential services provided - and paid for - by the private sector).

As indicated above, the publication of *The JI* is a costly endeavor.¹⁹ AAI and other professional societies which publish scholarly journals support their journal enterprise as an essential part of their mission. We know the cost, and wonder why, in this time of serious fiscal constraint, the federal government would create, and seek to expand, the use of central federal repositories, which are duplicative and inferior to publisher repositories (which include the full content of what is published, not just the content which is supported by federal funds). The notion that a government repository, which depends on an uncertain federal appropriation process, is more “reliable” or “permanent” than one managed by the private sector (in the case of AAI, for almost 100 years) is no longer valid.

AAI is also concerned that a centralized approach will (and has) increased administrative burden for our journal and our authors. Manuscripts accepted for publication in *The JI* are now submitted by *The JI* to PMC on behalf of the author, which adds staffing time and its associated cost. In addition, *The JI* staff must monitor PMC to ensure that articles are not posted prior to the expiration of the embargo period.

Despite these concerns, AAI recognizes that certain databases, including PubMed, Genbank, GEO, and the Protein Database, are useful tools for researchers and should continue to be supported. In addition, the NLM XML DTD has proven to be a valuable common tagging format for exchange of journal content. The federal government can continue to support projects such as these and new ones as the need arises, e.g., MIATA (<http://www.miataproject.org/>) is a new project that the government could consider supporting.

¹⁹ See Footnote 2, above.

As a professional association, AAI is in close touch with its members and therefore uniquely positioned to understand researchers' needs in relation to delivery of scholarly information. Although we are a not-for-profit organization, we must - like our for-profit competitors – have a viable business model if we are to remain operational. Part of our business model is to adopt innovative approaches to delivering information and making it more accessible to its users. Some innovations include:

- *The use of DOIs* (Digital Object Identifiers through CrossRef)²⁰: *The JI* enables interlinking of published articles via DOIs, a tool used by some scholarly publishers to “uniquely” identify content.
- *Linking opportunities*: *The JI* links from articles to relevant databases at the National Center for Biot Compliance with COUNTER requirements: All content in *The JI* is compliant with the latest COUNTER requirements. COUNTER (Counting Online Usage of Networked Electronic Resources) resulted from a collaboration among publishers, librarians and database providers, who created standards for measuring online usage data. COUNTER provides reliable usage data to librarians and publishers, allowing each to better understand how the online content is being used.

AAI will continue to strive to provide innovative products and services, unless our financial livelihood is undermined by harmful government policies.

(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?

Although efforts by publishers vary, all are undoubtedly concerned about, and committed to, archiving their product and ensuring its long-term accessibility and usefulness. The stewardship of archiving is secure when the publisher is secure. In addition, several entities already exist to ensure long-term availability of publications should a publisher go out of business (e.g., 1) CLOCKS <http://www.clockss.org/clockss/Home> - a not-for-profit joint venture started by libraries and publishers committed to ensuring long-term access to scholarly publications in digital format, and 2) Portico <http://www.portico.org/digital-preservation/> - a digital preservation service provided by ITHAKA, a not-for-profit organization with a mission to help the academic community use digital technologies to preserve the scholarly record. Articles are then all freely available to the public.

(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?

²⁰ CrossRef is a not-for-profit group founded by publishers in 2000 which maintains 50 million items (including articles, supplemental data, etc.)(see (<http://www.crossref.org/01company/pr/news111511.html>)). Through CrossRef, over 3500 publishers and societies assign Digital Object Identifiers (DOIs) to published content items, enabling seamless navigation of the research literature (see <http://www.crossref.org/01company/06publishers.html>), CrossRef enables users to link to the full text of a referenced article from the bibliography of another article.

AAI and many scholarly publishers have devoted extensive resources to promoting wide dissemination, and facilitating the use, of their publications. To encourage discoverability and interoperability, *The JI* participates in CrossRef, which provides users with seamless navigation of the research literature. (See *Footnote 20*) *The JI* is also crawled by major search engines such as Yahoo!, Google and Google Scholar, which further improves discoverability.

AAI makes the contents of *The JI* available online through HighWire Press, a subsidiary of Stanford University Libraries, which hosts the online platform for *The JI*. As mentioned in response to Question 3 (above), through HighWire, AAI has adopted COUNTER standards, a uniform way of counting downloads to enable librarians to analyze usage data. In addition, microdata markup tags,²¹ which are key to the semantic web, have recently been added to all journal article pages, improving the results of search engines which crawl the scholarly content.

Federal agencies can best foster private innovation in publishing by avoiding mandates, assigning liaisons within federal agencies to work with publishers on removing administrative obstacles, and continuing support for items that need a centralized system, such as those mentioned in Answer 3 (above).

(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 should facilitate private sector efforts to maximize public access to the federally-funded research in which taxpayers invest in a cost-effective and legal manner. It is important that the government remember that its investment is in research and not in publishing.²² Publishers invest in peer-review, editing, copy-editing, web-hosting, and numerous other services which transform the research results into a publishable article. Although federal agencies do not have the right, and should not have required publishers to provide public access, most private sector publishers already do make their articles available at no cost to the public, with varying embargo periods, because it is in their interest for their journals to be widely read.

Federal agencies should work collaboratively with the private sector to provide increased public access to both the findings of taxpayer-funded research and the final, peer-reviewed, value-added content available in private scholarly journals. Through such efforts, the government could also enhance interoperability between funding agencies and the content of private sector journals, whether or not that content is funded by a federal agency.

(7) Besides scholarly journal articles, should other types of peer-reviewed publications resulting from federally funded research, such as book

²¹ This addresses the June 2011 guidelines issued by microdata www.schema.org, a joint initiative of Google, Bing and Yahoo!

²² NIH grants provide grantee-authors with some funding for publication charges. But this funding does not cover the publisher's costs.

chapters and conference proceedings, be covered by these public access policies?

No. Private sector publishers also invest substantially in these other types of peer-reviewed publications. The same arguments which apply to scholarly articles apply to these publications, and a federal mandate requiring public access to these publications is an expropriation of private property.

(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, AAI believes that the copyright holder (in most cases, the publisher) has the right to determine the embargo period for a peer-reviewed manuscript or article. Publishers must be able to select the business model and embargo period which best suits the needs and demands of their field, with the goal of fostering accessibility without jeopardizing their journal's viability.²³

AAI has determined that for *The JI*, a twelve month embargo period is appropriate as it allows us to recover a portion of our costs by retaining the majority of our subscribers.²⁴ If the federal government were to require a shorter embargo period, it could erode our subscription base and potentially bankrupt *The JI*.

The cost of publishing *The JI*, the rates we charge our subscribers, and the costs to authors to publish in *The JI* are dynamic, varying from year to year and reflecting changes in the marketplace. Prices for supplies, such as paper, change considerably over time, just as new approaches, services, and technology²⁵ affect costs. Any excess revenue is used for product improvement, including new technologies.

AAI appreciates having this opportunity to comment, and thanks you in advance for considering our views.

Sincerely,

M. Michele Hogan, Ph.D.
Executive Director
The American Association of Immunologists
Executive Editor, *The Journal of Immunology*

²³ There is not one embargo period that will work for all disciplines, as each journal and field has different life spans. For example, articles published in the *The JI* have a cited half-life of 7 years, whereas other journals may have much shorter or longer cited half-lives.

²⁴ *The JI* has lost a considerable number of subscribers in recent years, which we believe is due to public access policies that undermine the need for a subscription.

²⁵ Since the 1990's, *The JI* has invested heavily in providing content in the digital arena. It was one of the first journals to put content online, beginning with its abstracts and *Cutting Edge* section, followed by full content in January 1998. Over the following decade, all content back to the first issue of *The JI* in 1916 was digitized and made available online. Working with others in the industry, *The JI* has dramatically improved functionality and efficiency for readers, who can now perform complex searches, immediately retrieve and either save to their computers or print full text articles, link instantly to other cited articles, export text and references to other databases and programs, and receive e-mail alerts or RSS feeds when new journal issues are released. In addition to its large screen website, *The JI* also has a website optimized for reading on mobile devices, allowing readers to use their time with even greater efficiency.



Phycological Society Of America

c/o Susan H. Brawley, President
5735 Hitchner Hall
School of Marine Sciences, University of Maine
Orono, ME 04469, USA
Telephone: 207-581-2973 Fax: 207-581-2801 brawley@maine.edu

December 22, 2011

Office of Science and Technology Policy

To Whom It May Concern:

The Phycological Society of America (PSA) appreciates the opportunity to respond to OSTP's request for information from scientific societies about questions concerning public access to federally-funded research, as described in the Federal Register solicitation of 4 November 2011 per Section 103 (b)(6) of the America COMPETES Reauthorization Act of 2010 (ACRA). "Phycology" refers to the study of algae. The PSA is a non-profit scientific society that is incorporated in the State of Maryland and was founded in 1946 to advance research and education in all aspects of algal science. Our membership is comprised of about 1,000 scientists, researchers, educators, and graduate students who do research in universities, industry, state and federal government, and NGOs. About two-thirds of them work or study in the US. Our members study a remarkable range of important topics, from key aspects of global carbon cycles to important health issues. For example, one of our members found that the mosquito-transmitted malarial parasite contains an essential, relic chloroplast (the part of a cell that carries out photosynthesis), opening new targets for treatment of the disease because human cells do not contain chloroplasts.

The PSA's *Journal of Phycology* is entering its 48th year (2012) of publication of basic and applied research on algae and their roles in the ecosystem. For example, the *Journal of Phycology* recently published:

- Unequivocal demonstration that the skin-eating alga *Pfiesteria piscicida*, which kills marine fishes, is a typical (not unusual) dinoflagellate, making its management more straightforward,
- Discovery of an ancient lineage of green algae in barely lit areas of the ocean,
- The first genetic modification (transformation) of the chloroplast of diatoms, which are among the most abundant algae of freshwaters and oceans and account for 25% of photosynthesis on Earth, and
- Many papers dealing with lipid biosynthesis pathways in algae, a field of research critically important for algal biofuels' development (Note: petroleum is composed of ancient algae).

PSA owns copyright to current and full back issues of the print and electronic journal, which is published 6 times/year. PSA controls all editorial decisions and content of our journal, which has been published since 1999 in association with Blackwell and now Wiley-Blackwell Publishers, the publisher of the journals of about 258 other US scientific and scholarly societies. Our partnership with Wiley-Blackwell (W-B) was particularly valuable in terms of early establishment of the electronic version of the *Journal of Phycology* (begun in 1999), which provides easier and wider access to research published in our *Journal*, and enables archive of digital data and detailed methodological information

associated with papers as supplementary electronic material. We are in a profit-sharing contract with W-B in which each party makes about \$100,000/year after expenses. Our “profit” is immediately invested in programs to support other society activities that build US science infrastructure (see below).

We wish to offer a general framework that is important as background to our responses to some of the questions posed by OSTP.

In PSA’s judgment, OSTP’s questions confuse two different products: the result of federally funded scientific research and the peer-reviewed publication that results from funded research. Despite the low-success rates (high competition) now typical for federally-sponsored grant competitions, the research papers that result are still vastly improved before publication by peer review from scientists who are expert in that field. This means that highly rated journals (e.g., the *Journal of Phycology*) that offer expert review of research in a particular field ---and their sponsoring scientific societies---should be valued by the federal government. Peer review sometimes finds that the investigator needs to use different or additional statistical techniques to analyze the data; sometimes the investigator has failed to compare his/her data to important research previously published in that area that influences interpretation of the new results; sometimes the investigator needs to do an additional confirmatory experiment in order for the results to be accepted by the scientific community as strong evidence in support or refutation of a hypothesis. Only peer review can force this improvement in the paper, and peer review by members of scientific societies is a critical contribution to the value of the research.

Scientists submit an average of 310 manuscripts each year to The *Journal of Phycology*. About half of these are ultimately published, but virtually no manuscript is accepted when it is first submitted. Instead, minor to major revisions are suggested by peer reviewers, and re-review of the revised work is required before a decision on acceptance is made. The *Journal of Phycology*’s review structure includes scientists serving as Co-Editors (4 CEs), Associate Editors (16 AEs), the Editorial Board (24 scientists), and several hundred other peer scientists who respond to requests from CEs/AEs to review manuscripts each year. Only the Managing Co-Editor receives stipend support (about 3 months’ salary) and only the Assistant Editor (editorial assistant) receives a full salary; total editorial office costs are \$106,000/year. These before-profit, publication expenses are paid by the partnership between PSA and W-B, based on income from subscriptions and downloads of articles by libraries and individuals (authors are rarely charged for publication in the *Journal*).

We estimate that the manuscripts that are reviewed each year at the *Journal of Phycology* require 5,580 hours of peer reviewer time and 2,640 hours of Co-Editor/Associate Editor time. Our conservative estimate of the value of this donated time is \$252,550.

The donation of time by scientists to society journals is an honored tradition to assure excellence in the field and because it is understood that scientific societies use the modest income from their journals to build scientific infrastructure. In 2011, PSA invested about \$30,000 in support of graduate student research and education; about \$22,000 in support of symposia (e.g., environmental genomics, algae and human health) at our annual meeting; about \$6,000 in public outreach activities (e.g., to K-16 teachers) through our Committees and website (www.PSAalgae.org); about \$3,000 to support the free, public algal database AlgaeBase (www.algaebase.org); \$5,000 for transition expenses in our Editorial Office; \$11,000 in 8 awards for meritorious research and publication prizes to senior researchers and students; and about \$10,000 in administrative expenses (e.g., accounting fees, travel by Society officers to a mid-year business meeting). PSA also “seeds” and co-publishes key

books in our field, including most recently *Algal Culturing Techniques* (2008), extensively used by researchers involved in experimental and commercial-scale grow-out of algae (e.g., for biofuels, food products, cosmetics) and *A Color Atlas of Photosynthetic Euglenoids* (2010), making it possible to identify algae that include biomonitors for eutrophication.

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?

As stated above, the PSA (and similar scientific societies) holds copyright of the peer-reviewed articles in its journal. Thus, we do not think the federal government has a role to play in developing and expanding markets for scientific journals. With respect to access to publications needed to support economic growth, the partnership between PSA and Wiley-Blackwell Publishers already results in articles in the *Journal of Phycology* (and similar journals) being readily available upon publication:

- 1) Abstracts to articles are freely available in the Wiley Online Library (and retrievable freely by searches using search engines such as Google); interested parties who are not associated with a subscribing university library can buy single copies of any article that appears important to them for \$30/each.
- 2) Individuals exploring commercial developments based upon algal research can become members of PSA and enjoy full *Journal* access (1965-2012) for \$80/year (electronic subscription rate).
- 3) Most US researchers will have free access to the *Journal* through the consortial journals' package sold to university/college libraries by W-B.

We have found our partners at Wiley-Blackwell to be innovative and aggressive in marketing of the *Journal of Phycology* to increase its availability while using W-B funding to scan issues printed as paper only (1965-1998) in order that the *Journal* is available to interested parties in electronic form. The PSA retains full copyright and ownership of the pre-1999 *Journal*, including the new electronic files. W-B also continually develops, and makes available to our editors, software that helps us manage the peer review system more efficiently and cost-effectively.

We do think there is a role for the federal government in “eternal archive” research. Paper is still the only truly archival medium. Various schemes of networked comparisons of electronic publications are being used, but many societies are considering electronic-only publication in the near future to reduce production costs and, therefore, subscription costs to both libraries and individual members. It does not seem to us that there is a really secure, permanent archive for electronic publications. The government might sponsor research to pioneer such developments.

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?

The best way to protect these intellectual property interests is to respect the copyright laws of the US. The key difference between “federally funded scientific research” and “peer-reviewed scholarly publications” must be understood by OSTP and Congress. Under no circumstances should the government jeopardize the value that peer review adds to the validity of scientific research, and the value and cost of this peer review should not be underestimated or ignored.

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?

Publishers such as W-B have developed Digital Object Identifier (DOI) tags for each journal article over the last decade, and CrossRef already maintains more than 50 million DOIs. CrossRef is a not-for-profit group established by publishers including Wiley in 2002, and this private enterprise-supported development provides a great degree of interoperability. CrossRef allows a reader to link from a reference in the bibliography of an article in one journal to the full text of referenced articles in other journals, including those published by other publishers.

We can not imagine the US government purporting to have the right or the capacity to take custody of all published (copyrighted) content.

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?

In discussing this question with our partners at Wiley-Blackwell, they told us about an important possibility in this regard, which would make it possible for funding agencies to be able to track all publications from each funded grant—and allow government funders to make a catalogue of this citation information for each grant available on agency websites. This would require that each journal article has metadata indicating the agency, agency division etc. that funded the paper. We understand that there is a good chance this feature may launch in 2012, because it has been endorsed by CrossRef and major science publishing trade associations with technical details now to be discussed between CrossRef, publishers, and agency representatives. This development would support accessibility and interoperability, and we urge the federal government to pursue its launch. It has the advantage of giving the public access to the various places in which peer-reviewed publications of the raw research results of each grant have been published.

The government could also develop a system that would link free research reports (funded grant abstracts or required final reports) from investigators on agency websites to the free abstracts of the final Version of Record of publications related to a particular grant.

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 publically available and can be easily found and linked to Federal science funding?

Please see answer to Q4.

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?

As pointed out above, tax payers have not paid for peer review, which adds strong value and validity to published research. Peer review has a cost. If the government wants to support the peer review system for federally-funded research, it could add suitable monies to each research grant to do so; most journals, including PSA's, have a mechanism to allow immediate free access to the entire published article if the authors pay a surcharge to cover costs of publication and (donated) peer review. In the case of the *Journal of Phycology*, authors' payments would allow us to reduce or eliminate subscriber fees if most authors paid for immediate open access.

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?

No. These materials also benefit from peer review (book chapters or collections of full research papers from conferences) and are copyrighted.

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?

In terms of the general issue of "shelf life" of research published in the *Journal of Phycology*, the ISI calculated "half-life" is greater than 10 years (and *J. Phycol.* has a 5-year Impact Factor of 2.86). This is an indication that our copyrighted product retains value for many years, due to the importance of the articles and the underlying excellence of peer review that helped make them important. This is one of the reasons that the PSA's partnership with W-B has been successful, because

W-B's early investment in electronic access for the pre-1999 *Journal of Phycology* opened wider access to the older literature published in the *Journal*.

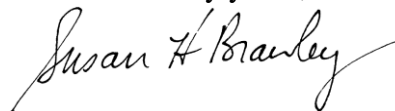
We will begin making 15% of our articles (the "Editors' Choices") open access in 2012; in part, this is an experiment to evaluate the effect on library subscriptions. Certainly, in order for the entire *Journal* to become open access, authors would have to pay the true cost of publication to replace income now derived from library and individual subscriptions and paid downloads of articles. The *Journal of Phycology* is a highly respected steward of much of the best research in algal science, but we are concerned that there would be substantial loss of library subscriptions if we made the full content of articles free soon after publication. What is "soon"? We are unsure, and must be cautious to avoid undermining our *Journal*, the focus and network of researchers it brings to research on algae, and the small income it provides to support PSA activities that support algal science.

It is important for researchers to continue to choose where they submit their research based on where they think it will have the best audience and where their manuscript will receive the strongest improvement in peer review, regardless of the ultimate decision on its acceptance or rejection for publication. This competitive system strengthens the scientific enterprise.

In summary, peer review provides a critical value-added benefit to the federally funded research on algae that is published in our journal, and *Journal of Phycology* content is therefore protected by copyright. We also believe, as documented above, that our content is widely available and that the free online abstracts allow interested parties to obtain full content important to economic development at very low cost.

These responses were reviewed by the PSA's Executive Committee. We would be happy to respond to further questions.

Sincerely yours,

A handwritten signature in cursive script that reads "Susan H. Brawley". The signature is written in black ink and is positioned above the printed name and title.

Susan H. Brawley
President



Springer

the language of science

Springer Science+Business Media, LLC
233 Spring Street
New York, NY 10013-1578 | USA
tel +1 212 / 460 1500
fax +1 212 / 460 1575
www.springer.com

William F Curtis
President
tel +1 212 / 460 1569
william.curtis@springer.com

Office of Science and Technology Policy
To: publicaccess@ostp.gov

22 December 2011

Re: FR Doc. 2011-28623

Dear Sir or Madam,

On behalf of Springer Science+Business Media, I am writing to respond to the OSTP Request for Information (RFI) published on 3 November 2011, regarding *Public access to peer-reviewed scholarly publications resulting from federally funded research*.

Springer Science+Business Media (Springer) is one of the world's leading suppliers of scientific and specialist literature. The company is headquartered in Berlin, with major editorial offices in New York (US), Heidelberg (Germany) and Dordrecht (Netherlands). It has operations in about 20 countries in Europe, the USA, and Asia, and some 5,500 employees.

Springer is the second-largest publisher of journals in the science, technology, and medicine (STM) sector, the largest publisher of STM books and the largest open access publisher worldwide. The group publishes some 2,000 journals and more than 7,000 new books a year, as well as the largest STM eBook Collection worldwide.

Springer appreciates the call for comments set out in the RFI and welcomes the opportunity to submit its position on some of the issues addressed. In the following, we have grouped our comments with reference to the specific questions set out in the RFI, however we do not believe it is helpful or necessary for us to respond to every question in detail.

Response to RFI questions

Springer is the largest open access publisher worldwide (with BioMed Central, SpringerOpen and Springer Open Choice). Springer was the first publisher to offer the open access option at article level for its

subscription journals in 2004, and ever since, we have continued to be both receptive and proactive in open access developments. We offer a variety of options to researchers and funding organizations to publish open access, including open access at article level and a rapidly growing number of full open access journals. In 2010, some 12% of articles in our journals were published open access.

While Springer fully supports and invests in open access as a business model, it should not be forgotten that the cost of formal scholarly publication needs to be recovered somewhere in the funding process. Publication in peer-reviewed journals has associated costs, regardless of the business model. Though taxpayers fund research through the US federal agencies, they do not fund the investments that produce the peer-reviewed scholarly publications that result from that 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?

With regard to public access to federally funded research, Springer supports specific steps that can be taken which respect the copyright of both publishers and scientists in peer-reviewed scholarly publications.

The first option is to encourage federally funded researchers to make their research progress reports publicly available. We recognize that researchers, when they produce scholarly articles themselves, need access to the peer-reviewed and published articles resulting from research. However, to achieve the goal of general public access to federally funded research results, the research reports before publication are useful.

The second option is to provide, as part of a research grant, funds for the article-processing charge associated with “gold” open access publication in a peer-reviewed journal. Springer supports new business models as long as they are sustainable. However, all publications must be funded, whether they are open access or subscription-based. Gold open access publishing generates funding by shifting payment from the reader-side to the author-side. As a consequence, publishers do not need the author to transfer copyright in this model.

“Green” open access policies, on the other hand, potentially undermine intellectual property rights. Green open access archiving does not cover the costs associated with formal publication, and poses risks both in terms of versioning and the sustainability of scholarly communications. Springer believes that systematic green open access self-archiving is a danger and should not be adopted as a policy.

(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?

Preserving the majority of content in one central archive does not necessarily ensure long-term availability or advantages for interoperability. In fact, a key factor in one of the prominent current archiving initiatives, CLOCKSS, is the idea that the distribution of content across multiple sources is an asset: In case one of the sources fails to deliver content, another will be able to provide back-up. While in the print era preservation was ensured by large library collections, publishers of digital journal and book content are actively engaged in preservation efforts. Springer has multiple arrangements with national libraries (e.g. Koninklijke Bibliotheek) and preservation initiatives such as CLOCKSS and Portico.

Search functions and interoperability are also areas on which STM publishers have worked intensively and successfully, have invested both as individual companies and in cooperation as an industry, e.g. through CrossRef. Springer does not see any reason why these issues should be managed centrally by a Federal agency.

(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?

Public access does not mean that this access must be free of cost. While taxpayers fund research through the US Federal agencies, they do not fund the processes necessary to produce peer-reviewed publications resulting from that research.

If Federal agencies decide that the peer-reviewed published results of the research they fund should be made available free of cost to the reader, the best way to achieve this is by making funds available to cover the costs of a sustainable gold open access publishing model. Gold open access publishing, unlike green open access self-archiving, does not pose risks in terms of versioning and the sustainability of scholarly communications. It also minimizes the burden on scientists as authors, as they are not required to perform tasks such as producing an "author version" of their publication and posting this in 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?

In the usage of digital content, Springer observes that digital search to some extent levels the distinctions between different types of content. However, in producing the content, the difference between publications deriving from the direct initiative of a federally funded researcher and publications commissioned and conceptualized by a scholarly publisher is still very distinct. This distinction needs to be very carefully observed in any public access policy.

(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 think systematic self-archiving by itself is sustainable and therefore do not support self-archiving nor embargo periods. However, if embargo periods were necessary, these would have to vary for different subject areas. For many journals, e.g. in mathematics or most social sciences, even a 2 or 3 year embargo period would be too short. Clearly, there is no one-size-fits-all appropriate embargo period.

Yours sincerely,



William F Curtis, Ph.D.
President

Springer Science+Business Media, LLC

From: Pierre Bierre, CEO
Spatial Thoughtware, Inc.

Dec 22, 2011

COMMENT 1

We are a small, innovative startup developing Algorithmic Geometry education for high school 11-12th graders. This is a best-in-world (in the 9-12 educational sector), cutting edge modernization of spatial problem-solving. Algorithmic Geometry is interdisciplinary Math-Computer Science, teaching the practice of mathematics as done in the spatial applications software industry. I'm attaching a recent NSF Proposal Summary.

The point of view we represent is probably the extreme in core mathematics curriculum innovation - what we call Math Education for the 21st Century.

The signal strategic issue: Our goal is to give U.S./Canadian students a quantum leap forward in their Math problem-solving knowhow (the "standard" math curriculum having fallen decades behind high-tech industry in its portrayal of how mathematics is actually practiced.). That is, the goal of the project is NATIONAL competitiveness. WE ARE NOT DOING THIS PROJECT TO DEVELOP STUDENTS IN INDIA, CHINA, RUSSIA, ISRAEL, TAIWAN, KOREA, IRAN, etc.

What does it mean to "Publish" results? How to maintain Comparative Advantage for U.S. students?

How can high-impact, cutting edge breakthroughs in Education be managed so as to give U.S. students comparative advantage? This question boils down to: How can we manage the dissemination of cutting edge breakthroughs so that they have time to benefit U.S. students before getting out to the rest of the world?

Clearly, the notion on "publishing" needs to be challenged as too non-selective a means of dissemination. Anything that is published is globally disseminated in this era where nations like China, India, Iran are constantly scouring the Internet to pick up innovations for free. U.S. taxpayers have every right to insist that the benefits of educational research be enjoyed preferentially by U.S. students. Why else would they be spending their tax \$ on R&D? To help kids India and China get ahead, and sell their knowhow underpricing American workers? You may want to do a poll or focus group with average taxpayers to get a clear justification for taking the complicated steps needed to design a national dissemination system.

How would a national dissemination system work?

The US needs a National Research Intranet (similar to how a large multinational manages its proprietary IP). At minimum, this is a password-controlled system that uses VPN communications, and has safeguards against theft of laptops, mobile devices and desktops. It needs a highly-skilled security team involved in its design and operation. There needs to be a legal framework (Congress needs to get involved) defining National IP Assets. These are IP assets owned both privately by US companies, jointly with the USG, or USG exclusive. There should be a National IP Director who reports to the President. Treaties and Trade Agreements must be explicitly approved by the National IP Director during their negotiation, to maximize long-term US Competitiveness.

Time spans. The US would keep the wraps on its most cutting edge scientific, engineering, and educational innovations for a reasonable period of time during which competitiveness was established and enjoyed in the US. This could be 5 years for some projects, 10 years for others. The time period should depend on how much time it takes to fully implement the innovation plus several years of exclusive advantage. Ideally, around the time just before the IP

is to become obsoleted by something much more powerful, that's when it should be released globally (or "published").

Graduate Student access to National IP. Grad Student education is one of the "leaky" faucets whereby U.S. taxpayers are subsidizing foreign competitor workforces. The rules need to be tightened up so that National IP Assets are safeguarded in our Universities. Many academics (secure in their employment) tend to be "globalists", insensitive to the economic challenges experienced daily by their fellow Americans who have to compete in a lowest-cost global labor market. The ultimate insult is to use these taxpayers money indiscriminantly educating grad students who will take the benefits home, from where they can undercut U.S. labor rates. The USG must separate itself from this unhelpful attitude on the part of academics.

The more transformative and powerful the breakthrough, the longer it should be kept under wraps. Much of the incremental (paradigm-extending) research could be published earlier. The challenge for the USG is, can USG scientists determine on their own what constitutes the more transformative projects vs. those of lesser impact? Usually, only the proponents of radical innovations can make this recognition early on. A "crowd" review within the Intranet is likely the best way to be deciding national R&D investment priorities in the future. The USG bureaucrats have to first learn how to engage and trust the broad scientific community in making wise choices. Clearly, a secure National R&D Intranet would be needed to begin harnessing crowd-power.

Summing up, the essential policy question is, how can the fruits of taxpayer-funded R&D maximally benefit the U.S. workforce (preferentially over global competitor nations who are actively, illegally stealing IP from the U.S.? A national R&D Intranet, modeled upon the best corporate Intranet designs, is the solution to this problem, combined with some "period of exclusive advantage" as National IP, after which innovations would be published to the world.

If you would like my help developing this further, just ask.

Pierre Bierre
Director, Algorithmic Geometry Project
CEO, Spatial Thoughtware
Math Education for the 21st Century

Subject: Public Access RFI response

Date: December 22, 2011 1:22:12 PM EST

December 22, 2011

To Whom it May Concern:

Comment 4: As President of Autism Speaks, the world's largest autism science and advocacy organization, I am writing to voice my support for the Public Access of full-text scientific journal articles that report research findings without cost or institutional affiliation. We are fully committed to removing obstacles to first-hand access for the broad population including clinicians, families, educators, researchers and businesses. Among other dissemination requirements, Autism Speaks has stipulated that any research findings resulting from our support must be freely accessible in a public database within a year of publication. Autism Speaks is the first non-profit health research and advocacy organization to require public access, even if the article might be co-funded by a federal agency such as the NIH. In this public/private effort, we greatly appreciate the assistance provided by the National Library of Medicine in establishing an Autism Speaks submission identifier that tracks research publications partially or fully supported by our research funding programs. To indicate the strength of this public/private endeavor, Autism Speaks has become a model for the Health Research Alliance and its 45+ other health non-profit organizations in establishing and disseminating Public Access policies and procedures for their grantees.

Comment 6: Autism Speaks applauds the efforts of the NIH in establishing Public Access requirements and enlisting the cooperation of publishers, researchers and research institutions. It is important that parallel policies be established by other federal funders as well. Expanding the policy across agencies would maximize the benefit to a broad base of stakeholders including clinicians, families, educators, researchers and businesses.

Sincerely,

Mark Roithmayr
President
Autism Speaks

22 December 2011

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: H. Frederick Dylla, Executive Director and CEO, American Institute of Physics
Tel. +1 301-209-3131; Dylla@aip.org

Electronically submitted to: publicaccess@ostp.gov

The American Institute of Physics (AIP) 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 AIP

The American Institute of Physics is a 501(c)(3) not-for-profit membership corporation created in 1931 for the purpose of “the advancement and diffusion of knowledge of the science of physics and its applications to human welfare.” AIP is an organization of 10 physical sciences societies representing more than 135,000 scientists, engineers, and educators and is one of the largest publishers of scientific information in physics, with activities extending well beyond publishing. AIP delivers valuable resources and expertise in education and student services, science communication, government relations, career services for science and engineering professionals, statistical research, industrial outreach, and the history of physics and other sciences.

As a publisher, AIP plays a central role in the process by which scientific research is developed, communicated, disseminated, and ultimately accepted by the scientific community. AIP publications include 15 journals (three of which are published in partnership with other societies), magazines, including its flagship publication *Physics Today*, and the *AIP Conference Proceedings* series. In addition to its own publication, AIP provides publishing services and expertise to five of its ten Member Societies. To accomplish this, AIP invests millions of dollars annually on peer review, editorial management, production, printing, shipping, distributing, and hosting its archival journals on a fully digital, highly reliable online platform, making the content available at all times to customers around the world in more than 70 countries.

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 at our Melville, New York, publishing center, in addition to more than 340 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 the substantial nature of the research findings. Managing peer review for approximately 30,000 articles submitted to AIP journals every year is a complex undertaking. It requires a large amount of sophisticated electronic resources, associated support personnel, a staff of professional editors—nearly all PhD physicists—and help from tens of thousands of referees. Each year AIP makes such necessary investments to fulfill its public nonprofit mission, generating an intellectual return through the dissemination of scientific research.

Introduction

AIP's highest goal is to achieve the widest possible dissemination of the research results it publishes, including any pertinent associated data and context information. Enabled by Internet technologies, AIP disseminates more information, more widely and more affordably, than ever before in its history, reaching more authors, subscribers, and users than ever before. This accomplishment requires heavy investments in technology and infrastructure (such as an online platform) and business-model innovation to deliver the option of free or low-cost access: open access, pay-per-view, or article rental, recognizing that the value of the final published article needs to be paid for to remain sustainable.

AIP believes that it would be in the best interest of the United States and its government, as well as in the best interest of all other stakeholders, to strike a balance between public access and sustenance of the scholarly publishing industry because of the impact and value it brings to the progress of science and its contributions to American society and economy. Such a balance can be achieved based on shared principles such as the importance of peer review, the recognition of economic realities through adaptable and viable publishing business models, the need to ensure secure archiving and preservation of scholarly information, 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 public access policies—an approach that engages all concerned parties, including federal agencies, scientists, university administrators, librarians, publishers, and the public.

Consistent with the recognition of economic realities, it is AIP's position that government agencies should develop their public access policies through voluntary collaborations with nongovernmental stakeholders, including researchers and publishers. Any policies should be guided by the need to foster interoperability of information across multiple databases and platforms. Agencies' efforts then could be directed toward facilitating cyberinfrastructure and collaboration programs with and between agencies and the 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. More detail on this is provided later in the document. AIP believes that any scholarly publication access policy needs to be flexible to accommodate agency-specific needs and have the capacity to evolve in response to the rapidly changing nature of scholarly publishing.

AIP 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 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 United States. New publishers, journals, and business models either evolve or emerge constantly, signaling a healthy, competitive marketplace.

The combination of investments in digital and online technology (by publishers as well as others) and the formation of library consortia (assisted by publishers in many cases) in the United States and around the world has accelerated and broadened access to peer-reviewed literature and dramatically decreased cost of such access. AIP serves approximately 2,000 research institutions, and every person affiliated with these institutions has instant access to AIP journal content.

There is a growing presence and diversity of business models in the scholarly market. It is our belief that the government should support and encourage this diversity through its actions and policies through sustainable partnerships with publishers that would contribute to the US economy and maximize the productivity of the scientific enterprise. (For AIP's suggestions of 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 the responses to Questions 4 and 5. 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 last year in productive discussions with subject matter experts within the NSF and DOE, two US federal agencies that fund substantial research in the physical and biological sciences and engineering.)

As stated in the 2010 *Scholarly Publishing Roundtable* report,³ many publishers have made the decision to move toward increasingly open structures and archives,⁴ as enabled by open access business models

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/ngen_public/article.asp?id=200&did=47&aid=24781&st=&oaid=-1.

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>.

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 re-use 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 journals.⁶ AIP echoes the *Roundtable Report* assertion that no existing digital business model has demonstrated its viability to the satisfaction of all, and cautions against government endorsement of any single approach.

As part of the market's evolution and scholarly publishers' commitment to community and distribution of results, 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 Nation's HINARI, AGORA, and OARE Research4Life programs, HighWire's Developing Economies Program, and JSTOR's Developing Nations Initiative. Additional programs include those of EIFL, INASP, and TEEAL. 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, AIP invests considerably in new technologies for viewing and sharing its journals. Within just the past two years, AIP developed a mobile phone reader for journals, a professional (and freely available) social networking site for physical scientists (www.aipuniphy.org), and an electronic book platform. AIP also launched a multimedia journal on renewable energy (<http://jrse.aip.org>) and one of the first community-style journals in the physical sciences (<http://aipadvances.aip.org>).

Such ongoing investments in existing products and services and the development costs for new products are funded through subscription fees or author payments. AIP and most other scholarly publishers offer an open access option for authors, no matter what type of journal they decided to publish in. Through AIP's Author Select, authors have the option to choose open access for their published article. Less than one percent of authors choose to do so. *AIP Advances*, AIP's newest journal, is an initiative based on community-style review, rapid publication, is fully open access, and employs a Creative Commons license.

This ability for scientific publishers 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. 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. AIP is collaborating with other scholarly publishers to identify and address any existing access gaps through

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.

initiatives such as the low-cost article rental scheme pioneered by DeepDyve, the Research4Life consortium for developing countries (mentioned above), the patientINFORM portal for patients or their caregivers, the Emergency Access Initiative offered to communities affected by natural disasters, and free or substantially discounted access for public libraries, journalists, and high schools.

Nevertheless, based on our experimentation with a modest-cost article rental model (through DeepDyve), AIP remains unconvinced that there is a large unmet demand for public access: only a few thousand members of the general public attempted to access our scholarly content over a year's time, compared to the nearly eight million visitors to AIP content on our online platform.

To maximize the effectiveness of its efforts, government has an important convener role to play in developing standards for data and metadata, and making research more readily searchable and discoverable. Publishers are already working in partnership to develop standardized information and collections through initiatives such as CrossRef.⁷ (For more detail on this, please see response to Question 5.)

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 public.⁸ 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 already do this (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.)

AIP has been a leading participant in organizing working groups that are proposing and planning partnerships with NSF and DOE on access, linking of grantee reports to publications, data mining across agency-publisher databases, tools and methodology for identifying publicly funded work, and potential pilot projects in the above areas. (More detail on this can be found in response to Question 5.)

Government mandates for public access come at a significant cost to the US economy and to the scientific enterprise. The National Institute of Health's (NIH) PubMed Central (PMC) data indicates that two-thirds of its users are from overseas. This suggests that critical export opportunities for the industry may be compromised, resulting in loss of US jobs.⁹ Significant economic value added generated by the publishing industry could be wasted if revenue derived from sales in the global market is compromised

⁷ CrossRef (www.crossref.org) is a not-for-profit group founded by publishers in 2002 and maintains 50 million items. Almost 1000 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/porfaqs.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/>.

or eliminated. Furthermore, mandates often result in additional costs for publishers. For example, although only a very small portion of AIP's content is subject to the NIH public access mandate (AIP is primarily a physical science publisher), AIP had to incur costs to modify formats and procedures in order to deposit manuscripts into PMC. AIP remains concerned that PMC is shifting readers from the publishers' sites to PMC despite linking arrangements, thus undermining the value of the publishers' investments.

AIP has concerns about any government policy affecting global trade balance. The number of papers submitted to AIP journals from China exceeded the US submissions two years ago. In response, AIP opened an editorial and marketing office in China to help promote established AIP journals in the physical sciences in China, rather than see China develop competing international journals. Free and unimpeded access to US journal content, even if one were to factor in a short embargo period, will undermine our and other US publishers' needed revenue to establish business relationships in potentially lucrative and large global markets such as China.

In summary, AIP believes that publishers should continue to be free to experiment with various business models in the marketplace of ideas and economics. AIP endorses the *Scholarly Publishing Roundtable* 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 process does 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?

Scientific publishers, such as AIP, 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 AIP's position that agencies should provide free public access to final research reports and link to the peer-reviewed journal articles, which are available through a variety of access mechanisms. 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.

AIP 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 millions of dollars in annual losses to US publishers.

Nearly all scholarly publishers adopt liberal copyright policy, allowing authors to post copies of their manuscript on their individual and institutional websites with very little restriction, share copies with colleagues, and to 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?

Although a centralized data platform has some obvious advantages of simplicity of operation, the use of a centralized, government-controlled platform for a large corpus of scholarly content has significant downsides, including increased 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 and promote the widespread use of such standards.

AIP supports the recommendation of the *Roundtable Report* that states that government policies should be guided by the need to foster interoperability and encourages “. . . 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 National Science Foundation, the Department of Energy, 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 PubMed Central, though by some measures successful, is not a model that is universally applicable or necessarily the best model for the future. Indeed, the success of the World Wide Web 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 AIP'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, 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 CrossMark⁹ service, which electronically watermarks an article's Version of Record (VoR), and DataCite,¹⁰ 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.

AIP believes that it is unlikely that one optimal procedure for preservation and stewardship would emerge to become applicable across all of scholarly publishing. For now, AIP 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, Portico¹¹ and CLOCKSS.¹²

Long-term stewardship of content comes at significant cost that is being borne by publishers. In an era of dwindling federal resources, central federal repositories are 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

⁹ 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.

¹⁰ 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.

¹¹ 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.

¹² 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>).

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 common core properties that promote access to and interoperability among the content in all public access databases. Specifically, AIP 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. Agencies should fully engage researchers, institutions, and publishers working in fields that coincide with the agencies' missions, both in establishing initial public access policies and in modifying those policies as appropriate over time.

Many scholarly publishing organizations, such as AIP, were founded by scientists for scientists and fully embrace providing publishing and other services as their primary mission. As part of this, AIP's CEO was an active member the Scholarly Publishing Roundtable and subsequently helped organize working groups of nonprofit and commercial publishers to propose and implement joint projects with both the DOE and NSF with a mutually agreed-upon goals. Moreover, AIP is a cofounder of CrossRef and participates in a number of standards organizations such as the National Information Standards Organization (NISO—www.niso.org), National Federation of Advanced Information Services (NFAIS—www.nfais.org), and the newly formed consortium Open Researcher and Contributor ID (ORCID—www.orcid.org), with a purpose to develop unique researcher identifiers.

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. AIP has promoted the recommendation that publishers develop, in collaboration with funding agencies and CrossRef, a means of standardizing funder information and make 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 the specific grant that funded the research. It would be very expensive for agencies to obtain this information through data mining existing publisher databases.

This proposal has been endorsed by CrossRef and the major scientific, technical, and medical (STM) publishing trade associations: 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 and CrossRef, 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 Digital Object Identifier (DOI)] to the publisher's platform where article abstracts are freely available and the Version of Record (VoR) (maintained by the publishers) is available through a variety of access mechanisms, including innovative rental access models, which give the public instant access for a modest fee. More than 40 scholarly publishers are currently testing this access mechanism.

Standards and Identifiers: DOIs for Data Sets and Supplementary Material

Increasingly throughout the world, grant 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 Disambiguation

Name ambiguity and attribution are persistent, critical problems embedded in the scholarly research ecosystem. AIP encourages agencies to work in collaboration with publishers as well as universities, funding organizations, and corporations from around the world to eliminate this problem through ORCID. ORCID is a newly 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.

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 profile data contributed to ORCID by researchers or claimed by them will be available in standard formats for free download (subject to the researchers' own privacy settings) that is updated once a year and released under a Creative Commons license. All software developed by ORCID will be publicly released under an open-source software license approved by the Open Source Initiative (OSI). For the software it adopts, ORCID will prefer open source. ORCID is governed by representatives from a broad cross section of stakeholders, including publishers, library organizations, research institutions, and funding agencies (see <http://orcid.org/board-of-directors>).

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, 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 a 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, which includes an article processing charge paid by the author or their institution, delivers immediate and unrestricted online access to the final published article (defined by NISO as the Version of Record).

AIP suggests that agencies could work with publishers to set up experiments 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 of their articles? If not, how can authors be encouraged to utilize the available funds? (Several methods/messages could be tested.)
- 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

AIP 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 pilot would 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 Version of Record of articles connected to a particular research report or funded project.

If successful, this would result in interoperability between onsite 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 the pilot could 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, 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 (Versions of Record) 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.

(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 Office of Scientific and Technological Information, 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) 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 Version of Record, 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.

In addition, please see the response to Question 5 above for specific agency initiatives.

(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?

AIP believes that a uniform access policy or mandate for scholarly publications would be an ineffective approach. An 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.

AIP analyzed related industry data using the “cited half-life” metric as a relative indicator for how long journal titles within scientific categories are being accessed and cited, thus reflecting economic viability. The findings could help inform considerations related to embargo periods. Based on the evidence related to AIP journals and to journals covering physics and related sciences, significant economic uncertainty remains with the assignment of minimum embargo periods. In looking at a sample of several physics and related topics and AIP journals within those categories, AIP found that physics journals have a longer cited half-life compared to some other scientific disciplines, and furthermore, AIP Journals have a longer cited half-life than their respective physics category averages.

The chart on the following page provides are some examples.

Cited Half-Life of AIP Journals/Physics vs. Other Scientific Disciplines

The cited half-life for the journal is the median age of its items cited in the current year. Half of the citations to the journal are to items published within the cited half-life.

| Scientific Categories | Average Cited Half-Life within Sub-Category (Years) | Average Cited Half-Life of AIP Journals within Sub-Category (Years) | AIP Journals in Category |
|--|---|---|--|
| Applied Physics | 5.6 | 7.8 | <i>Applied Physics Letters, Journal of Applied Physics, Journal of Low Temperature Physics, Review of Scientific Instruments</i> |
| Chemical Physics | 7.1 | >10 | <i>The Journal of Chemical Physics</i> |
| Physics—Fluids & Plasmas | 6.6 | 7.6 | <i>Physics of Fluids, Physics of Plasmas</i> |
| Mathematical Physics | 6.5 | >10 | <i>Journal of Mathematical Physics</i> |
| Medicine, Research and Experimentation | 5.4 | | |
| Cardiac & Cardio Systems | 4.9 | | |
| Emergency Medicine | 5.7 | | |
| Robotics | 5.5 | | |

Source: Thomson Reuters, ISI Web of Knowledge, Journal Citation Reports, Year 2010

In lieu of trying to solve the public access problem by imposing a one-size-fits-all solution with a fixed embargo length for all articles that have some component of federal funding or introducing a complicated scheme for varying embargo lengths (as necessary to address field-specific conditions), AIP proposes a simpler system that allows government to accomplish public access in a way that is not only effective, efficient, and sustainable, but also keeps the US scientific enterprise thriving as it moves into the future.

To summarize the key components, AIP and a number of our colleagues from the scientific publishing community propose the following scheme to improve public access to the results of publically funded research:

1. Scholarly publishers as a group have proposed modifications to their author submission software so that all journal articles written after the implementation date would include funding agency information along with the standard metadata that is already being deposited in CrossRef and other standard bibliographic databases. This new metadata, which specifically tags the funding agency(s) responsible for the research leading to the journal article, would be deposited in the CrossRef database. (The CrossRef database has been developed and

maintained by this nonprofit consortium for the past 12 years and now contains the metadata for more than 50 million scholarly articles and related content.) Funding agencies can procure a license to this database at modest cost—many already have. Such a license provides access to the article metadata, including the critical article identifier (the DOI).

2. With the successful implementation of this funding identity proposal by scholarly publishers and CrossRef, agencies would have access to the 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 Digital Object Identifier (DOI)] to the publisher’s platform where article abstracts are freely available and the Version of Record (VoR) (maintained by the publishers) is available through a variety of access mechanisms, including innovative rental access models, which give the public instant access for a modest fee. More than 40 scholarly publishers are currently testing this access mechanism.
3. Scholarly publishers have proposed and initiated pilot projects with funding agencies to link agency research reports and related content on agency sites to publisher content tagged with the same funding information, thus expanding interoperability between agency and publisher databases and access to the linked content.

**Subject: Response to REQUEST FOR INFORMATION: PUBLIC ACCESS TO
PEER-REVIEWED SCHOLARLY PUBLICATIONS RESULTING FROM
FEDERALLY FUNDED RESEARCH**

Date: December 22, 2011 6:38:32 PM EST

This is in response to the RFI: Specifically it is in response to: "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."

As introduction, I am a Professor of Radiology at UCSF, I've been doing research for more than 40 years, and I am currently Principle Investigator of the Alzheimer's Disease Neuroimaging Initiative (ADNI) which is the largest grant funded in the world concerning Alzheimer's disease. (AD) ADNI is a multisite longitudinal clinical observational study aimed at validating imaging and biomarkers for diagnosis, early detection, and as inclusion and outcome measures in AD clinical trials. ADNI shares ALL raw, processed, and analyzed data with all qualified scientists in the world through its website UCLA/LONI/ADNI, which (to our knowledge) is unprecedented!. This data sharing has led to almost 300 peer reviewed publications. This experience has led me to be a passionate advocate of widespread sharing of all raw scientific data, after publication.

Although currently, most scientists do not share raw data, one big problem is that for those scientists who wish to share their data, there are no easy ways to do this. There is no federally funded mechanism for sharing raw data. Universities do not provide mechanisms for this. In fact there is not a lot of available software for data sharing, although DATAVERSE from Harvard (Gary King is PI) is an excellent open source tool for data sharing.

I believe that widespread sharing of raw scientific data which is described more extensively in the attached document, will have numerous benefits, and thus should be a national repository. Specifically there should be a National Raw Scientific Data Repository, where any scientist can deposit their data. The data could easily be linked to publication in PubMed Central. The cost of such a repository would not be huge. I believe that the benefits could be quite substantial leading to new discoveries, less scientific fraud, and a shift of the scientific culture leading to more cooperation and interaction. All of the benefits and issues are described in more detail in the attached document.

Sincerely

Michael W. Weiner, M.D.

Director, Center for Imaging of Neurodegenerative Diseases (CIND)
<http://www.cind.research.va.gov/>
San Francisco VA Medical Center
4150 Clement Street (114M), San Francisco CA, 94121 USA

Professor of Medicine, Radiology, Psychiatry, and Neurology
University of California, San Francisco, USA

Principal Investigator: Alzheimer's Disease Neuroimaging Initiative (ADNI)
<http://www.adni-info.org/>

Principal Investigator: Resource for MRI of Neurodegenerative Diseases
<http://www.rrmind.research.va.gov/>

Scientific Data Sharing Project
<http://scientificdatasharing.com/>

Plan to Achieve Widespread Sharing of Scientific Data

Michael W. Weiner M.D., Director, Center for Imaging of Neurodegenerative Diseases
Professor of Medicine, Radiology, Psychiatry, and Neurology, UCSF

Overall Goal: The overall goal of this project is to achieve widespread voluntary sharing of scientific data at the time of publication. The impact of widespread sharing of scientific data will be: to greatly increase the amount of information and knowledge which will be available to all scientists, to accelerate development and facilitate new discoveries of improved diagnostic and therapeutic methods leading to improved health and quality of life for society, and to stimulate the economic sectors which will use this information including the pharmaceutical, biotechnology, chemical, engineering, and computer/internet industries. In addition to the scientific and economic gains, substantial benefits from data sharing are also expected for education, scientific culture and communication. This goal will be achieved by a parallel approach of individual and institutional initiatives as discussed below.

Background: My experience as the Principle Investigator of the Alzheimer's Disease Neuroimaging Initiative (ADNI) has led me to the conclusion that widespread sharing of scientific data can be achieved now and that great scientific and economic benefits will ensue. ADNI is the largest NIH grant funded for Alzheimer's research (\$140 million total funding thus far) and all our raw data is immediately shared with all scientists in the world without embargo. The success of this project (more than 160 publications and 80 more submitted) demonstrates the feasibility of this approach and reinforces the success of other projects which share data, e.g. the Human Genome project. Currently data sharing is mostly done by large, well funded multi-investigator projects. There would be great benefit if much more raw data were widely shared, especially data from individual investigators in all fields of biological/medical science (and other areas of science as well). This would be best done at the time of publication, with the raw data being linked to papers in PubMed Central. It should be mentioned that in many fields within the social sciences (e.g. economics and political science), sharing of raw data at time of publication is already widely done, and is de rigeur. Widespread voluntary sharing of raw data will be achieved using two interlinked approaches:

Individual initiatives: Our laboratory at the Center for Imaging of Neurodegenerative Diseases will begin to share data at the time of publication during 2011. The raw data, e.g. individual subject data including numerical maps and images, will be available on a website, and access to the data will be achieved using available software (Dataverse) which allows the investigator control over data release. In addition to the raw data, a description of how the raw data was processed and analyzed, leading to the findings in the publication, will be provided. All data sharing will be performed with permission of the Institutional Review Boards and other university and governmental authorities concerned with human subjects and privacy protection. As this is being achieved we will identify other scientific groups who are sharing data and post them on our website scientificdatasharing.com. We will relate our experiences to the following: 1) Other collaborators in ADNI. ADNI scientists will be encouraged to share the raw data of their ADNI papers, and other papers from their laboratories. This should impact the field of Alzheimer's

leading to a greater acceptance of data sharing in the medical imaging and neurology fields. 2) Other faculty in the Department of Radiology at UCSF and our collaborators in Neurology and Psychiatry at UCSF. 3) If there is sufficient interest at UCSF, the Chancellor, Deans, and Department Chairs will be urged to make more widespread voluntary sharing of scientific data a UCSF priority/policy. Such actions would include providing storage space for shared data, and development of policies which would reward data sharing in the hiring and promotion process. The example of UCSF should urge the entire University of California system to encourage data sharing. 4) Other collaborators and colleagues in other universities around the world will learn about the work done at UCSF and in ADNI and will adopt similar policies (evidence of this happening is already available). 5) We will develop and test a “data sharing impact factor” which would allow scientists to cite the utilization by others, of data they collected.

Institutional mechanisms: Efforts are already underway to encourage increasing involvement by the NIH, NSF, and the National Library of Medicine (NLM), to promote and facilitate sharing of scientific data. These efforts will be strengthened as we gather increasing evidence of data sharing by our laborator and others at UCSF and elsewhere. First, the NIH and NSF will be encouraged to emphasize and expand their existing policies concerning data sharing and notify the scientific community of this greater emphasis. Second, the NIH should establish a small group of committed individuals who can help formulate policy in this area and suggest specific steps including generation of budgets to achieve specific goals. One approach to this would be for a few Institutes (such as the NIA, NIBIB, NCR, NIMH, NINDS, NIAAA/NIDA, Neuroscience Blueprint) to take a leadership role in creating a policy framework that favors open availability of scientific data. Third, and hugely important, would be to establish technical mechanisms for data sharing, such as a national system for storage of all raw scientific data, such as a national data repository or data bank. This can be achieved by the National Library of Medicine, with links on PubMed Central publications to the raw data. Another approach would be repositories supporting universities, foundations or private companies, using systems like Dataverse. The advantage of an NLM national repository is that its long term existance would not be in doubt. Fourth, means should be developed at NIH and NSF to incentivize scientists and institutions to share their raw data. This could be done: 1) by requesting reports in non competitive reviews, competitive reviews and/or new applications; 2) through the grant review process by instructing the reviewers to consider data sharing in assessing priority scores; 3) through special acknowledgements in publications; 4) by providing affordable access to infrastructure, i.e. software and media, which facilitates data sharing. Fifth, the NIH should provide funding for small grants aimed to promote and take advantage of shared data. New methods are already under development for putting pieces of data from different sources together and making a new whole collection of data which is greater than the sum of its parts. Having all the raw data on the internet will enable: 1) data mining: computer methods which troll the internet searching for particular types of information of interest; 2) cloud computing: computer methods which assemble different “clouds” of data and derive new knowledge using the combined information. Knowledge based industries are likely to benefit because of the increased accessibility to large amounts of raw data - especially the pharmaceutical and health care industry, chemistry, technology, engineering, etc. But ultimately the entire economy would benefit from improved knowledge. New technologies and new companies are expected to be developed to take advantage of the new information being made widely available. It is hard to predict the exact benefits of this type of activity. On the other hand, the costs would probably be relatively modest.

Its important to re-emphasize that the gains to be achieved by promoting widespread sharing of raw scientific data promise to greatly outweigh the relatively small costs involved in developing the necessary infrastructure. There is reason to believe that there will be substantial economic and public

benefits gained by widespread sharing of scientific data, because of the ability to link data sets, and make discoveries not related to the original goals of the data collectors.

In conclusion, we propose a two-fold plan (individual initiatives and institutional mechanisms) to achieve widespread sharing of scientific data which will speed the development of improved diagnostic techniques and treatments for a wide range of disorders, and to facilitate the growth of the knowledge based economy. Of course we would be happy to discuss this in more detail, and present a (modest) budget to implement this plan.

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

Date: December 22, 2011 10:16:12 PM EST

I declare myself to care deeply about this issue on behalf of the scientists at the Geophysical Institute (GI) and the students at the University of Alaska Fairbanks where the GI is located. I feel strongly that publicly funded research should be published Open Access.

The GI in my opinion does excellent research, much of it funded by a federal agency. The results are published but are not necessarily Open Access. Therefore many who might be interested in the research results never see them. This is to the detriment of:

the public who would like to read the research results;
the federal agencies which believed the work worthy of financing as the results are not disseminated as broadly as possible;
the scientists who did the research as the results are not disseminated as broadly as possible;
the University of Alaska Fairbanks where there is excellent research carried out but the results are not disseminated as broadly as possible

1 Re productivity of the American scientific enterprise.

UK and European institutions already are far along with Open Access policies. Just an hour ago I was looking at a list of scholarly societies which have OA journals and noted that the European Geosciences Union has several. Meanwhile the American Geosciences Union journals are priced beyond the reach of many libraries, e.g. community colleges, small liberal arts colleges such as Vassar College where I used to be Science Librarian. It is my firm belief that Europe and other countries will see their science efforts having a higher impact simply because everyone can access their results. To me it is highly significant that CERN has mandated OA and thus their scientists do not publish in Nature which doesn't allow making a PDF of a publication accessible in a repository. Their science hits the front page weekly!

2. The Creative Commons has addressed these questions very thoroughly. Since the publishers do not have to pay authors for their work I do not see the need for the publishers to obtain ownership. There are several Author Addenda in existence which confer all the rights to the publisher to publish and to republish should there be some future format change but allowing authors to keep ownership and rights to disseminate as they wish.

3. Lots of Copies Keep Stuff Safe. While ArXiv is facing financial issues there seems to be the will in the physics community to solve these issues. If it was run by a single Federal Agency it would be at risk of being axed at some budget cutback or being unavailable during ever more frequently threatened government shutdown.

4. I think there have been several successful partnerships already - Highwire Press with Pubmed, for example.

7. Yes

8. The ideal is no embargo period as CERN mandates. 1 year at most, 6 months would be better as is the case with PNAS, Proceedings of the National Academy of Science which makes all its articles OA after 6 months. No-one would dream of cancelling their paid subscription to PNAS because the information is free after 6 months. Springer allows OA to accepted papers while they are "In Press" but again no library will cancel a Springer journal just because the material is free if you are fast enough.

I do not speak for the Geophysical Institute nor the University of Alaska Fairbanks but I know every student who has discovered the ideal research paper but then can't access it in our library or using Google is frustrated but then digs some more and uses another research paper which is available. Or else the student makes an InterLibrary Loan request which means a delay in getting the information. Why should federal funding agencies pay for the research to be done without insuring access to the results? How ironic when it is government scientists who sometimes can't access the results without delay!

As a librarian, as an Alaskan resident, and as a tax-payer I whole-heartedly endorse OA for all research but believe it most imperative for federally funded research.

Thanks for allowing me this opportunity to share my ideas.

Flora Grabowska

--

Flora Grabowska
Librarian, Keith B. Mather Library
Geophysical Institute, University of Alaska, Fairbanks

December 20, 2011

To Whom It May Concern:

I am a Nobel Laureate in Physiology or Medicine and a working scientist in a commercial enterprise. I am writing to respond to the request for public comment by the White House Office of Science and Technology Policy on the important issue of access to the published results of our nation's investment in science and technology. I wish to express my whole-hearted support for the expansion of the successful public-access policy of the National Institutes of Health across federal science and technology agencies.

Scientific research and the use of scientific findings are increasingly conducted in a digital world. Our nation has invested heavily in ensuring that digital content is available via the internet and there has never been a greater need for policies to ensure that new research tools and techniques can be fully exploited. The removal of access barriers and the enabling of the expanded use of research findings is already transforming biomedicine. It has the potential to do the same for climate change, energy research and a myriad of other applications. Having witnessed the effect of open access on my own research area I would urge you to support timely public access to the results of all research, irrespective of discipline, that is funded with taxpayer dollars. My own research interests extend beyond medical applications and increasingly interdisciplinary work is becoming the key to progress.

I believe that Open Access should be immediate and free and should include the rights to use and re-use all published material without hindrance. I do not favor any embargo periods on such use since they impose unacceptable delays on my ability to keep up with the cutting edge of research and hinder my ability to build on those results in a timely fashion.

The open availability of federally funded research for broad public use in open online archives is a crucial building block in laying a strong national foundation to support accelerated discovery and innovation. It encourages broader participation in the scientific process by providing equitable access to high-quality research results to researchers at higher education institutions of all kinds – from research-intensive universities to community colleges alike. It can empower more members of the public to become engaged in citizen science efforts in areas that pique their imagination. It will equip entrepreneurs and small business owners with the very latest research developments, allowing them to compete more effectively in the development of new technologies and innovations. Open availability of this research will expand the worldwide visibility of the research conducted in the U.S. and increase the impact of our collective investment in research. Furthermore, it will encourage the adoption of such policies in all countries and show our recognition that scientific research knows no national boundaries.

I urge you to establish a firm policy to require that the results of the U.S. investment in science and technology be made openly available immediately upon publication.

Yours sincerely,

A handwritten signature in black ink that reads "R. J. Roberts". The signature is fluid and cursive, with a long horizontal stroke extending to the right from the end of the name.

/KO

Sir Richard Roberts Ph.D., F.R.S.
1993 Nobel Laureate in Physiology or Medicine
Chief Scientific Officer, New England Biolabs
240 County Road Ipswich, MA 01938-2723 USA
Tel: (978) 380-7405 / Fax: (978) 380-7406
Email: roberts@neb.com

Assistant: Karen Otto
Tel: (978) 380-7206 / Fax: (978) 380-7406
Email: otto@neb.com



Task Force on Public Access to Scholarly Publications
National Science and Technology Council
Executive Office of the President
Washington, DC 20502

23 Dec. 2011

Response to Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research (76 Fed. Reg., 68518, 4 Nov. 2011)

Dear Members of the Task Force on Public Access to Scholarly Publications:

With nearly 400,000 members in more than 160 countries world-wide, IEEE is the world's largest professional association dedicated to advancing technological innovation and excellence for the benefit of humanity. In addition to our conferences, standards and other activities, IEEE publishes more than 150 transactions, journals and magazines, which represent more than 30% of the world's annually published literature in electrotechnology, computing and related fields.

IEEE is a strong supporter of public access within a context that protects and advances other important societal interests inherent in scholarly publishing, including peer review, protecting the integrity of the research archive, and preserving the intellectual property rights of authors and publishers. IEEE's long-standing policies permit authors to "self-archive," i.e., to post their articles either on their personal web sites or their employers' web sites, consistent with open access practices. IEEE is also experimenting with several open access business models with the goal of supporting sustainable open-access publications. Our views were outlined in greater detail in a 2007 IEEE Position Statement on "Scholarly Publishing", which is available on-line at:

http://www.ieee.org/documents/IEEE_Publishing_Principles.pdf

IEEE believes that learned societies serve an important role as the unbiased custodians of peer-reviewed scholarly literature, archived as a trusted version of record. Among the principles that should be respected in the dialogue on public access are the following:

- It is important that any approach be financially self-sustaining. All publishing and archiving activity carries costs that require offsetting revenue models. The development of any repositories – public or private – must account for funding to support costs. Sustainability is enhanced when US agencies avoid investing in information delivery systems that duplicate services already provided by publishers. This is especially important at a time when the US government is trying to control its operating costs and reduce the national debt.

- Any approach must be protected from censorship by any source, including government policy agendas.
- Any approach must protect the rights of authors to publish in the forum that they believe is most appropriate for their work.
- Interoperability is an excellent way to achieve these goals because the research results could continue to reside in a version of record at the publisher's site but be more easily reached from other indexing or search services. Indeed, the IEEE already supports these efforts through its participation in industry organizations like CrossRef and NISO.

Dealing with these complex issues will require more than just the exchange of written documents. IEEE strongly supports formal collaboration among stakeholders in public-private working groups, as recommended in the 2009 Scholarly Publishing Roundtable Report. We would be pleased to provide representatives to serve on collaborative groups convened to deal with these issues.

Please see detailed responses to OSTP questions following this letter.

Sincerely,

A handwritten signature in black ink, appearing to be 'Moshe Kam', written in a cursive style.

Moshe Kam, Ph.D, P.E.
IEEE President and CEO

IEEE Response to OSTP Questions on Public Access to Scholarly Publications

With respect to the specific questions posed in the Request for Information issued 3 Nov. 2011 by the Office of Science and Technology Policy on behalf of the National Science and Technology Council, IEEE is pleased to provide the following input:

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 set policies that assure long-term sustainability of the distribution mechanisms favored by public policy. Free access to research results benefits those who can exploit those results, either to advance the research further or to create commercially viable new products. However, without sound business models to support continuous publication and archiving of new results, over time dissemination will become less reliable. A proliferation of publications and archives that lack strategies to support ongoing costs poses a near-term threat to established journals supported by conventional subscription plans. In the long term, even soundly established open access publications may suffer in a field that is largely populated by underfunded and unsustainable journals and archives.

One of the priorities in any Federal public access policy should be to ensure affordable access to research results by small and medium-sized business, especially start-ups and other known job creators.

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?

Peer review must continue to be available, free of government influence and management. Each published paper should be available in a single final authoritative version. This version must be unable to be altered by the author or anyone else. Authors should have the freedom to choose where to publish and readers should have the freedom to choose where to read.

Policy should allow for public access to the author-prepared manuscript of an article, as accepted by the publishing journal and incorporating the author's changes that result from the publisher's peer review process. However, this publicly available version should always include a Digital Object Identifier (DOI) or a link to the publisher's version of record. In no

case should policy permit the harvesting or public posting of the final published version, in whole or in part, of an article.

Respect for US copyright law would also logically imply that when funding agencies make copyrighted work available to the public, they do so without encouraging or condoning the reuse of that work without permission of the copyright holder. In other words, no provisions of existing copyright law are to be abrogated by government action.

Peer review is only one part of the value added by publishers to scholarly research, which also includes editorial review, abstracting, indexing, and provision of accessible on-line portals for subject-matter searching. Increasingly, publishers are expected to share with the library community the responsibility for archiving published works in perpetuity. Public access policy should recognize and preserve the full range of value added by scholarly publishers.

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?

Through appropriate regulations and standards, interoperability can be fostered among publishers and funding agencies to ensure sharing and archiving of funded research results to the benefit of the public good.

On the other hand, there is no compelling argument for the US Federal Government to make substantial investments to duplicate the considerable and expensive infrastructure already developed over many years by scholarly publishers. There is also no assurance that a central government operation can provide the archiving capability already long established by publishers and libraries.

The government's stewardship interest is adequately served by a combination of distributed publication of articles supported by interoperable search and retrieval mechanisms, and by maintaining custody of research reports filed by principal investigators with their Federal funding agencies.

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 least costly approach to enhancing public access to government research captured in the form of scholarly publications is to collaborate with scholarly publishers. Collaborations can be explored in several areas, including literature and data hosting, interoperability standards, and common search tools and portals.

Government should work with current owners of content repositories (e.g., not-for-profit learned societies, commercial publishers, universities) to enable public access to research funded by Federal monies. Government investing in the development of its own content repositories makes little sense from an economic perspective. The cost and time required would be significant and not necessarily viewed as prudent, especially in the current economic times. Government could far more easily leverage the expertise of not-for-profit learned societies, commercial publishers and universities in operating large repositories of scientific papers by, for example, providing funds to enable the publishers to appropriately tag content under a public access policy and thus ensure transition from closed to open status within a single hosting platform.

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?

Rather than establishing new content repositories, the US Government should support and cooperate with the scholarly publishing industry in its efforts to set standards for interoperable search and improved discoverability of content. CrossRef, a not-for-profit association founded by scholarly publishers in 2002, has pioneered the development of the Digital Object Identifier (DOI) as a unique identifier to locate every published scholarly work. IEEE and other scholarly publishers would support an initiative to enhance article metadata with information clearly identifying the agency responsible for funding the described research. Agencies would save considerable effort and expense by supporting improvements to DOI metadata that would automatically capture this essential information.

Similar collaboration between public and private sectors will lead to success in efforts to provide identifiers for data sets that are created in connection with funded research. Examples include DataCite (www.datacite.org) and the NISO/NFAIS Working Group on Supplementary Journal Information (www.niso.org).

Another example of a collaborative approach among publishing partners is the Open Researcher & Contributor ID (ORCID) project (www.orcid.org), a successful public-private partnership with 275 participating organizations. This project addresses name ambiguity among individual authors and the resulting difficulties in consistent author attribution.

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?

These goals can be advanced by fostering the development of standards for interoperability among content creators and agency repositories, the US Government and its agencies can avoid costly duplication of services while delivering to taxpayers the results of funded research.

It is important to remember that publication is an integral part of the research process. Nothing is free. Supporting the publication costs of researchers publishing in an open access environment as part of the research grant is essential.

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?

Both types of identified content – books and conference proceedings – have unique characteristics that differentiate them from scholarly journals.

Conference proceedings provide early reports of results by researchers, and therefore may be seen as highly desirable content to be made publicly accessible. In the technology fields in which IEEE participates, the conference record often plays a more important role than it might in basic science. As a result, the differences between a conference article and a journal article are narrowing. Therefore, if conference proceedings are to be considered for public access, the same principles of interoperability contemplated for journals should also apply to conference articles, as should the principle of directing users to the publisher's site to obtain the version of record. Finally, sustainability in public access to conference articles should be encouraged by use of grants to support publishing costs.

Books and book chapters are an entirely different class of publication, usually synthesizing the work of an author from a long period of time and generally more difficult to connect to specific research grants. In terms of financial investment, books present publishers with more concentrated risk than do journals, and therefore federally mandated public access to book content is likely to be far more problematic to publishers. From an author's standpoint, books generally involve enhanced levels of intellectual property, including detailed publisher contracts and ongoing royalty arrangements that would be impacted by a public access mandate.

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

The answer to any question of appropriate embargo period is complex and may vary according to the field of science, engineering or technology researched and how the public

interest in free access is defined, including the mission and goals of the particular federal-funding department or agency.

The current NIH model, allowing 12 months' delay to public access, is motivated by the rapid advances characteristic of health-related research combined with the understandable desire of patients to get timely access to news of potential medical treatments and newly discovered health risks.

By contrast, most engineering, science and other types of technical information have a longer useful life. The IEEE's premier journal, the *Proceedings of the IEEE*, has a cited half-life of over 10 years (Thomson Reuters Journal Citation Reports, 2010 Science Edition). As a further example of the long useful life of content in the engineering field, an analysis of usage in 2010 of the IEEE Xplore Digital Library – the full offering of IEEE's archived content – shows that usage of articles published in 2009 and 2010 amounted to about 34% of total usage. In other words, nearly two-thirds of 2010 usage was of articles published before 2009; an embargo period of 12 or even 18 months is not adequate in this context.

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.

IEEE has no additional items to propose at this time, but is eager to participate in public-private working groups to address public access issues.

In conclusion, IEEE appreciates the opportunity to provide input to the Task Force's deliberations and stands ready to answer questions and provide additional information as needed. If we can be of any further assistance, please contact Kenneth Moore, Director, IEEE Book & Information Services (e-mail: k.moore@ieee.org; tel: 732-562-3954).

TO: Office of Science and Technology Policy publicaccess@ostp.gov
FROM: Daniel Lee, Tucson, Arizona
RE: Request for Information: Public Access to Peer-Reviewed Scholarly
Publications Resulting From Federally Funded Research
DATE: Friday, December 23, 2011

The following comments are in response to the request for information issued November 3, 2011, by the Office of Science and Technology Policy (OSTP) regarding recommendations on approaches for ensuring long-term stewardship and broad public access to the peer-reviewed publications arising from federally funded research. I would like to thank OSTP for the opportunity to respond and contribute to the conversation. My comments follow.

(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 important step agencies could take to leverage the results of research funding to help maximize economic growth is to require publications that follow from such funding be made openly available to all. A truly open solution would allow anyone to read, revise, redistribute, and reuse both the content and the publication.
- On the one hand, this would get the research findings out to small businesses and other entrepreneurs, who currently have little or no access to the literature, in order for them to build on them and create new products, services, markets, and jobs.
- On the other hand, this would allow for the creation of new markets in products and services making use of the publications themselves to expand the usefulness of the literature. These new value increasing ventures would further increase exposure of research findings leading to greater impact of the federal investment.

(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?

- First, it is important to recognize that copyright resides with the author/creator unless and until s/he transfers that right to another party such as a publisher. A research funder could reasonably require authors to adjust these rights in some way or other. Research funders that are the author's employer might indeed be the rights holder themselves and thus impose even greater control over how rights are managed. Given that scientist/authors and such funders are usually more interested in getting read and having impact than in controlling rights, protecting the intellectual property of publishers doesn't seem like a helpful place to start.
- This isn't to say that publishers don't add value and that their investment in that value doesn't need protection of some sort. They do and it does. Some of that protection will come from being the first one to release a publication, and some protection will come from being responsible for the version of record. More specific answers to this question will come out in the answers to other questions posed here regarding possible embargo periods and the possibility of distributed archives.

- In this same vein, though, it should also be noted that fee based can certainly compete with free (and that the entertainment industries contrary cries of lost revenue from file sharing are grossly overstated). Most often this can be done by adding value, but in many cases a fee based publisher's copy could likely compete merely by being the source for the imprimatur of quality and being the known source. A good example is UMI/ProQuest's ability to generate sufficient revenue from the services they provide around dissertations while many of the universities that supply them with these dissertations make copies openly available through various repositories. The point here is that Federal agencies may not need to provide much protection of publisher investments.

(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 may be possible to achieve broad public access to peer-reviewed scholarly publications that result from federally funded research through a metadata repository that points to openly available copies on publishers' websites. However, some mechanism or incentive would have to be in place for the funding agencies to ensure compliance with the agencies' interest in broad public access and the "long-term stewardship" of those publications.
- As funding agencies relationship is with the scientist/authors and not publishers, it's difficult to conceive of what that would be. Assuming the broad possibilities for the multitude of audiences, including machine audiences, intended by the full interpretation of "open" in the first comment, publisher compliance would likely entail adjustments to publisher web platforms. Full compliance should also include making copies available in a standard mark-up language such as XML to support wide range of discovery and re-use tools.
- The strongest incentive would seem to be that driving audiences to publisher web platforms in this manner would be one strategy to protect and exploit their investment in the original publication. However, the difficulty in ensuring full compatibility of multiple platforms strongly suggests funding agencies may indeed require custody of copies of publications to achieve both goals of broad public access and long-term stewardship.

(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 comment 3.

(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 metadata repository that points to openly available copies on publisher websites envisioned in comment 3 – or any cross agency repository - would require the interoperability suggested here. By being responsible for the metadata that leads to persistent connections to peer reviewed publications agencies would be able to make certain the necessary metadata was present. Standard fields such as author, author affiliation, paper title, journal title, date, volume, issue (where applicable), pagination, funding agency, grant number/program, and persistent link

seems like a good start. For discovery, all the metadata as well as the full text of publications should be available for indexing by any viable enterprise from general search engines to indexing and abstracting services targeting specific research areas.

(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?

- This is a very important goal. Simplifying processes that funded scientists and their institutions go through to meet agency requirements for accountability would lower overhead costs and allow for more research to occur. Common, consistent requirements and processes across funding agencies would also contribute to greater return on investment. Researchers want to do research and share the results, not satisfy bureaucracies.
- One way to simplify compliance by funded scientists/authors would be to include submission of copies of manuscripts or published papers, or links to already openly available publications, in the standard reporting mechanisms. Simple procedures that fit into an existing workflow have the best chance of achieving desired ends with minimal additional burdens.
- This also raises the need for agencies to include funding for publication costs in grants. Quite often papers are published about the results of the funded research after the funding period (and sometimes funds) has run out. This suggests that allowances need to be built in to grants to allow for spending on publication costs after the normal grant period.

(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?

- Any peer reviewed publication that directly documents the analyzed findings of research that is funded by Federal agencies should be considered for inclusion.

(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?

- Immediate open access to peer-reviewed publications that result from federally funded scientific research is the policy ideal. However, different business models to pay the costs of publication exist and likely should be accommodated to ensure a wide range of distribution channels. Many journal publishers already track their revenue streams and are well aware of the point of reasonably diminishing returns. Some of these found that point to be one or two months. Any embargo that is allowed has the possibility of protecting the investment made in publication. However, no embargo should be allowed that exceeds the point where exclusivity demonstrably protects that investment. Given the six month maximum embargo already in place by policy in Canada, the United Kingdom, and the European Union, this would appear to be a useful cap to achieve consistency across research funders.

Access to and preservation of the scholarly record is fundamental to the scientific enterprise in the United States. Nonetheless, striking a balance between that ideal and the practicality of administering a complex system with many stakeholders is a difficult problem. As the America COMPETES Act states, the government needs to “invest in innovation through research and development, and to improve the competitiveness of the United States.” The University of Pennsylvania Libraries grapples with this challenge every day, and we feel that there are many potential ways federal agencies can help to solve this difficult issue.

Our peer institutions have been able to accomplish much with the greater availability of content that open access allows. Some universities such as Rice University (<http://cnx.org/>) and the University of Michigan (<http://www.lib.umich.edu/spo/reprints.html>) have experimented with ways of providing books and articles for free and charging for value-added services such as reprints. No doubt commercial publishers will soon follow the same model. Yet, to do so will require them to relinquish intellectual property rights. Traditionally, scholarly publishers have helped to sustain scientific research by obtaining scholars’ copyrights over their articles. In an era when it was very difficult to reproduce information (other than by copy machine), doing this made sense, but in the internet age, that model is breaking down. What is required is a shift in thinking about how scholarly publishing works. Academic publishers in fact have been shifting their thinking for years. ProQuest Information and Learning (<http://www.proquest.com/en-US/>) provides access to databases of content of public domain and other materials for which they do not own copyright (including dissertations, newspapers, and early printed books) and EBSCO (<http://www.ebsco.com/>) indexes the work of other publishers so that it can be searched in ways that Google and other less specialized search engines cannot provide. In other words, some publishers are making money utilizing content which is already freely available.

How can federal agencies help to facilitate this shift in thinking about scholarly publishing? In some ways, they already are. The National Endowment for the Humanities (NEH) provides “digital humanities” grants to fund new kinds of scholarship

(<http://www.neh.gov/ODH/GrantOpportunities/tabid/57/Default.aspx>), and, perhaps more importantly,

the National Science Foundation (NSF) has an entire office of cyberinfrastructure

(<http://www.nsf.gov/dir/index.jsp?org=OCI>) engaging with the larger issues of scientific publishing. In

other countries agencies of the central government like the Joint Information Systems Committee (JISC)

in the UK (<http://www.jisc.ac.uk/>) and the Canadian Research Knowledge Network (CRKN,

<http://www.crkn.ca/home>) have created fairly broad collaborative approaches to scholarly publishing

issues. In the United States, there has been a more decentralized, non-government approach. Groups like

the Association for Research Libraries (ARL -<http://www.arl.org/sc/>) and Ithaka (<http://www.ithaka.org/>)

have helped to develop smaller scale projects like SCOAP3 (<http://scoap3.org/>) for Physics and JSTOR

(<http://www.jstor.org>) for humanities and social sciences.

In addition to continuing the work that many agencies have already begun, the United States has the

opportunity to create the best of both worlds: a hybrid of decentralized entrepreneurial initiatives like

ARL and Ithaka, and more centralized government approaches such as the JISC and CRKN. If the various

federal funding agencies (like NSF, NIH, NEH, etc.) were to form a joint task force on publishing,

comprise up of multiple stakeholders, such a group could investigate the current publishing efforts in the

US, recommend new models and new ways for collaboration, and, perhaps most importantly, recommend

strategies for interoperability. There are many standards for creating interoperable metadata, but the issue

goes beyond metadata. Electronic technology allows new kinds of scholarship and requires new methods

for making it available and sustainable. In all, what the United States needs in order to compete with

countries like the UK and Canada is more collaboration between stakeholders in addition to the current

ad-hoc, decentralized approach. Open Access mandates such as the NIH are an important step, yet they

also encourage many uncoordinated approaches to a similar problem. The federal government could

provide a great service by coordinating the competing parties in a way that no one else can, and it already has several models both within the US and outside that it can use to create such a dialogue.

From: Dr. Richard Epand <epand@mcmaster.ca>
Date: December 26, 2011 3:34:03 PM EST

I would like to respond to your request for information.

While complete public access to publications of government funded research appears to be a just and valid goal, there are some aspects that put this conclusion into question. Whatever the system, publication of research is not without cost and there has to be a mechanism to pay for these costs. Currently in open access journals, a larger fraction of the cost of publication is passed on the author. This results in valuable government research grant funds being used for publication costs. There is also an issue of the efficiency of the dissemination of research results in terms of cost. Commercial publishers, while extracting a profit, do publish scientific results at a relatively low cost and without an over-riding system of author verification of granting agency policy and separate depositing of manuscripts. These publishers also maintain their own historical databank, which is not always the case with open-access journals.

There are now established open-access journals as well as journals published with initial restricted access. These restrictions are only for a limited period of time and these journals tend to be readily available in libraries. This system gives authors a choice, it is simple and efficient. Any increased regulation requiring free access to all publications will be detrimental to science in the long run.

Sincerely,
Richard M. Epand
McMaster University
Dept. Biochemistry and Biomedical Sciences

Name: Thomas W. Jeffries

Email: twjeffri@wisc.edu

Affiliation: University of Wisconsin-Madison and the USDA Forest Products Laboratory

Address: Institute for Microbial and Biochemical Technology
One Gifford Pinchot Drive
Madison, WI 53726

I am writing in response for a request for information on “Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research” [Federal Register Volume 76, Number 214]

I am a senior scientist in the USDA (Microbiologist, ST) and have over 30 years of experience as a reviewer, editor and senior editor for scientific journals. I am on the editorial boards for *Applied and Environmental Microbiology*, *Applied Microbiology and Biotechnology*, *FEMS Yeast Research*, *Metabolic Engineering*, *Enzyme and Microbial Technology* and *Protein Expression and Purification*, and I am a Senior Editor for the *Journal of Industrial Microbiology and Biotechnology*. I am a periodic or ad-hoc reviewer for *Science*, *PNAS*, *Nature Biotechnology* and a number of other journals. As such I have considerable interest in the editorial, review and publication process.

Your request for information concerning “...approaches [to ensure] long-term stewardship and broad public access to the peer-reviewed scholarly publications that result from federally funded scientific research...” seems appropriate in the light of the current rapid expansion of on-line journals.

While broad public access certainly has increased as a result of the “Public or Open access; online journal” model, I am less confident about the issue of long-term stewardship, and I am very concerned about the quality of papers published in such journals. To be clear, online journals provide a rapid, inexpensive way for publishers to distribute scientific articles to their readership. Increased revenue from the Open Access publication fee benefits the publisher, and the author benefits from a higher citation frequency – regardless of the quality of the paper.

The problem with this model is that the costs associated with publication represent a very small fraction of the cost of generating and distributing scientific data. The major costs are the experiments themselves, along with writing, editing and reviewing. The Publisher basically gets the Authors’ time for writing the papers, the Editors’ time for handling the manuscript and the Reviewers’ time for reading the papers at little or no cost. While this assures that the scientific data are widely distributed, it gives a strong monetary benefit to Publishers to create more on-line journals in order to collect more processing fees. It also creates a very strong demand for qualified reviewers. Unfortunately, with so many new on-line journals the demand for qualified editors and reviewers greatly exceeds the supply. Consequently, more papers of lower quality are accepted for publication and widely distributed despite their questionable quality.

In fact, the Editor in Chief of a respected peer-reviewed publication told me “Originality is not a criterion for publication for on-line journals.” He was not expressing this as an opinion in derision so much as in support of starting a new on-line journal in order to raise additional funds. The publisher for his current journal was encouraging him to start an on-line journal as a way to create new revenue. He saw this chiefly as a money making proposition.

The Open Access aspect virtually guarantees a bump of two or three points in the citation frequency of a manuscript – regardless of the quality of the contents. Thus, it is no longer possible to judge the quality of a journal simply by citation or download frequency.

It is not at all clear to me that the many new publications that are coming on-line have the capacity to maintain records and public access over the decades in which public access is necessary. Nor is it at all clear that these new journals are serving a useful function.

In the past three years, we have seen a flood of new manuscripts from China and India. While a few are well written and deal with novel topics, most are simply a rehash of things that were published many years before.

Open Access and electronic publishing certainly have given opportunities to specialized scientific areas, and I am certain that there are useful models that could assure the distribution of high quality literature while maintaining low publication costs, but the economic pressures to increase publication frequency is currently straining the editorial and review processes.

Subject: open access comments

Date: December 26, 2011 8:49:35 PM EST

A bit about me before responding to the several questions concerning future open access and information dissemination sustainability. I am an executive editor for a methods journal and handle the assignment of reviewers and subsequent disposition of 50-60 manuscripts a year. I am also a retired professor emeritus of genetics. My telecoms are listed below.

Question 1: Open access greatly enhances the scientific research and enterprise of the country as well as facilitates the development of new technology driven businesses. One extremely annoying problem with top journals in the sciences, Nature journals in particular, is that there is no free access to archived papers. I recently tried to access the original DNA articles in Nature as well as a paper demonstrating the genetic code is a triplet code but could only have obtained if purchased for several dollars each. Yet, I can go to my university library and copy for a few pennies charged to a research account. This is patently absurd.

Solution 1: I cannot formulate a specific policy apart from federally funded research agencies disallowing publication in such high end journals unless they change their free access to archived articles. How long should access be delayed after print or online publication? 30-60 days. Perhaps the PTO policy with regard to submitted patent publication might be looked into. The journal Science has a good policy in this regard.

Question 2: Protecting IP is reasonable and expected of any policy and I think it is covered by most university policies at the moment. At my own university, public disclosure, e.g., a seminar or poster or other meeting presentation, can be held up for 60 days until the IP is provisionally or otherwise submitted. This is necessary, especially for young scientist in training - undergraduates, graduate students and postdoctoral fellows - to ensure the freedom to publish their findings to enhance their careers.

Question 3: I think the government and federally funded agencies should keep their noses out of it as much as possible except as noted above. The critical issue is trying to protect the publishers/ journals' interests as well as the scientists and the IP if applicable while maintaining relatively rapid open access. Instantaneous open access through ejournals will probably evolve and occur in the future, but the current publishing model has served the scientific enterprise

well and deserves some consideration during the transition. Said another way, what will happen will happen as scientists collectively adjust and use the new electronic access features and the print model will die away.

Question 4: Not that I am aware of however Google's copying of library contents may make this a moot point.

Question 5: No comment apart from what was said above.

Question 6: As noted in Solution 1 - develop a policy with regard to the time lag after publication when public access is available online. Libraries provide the dissemination point for most research scientists and companies have their own intralibraries of critical journals. Perhaps moving libraries and their ability to license a large collection of journals through individual publishers is the way to pave to free online access. Library of Congress can do this, too?

Question 7: Yes.

Question 8: Answered above. And this is the critical, essential question that needs to be resolved. Good luck and thanks for asking.

Bob...

Robert Ivarie
Emeritus Professor
Department of Genetics
University of Georgia
Athens GA 30602-7223



MAX PLANCK SOCIETY

MPI for Developmental Biology • Spemannstr. 35 • D-72076 Tübingen

Dept. of Molecular Biology

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

Dr. Detlef Weigel
Tel.: 07071/601-1411
Fx.: 07071/601-1412
Mb.: 0179/676-9032
Email: weigel@weigelworld.org
URL: <http://weigelworld.org>

December 27, 2011

Dear Colleagues,

I am writing in response to the RFI on Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research, published in the Federal Register Volume 76, Number 214, on Friday, November 4, 2011. I am currently based at the Max Planck Institute in Germany, and am writing as a US citizen whose previous research in the US was supported by various Federal agencies, including the NIH, NSH, USDA and DOE. I am an elected member of the US and German national academies and a Foreign Member of the Royal Society of London. I have served on the Editorial Boards of several scientific journals, both ones published by for-profit entities such as Elsevier Inc. and ones published by not-for-profit scientific organizations such as the American Association for the Advancement of Science. Since September of this year, I am Deputy-Editor-in-Chief of a new open access journal, eLife, jointly sponsored by the Howard Hughes Medical Institute (US), Wellcome Trust (UK) and Max Planck Society (Germany).

The pace of science has greatly accelerated in the 25+ years that I have personally witnessed. Ironically, although electronic dissemination on the Internet should enable the entire scientific community to rapidly access the latest results, only elite institutions with deep pockets can afford to pay the often very steep subscription charges that apply to the majority of scientific publications. Rapid access, in turn, is not only absolutely essential for all of us to stay competitive, but it is also important to avoid duplication of research efforts and ensure that Federal grant monies are used as efficiently as possible.

As other areas, large parts of science are currently transformed by "Big Data", where those who produce such data often analyze only a small fraction of these data. This creates tremendous opportunities for anyone with access to the Internet and clever ideas for data analysis, thus greatly leveraging the initial investment, which often was in the form of Federal grants. Again, not being able to access primary publications in a timely manner impedes the efficient use of such data.

For these reasons, I wish to express my strongest support for an official Federal policy in support of open access to the scientific literature.

Yours truly,

Detlef Weigel
Director, Max Planck Institute

From: Aaron Presnall
Subject: Response to RFI on Public Access to Digital Data
Date: December 27, 2011 7:13:40 AM EST
To: publicaccess@ostp.gov, digitaldata@ostp.gov

In response to the RFI on Public Access to Digital Data, I would like to endorse the National Digital Stewardship Alliance (NDSA) statement. The link is http://digitalpreservation.gov/documents/NDSA_ResponseToOSTP.pdf Also, find the statement as a pdf file attached.

Aaron Presnall
President
Jefferson Institute
Washington, DC



**Response to Office of Science and Technology Policy Request for Information on
Public Access to Digital Data Resulting from Federally Funded Scientific Research**
Submitted by the National Digital Stewardship Alliance (NDSA)
January 2, 2012

Introduction to the NDSA

The National Digital Stewardship Alliance (NDSA) was founded in July 2010 to extend work begun in 2001 by the National Digital Information Infrastructure and Preservation Program (NDIIPP) at the Library of Congress. The Alliance has over 100 members from educational institutions, non-profit organizations, businesses and local, state and federal government agencies, as well affiliations with international organizations. Its mission is to establish, maintain, and advance the capacity to preserve our nation's digital resources for the benefit of present and future generations. [1] Members of the Alliance are taking action to preserve access to our national digital heritage by:

- broadening access to our nation's expanding digital resources
- developing and coordinating sustainable infrastructures for the preservation of digital content
- advocating standards for the stewardship of digital objects
- building a community of practice around the management of distributed digital collections
- promoting innovation
- facilitating cooperation between government agencies, educational institutions, non-profit organizations, and commercial entities
- fostering the participation of diverse communities and relationships across boundaries
- raising public awareness of the enduring value of digital resources and the need for active stewardship of these national resources.

Supporting communities of practice for preservation and access

The values of the Alliance are highly relevant to establishing approaches for ensuring long-term stewardship and encouraging broad public access to unclassified digital data that result from federally-funded scientific research. When applied, these values support the practical development of communities of practice capable of gaining consensus to support preservation and access to digital data. The shared expertise and common experience of these communities result in stakeholder buy-in and adoption of policies and

standards. The National Digital Stewardship Alliance member organizations are bound as a community by the following values.

Stewardship. Members of the NDSA are committed to managing digital content for current and long-term use. The members of the NDSA are actively ensuring sustained access to the digital content that constitutes our national legacy and empowers us as leaders in the global knowledge economy. Individually, these organizations support the management of digital resources; the Alliance is committed to protecting our nation's cultural, scientific, scholarly, and business heritage.

Collaboration. Collaborative work is the centering value of the Alliance; it is a value shared by all members and a priority in work with all organizations and associations. Approaching digital stewardship collaboratively allows the NDSA to coordinate effort, avoid duplicate work, build a community of practice, develop new preservation strategies, flexibly respond to a changing economic landscape, and build relationships to increase capacity to manage content beyond institutional boundaries.

Inclusiveness. The NDSA is a collaborative effort to preserve a distributed national digital collection for the benefit of current and future generations. We value the range of experience, the potential for innovation, and the fault-tolerance that heterogeneity brings. We believe the preservation of digital information is a pervasive challenge and that engaging across different communities strengthens the nation's digital preservation practices and increases the likelihood of preserving content now and into the future.

Exchange. Members of the Alliance encourage the open exchange of ideas, services, and software. This leverages the commitments of each member to increase the capacity of the entire stewardship network. Participation and engagement result in innovations and benefits that can be shared by all. The Alliance is committed to transparency and all products generated or produced by the Alliance will be circulated under open licenses.

Support sustainable action through policy that catalyzes collaborative work on preservation and access within and across scientific disciplines

Community-based approaches to the challenges of rapid change and high volume within the data domain have proven to be the most successful in the long term. The Blue Ribbon Task Force on Preservation and Access recommended that for research data “Each domain, through professional societies or other consensus making bodies, should set priorities for data selection, level of curation, and length of retention.” [2]

The report validated experience over the last ten years of digital preservation work. A study of the networks developed through the NDIIPP program indicated that participating institutions bring to the network their own resources, interests, and organizational culture. Under the auspices of a neutral convener and honest broker, natural networks emerge over time through participation in shared activities and problem solving. As these networks form, the larger network becomes more complex, but also stronger and better able to withstand stresses and strains. [3]

The Opportunities for Data Exchange (ODE) project supported by the Alliance for Permanent Access and the European Union also takes a cross-cutting community approach to preservation and access to digital data. “The potential answers to grand challenges of our times require...the inclusion of an interoperable data sharing, re-use and preservation layer to the emerging eco-system of e-infrastructures...All stakeholders in the scientific process must be involved in the design of this layer; policy makers, funders, infrastructure operators, data centers, data providers and users, libraries and publishers...” [4]

An exemplar of collaborative community efforts is the Dataverse Network project [5] recently described by the National Research Council of the National Academies as the “State of the Practice in Data Sharing.” [6] The Dataverse Network is “unique in being designed to explicitly support long-term access and permanent preservation. To this end the system supports best practices, such as format migration, human-understandable formats and metadata, persistent identifier assignment and semantic fixity checking. In addition, many threats to long-term access can be fully addressed only by collaborative stewardship of content, and the system supports distributed, policy-based replication of its content across multiple collaborating institutions, to ensure the long-term stewardship of the data against budgetary and other institutional threats.” [7]

Foster public values and support for stewardship of digital data beyond mandating data management plans.

Policy should assert the value of research data and provide mechanisms to support the preservation, discoverability and access. To relieve frustration and confusion about actions the policy should provide a clear direction for funders, researchers and stewardship organizations. The Blue Ribbon Task Force recommended “Funders should impose preservation mandates, when appropriate. When mandates are imposed, funders should also specify selection criteria, funds to be used, and responsible organizations to provide archiving. They should explicitly recognize “data under stewardship” as a core indicator of scientific effort and include this information in standard reporting mechanisms.” [8]

Leverage substantial national and international efforts for common practices that support interoperability.

Substantial efforts have been made to pave the way for interoperability, re-use and re-purposing. Emerging practices for data citation, licensing and protocols for data sharing and sustainable re-use are becoming enough to adopt more broadly. Notable in these areas are work on the Data Seal of Approval by the Data Archiving and Networked Services that promotes sustainable access to digital research and provides training and advice about archiving and reuse.[9] LOCKSS is a community initiative that provides libraries with digital preservation tools and support so that they can easily and inexpensively collect and preserve their own copies of authorized e-content. [10] The

Data-PASS organization promotes collaborative, institutional stewardship of research data, permanent data archiving, and citation that permits results to be verified and re-purposed. [11] DataCite collaboratively addresses the challenges of making research data visible and accessible through data citation.[12] The Creative Commons project, Science Commons, has focused on protocols for sharing scientific data that includes licensing and mitigating legal barriers.[13]

Summary of Major Recommendations

- Support sustainable action through policy that catalyzes collaborative work on preservation and access within and across scientific disciplines
- Establish policy that catalyzes collaborative work on preservation and access within and across scientific disciplines
- Foster public values and support for stewardship of digital data beyond mandating data management plans.
- Leverage substantial national and international efforts for common practices that support interoperability.

Additional Responses on Selected Questions

The principles and recommendations above apply broadly to the set of questions posed by the RFI. The responses below exemplify how the principles can be applied to the individual questions, and highlight relevant NDSA activities in these areas.

(1) What specific Federal policies would encourage public access to and the preservation of broadly valuable digital data resulting from federally funded scientific research, to grow the U.S. economy and improve the productivity of the American scientific enterprise?

The most effective policies in this regard would mandate data deposit into publicly accessible repositories. In the absence of such a policy, there are already cases of data which have been lost. The Federal policy framework should move public access to data away from the current idiosyncratic environment to a systematic approach that lowers barriers to data access, discovery, sharing and re-use.

Many members of NDSA provide repository services at low cost or through cooperative arrangements. Members of the NDSA also provide repository services that provide legal, technical, procedural and statistical controls necessary to protect data confidentiality while ensuring long. And the NDSA provides a model of institutional collaboration that supports stewardship, discovery and accessibility. An example of a free access service is ViewShare.org, a platform for empowering curators, archivists, and librarians to provide access to the digital collections they are preserving through a shared interface. This

service provides the dual benefit of making data more broadly available and accessible while also making it easy for end users to copy and make use of the data in other environments. [14] The NDSA content working group is also working toward developing a clearinghouse for at-risk digital collections to help match data to potential preservation partners.

(3) How could Federal agencies take into account inherent differences between scientific disciplines and different types of digital data when developing policies on the management of data?

Each domain and discipline should be empowered to set priorities for data selection through, level of curation, and length of retention, through professional societies or other consensus making bodies.

Notwithstanding, there are still baseline conditions or requirements that apply to all data regardless of discipline, particularly as they relate to archiving and preservation. For most data, “open access” is needed not only for the short term, but for the long term. And scientific disciplines have focused primarily on short-term access. There are critical standards for metadata exchange, fixity information and verification, and persistent citation that can support long-term access to data, preservation, and the long-term reproducibility of public results. Such baseline standards should be applied all scientific data. Among the range of important new standards for preservation and access there is still little knowledge about which standards are being implemented in which situations. The NDSA Standards working group is working on inventorying these standards and exploring how they are currently being used by NDSA member organizations. More than advocating the need for standards there is a clear need to understand which standards are being used in which situations and use that information to promote the usage of standards that are leading to results.

(5) How can stakeholders (e.g., research communities, universities, research institutions, libraries, scientific publishers) best contribute to the implementation of data management plans?

There exists diversity in approaches for data management within various scientific communities, which is healthy for various reasons. In cases where communities have resources for data management, it is worthwhile to build upon existing infrastructure (e.g., the Data-PASS archives). However, it is critical that even in these cases the community service provider demonstrates rather than assert capability. Far too often, terms such as archiving or preservation being used loosely without associated evidence of meeting specific requirements. Memory institutions such as archives, libraries and museums have an extensive track record with these functions and collaborative organizations such as NDSA could serve the essential purpose of developing or implementing frameworks that thoroughly test and certify assertions. In this respect, work from the NDSA innovation working group toward developing a “Neighborhood Watch” system for repository quality assurance could serve as the basis for establishing clear, externally verifiable reporting. [14]. The group has identified a pressing need for

an objective, repeatable, independently verifiable and simple way for an external agent to periodically retrieve content, verify its bit level integrity and publicly announce the results. This is a clear example of how assertions about data management could be tested and certified.

(6) How could funding mechanisms be improved to better address the real costs of preserving and making digital data accessible?

The most important step would be to communicate that the real costs of preserving and making digital data accessible are indeed legitimate and necessary costs of the overall research enterprise. Researchers routinely include publication costs within their research proposals -- the costs of ensuring long-term access reuse of data should be treated in the same way.

(7) What approaches could agencies take to measure, verify, and improve compliance with Federal data stewardship and access policies for scientific research? How can the burden of compliance and verification be minimized?

One of the key points in this context is that it is easier to verify compliance through systematic approaches. It is easier to verify compliance of library-based or community-based data archives than to check thousands of individual researcher hard drives. Technical infrastructure components such as persistent identifiers and appropriate licenses represent critical mechanisms through which compliance and verification can be automated thereby reducing costs.

(13) What policies, practices, and standards are needed to support linking between publications and associated data?

There is widespread consensus within the research community that it is essential to link publications and underlying or associated data. Many NDSA partners are leaders in this area. The peer-reviewed publication is viewed as the final “snapshot” of the research process and outcome. One of the most important considerations from a policy, practices and standards is a requirement to use persistent, unique identifiers for publications, data, authors, etc. These identifiers not only bolster the linking of publications and data, but also help foster the re-use and development of new services by people and machines. While there are multiple identifier schemes, at this point, perhaps the most important policy decision would be to require using persistent identifiers instead of relying upon existing mechanisms such as website URLs.

References

[1] The National Digital Stewardship Alliance: <http://www.digitalpreservation.gov/nds>

[2] Berman, Francine, and Brian Lavoie, et al. 2010. *Sustainable Economics for a Digital Plant: Ensuring Long-term Access to Digital Information*. Final Report of the Blue

Ribbon Task Force on Sustainable Digital Preservation and Access supported by the National Science Foundation, et al. Washington, DC:
http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

[3] Library of Congress. 2010. *Preserving our Digital Heritage: The National Digital Information Infrastructure and Preservation Program (NDIIPP) 2010 Report*. Washington, DC: <http://1.usa.gov/hmw2lj>

[4] Alliance for Permanent Access. 2011. “Opportunities for Data Exchange (ODE) Project”: <http://www.alliancepermanentaccess.org/index.php/current-projects/ode/>

[5] King, Gary. 2007. An introduction to the Dataverse Network as an infrastructure for data sharing. *Sociological Methods and Research* 36:173-99.

[6] National Research Council. 2011. *Communicating Science and Engineering Data in the Information Age: Panel on Communicating National Science Foundation Science and Engineering Information to Data Users*. Preprint. Washington, D.C.: National Academies Press: <http://bit.ly/NCSES>

[7] Altman, Micah and Jonathan Crabtree. 2011 “Using the SafeArchive System: TRAC-Based Auditing of LOCKSS,” Archiving 2011 Final Program and Proceedings, May 16–19, 2011, Salt Lake City, Utah: 165–170. Society for Imaging Science and Technology: <http://bit.ly/tLzUmr>

[8] Berman et al. 2010.

[9] Data Seal of Approval: <http://www.datasealofapproval.org/>

[10] LOCKSS: <http://lockss.org>

[11] DataPass: <http://data-pass.org/>

[12] DataCite: <http://datacite.org/>

[13] Creative Commons project, Science Commons: <http://creativecommons.org/science>
<http://wiki.creativecommons.org/Science>

[14] ViewShare: <http://viewshare.org>

[15] Abrams, S, Cruse, P, Kunze, J, Minor, D, Smorul, M. 2011. “Neighborhood Watch” for Repository Quality Assurance. Presented at Designing Storage Architectures for Preservation, Washington, DC: <http://1.usa.gov/uXj2Mf>

Beth A Sullivan, PhD

beth.sullivan@duke.edu

Assistant Professor, Duke University School of Medicine

Associate Editor, PLoS Genetics, PLoS ONE

Durham, NC

Currently, only 8% of all scholarly papers are Open Access (OA). Thus, 92% of papers that are published are unavailable as a “market opportunity and size” for new innovative tools to start new business more jobs. Most importantly the majority of papers funded by tax dollars are not accessible to the public who has funded them. In my opinion, this is a serious breach of the trust between scientists and the American people who are the financial benefactors in scientific research endeavors.

Comment 1: NIH has implemented a public access policy that has proven to be extremely cost-effective at \$3.5-4.6 million per year. This means that an investment of approximately 1/100th of 1 percent of NIH’s overall budget results in access to 2.2 million scientific articles. The database for searching and accessing the articles is accessed daily by over 500,000 users. Most of these users are actually from OUTSIDE of education, emphasizing the demand for this information throughout the public sector.

Scientific research is an inter-dependent process in which each experiment is informed by the results of previous experiments. The scientists who perform research and the professional societies that represent them have a great interest in ensuring that research results are disseminated as immediately, broadly and effectively as possible. Electronic publication of research results offers the opportunity and the obligation to share research results, ideas and discoveries freely with the scientific community and the public. OA articles provide scientists with new tools (machine reading, computational tools, analysis programs) to obtain and read more information faster. Such new uses will encourage opportunities for commercial development through private investments to capitalize on a public resource. Importantly, OA expands the potential for interdisciplinary research that increase the breadth, depth and value of our public investment in science.

The type of OA is a crucial issue. In my opinion, there should be full OA, with free immediate access coupled with rights for full digital re-use. Mechanisms are necessary to enable full use, such as distribution, reuse, text and data mining, computations, and creation of derivative works. Appropriate licenses should be implemented, such as Creative Commons (CC-BY) licenses.

Comment 3: If a federal agency, such as NIH or NLM, maintains custody of all published content, it should make it easier to ensure accountability across the various institutes that funded the research. It would also assist grant reviewers in identifying publications during the peer review process to evaluate previous work and progress of grantees. In my opinion, distribution of information by a private source is not ideal, unless there is a federal mandate of how content of databases is structured and maintained, and curated.

Otherwise, distribution of information by multiple private sources is at risk of compromising standards and archiving of data. In other words, each source for archiving will set up the databases as they wish. It seems more logical that there be consistency in how the information is collected, tagged with identified, maintained and accessed.

Comment 4: Publishers are in an important area to participate in public/private partnerships by providing approved repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles.

Publishers need to commit to providing an open access option for any research article published in any of the journals they publish and declare a specific timetable for transition of journals to open access models. They should work with other publishers of open access works and interested parties to develop tools for authors and publishers to facilitate publication of manuscripts in standard electronic formats suitable for archival storage and efficient searching. Finally, they should ensure that open access models requiring author fees lower barriers to researchers at demonstrated financial disadvantage, particularly those from developing countries. This specific issue is going to be difficult to implement since many of the high profile biomedical journals still cling to impact factors as measures of standing and clout in the field. A broader issue as well is antiquated university tenure system that rates faculty progress and stature based on journal impact factors rather than article impact metrics. Article levels metrics measures how often an article is downloaded and cited in other scholarly works. Article level metrics (ALMs), such as those adopted by the PLoS journals, appear to be a more accurate method by which to evaluate the impact and reach of a specific journal article, rather than simply citing the impact of journal as a whole. ALMs could also be directly linked to grants and funding agencies to emphasize and track the direct impact of federal funding by a) individual investigator, b) research initiative, c) research topic, or d) technology (to name a few).

Comment 5: Metadata affords ways to enable specific actions and parameters, rather than basic item description. Parameters could include journal name, conference, publication title, researcher name, researcher grant code, institution publication data, and other identifiers such as subject/topic, technique/technology or method. Provenance metadata is a distinguishing feature of this effort as a means of facilitating specific desirable actions for use, reuse and analysis of publications. It should be a priority to develop machine-readable/machine-operable metadata for rights of downloading and reuse.

Comment 6: For OA models to work effectively, there must be consistency of requirements and mandates for depositing peer-reviewed literature across all funding agencies. Consistency in policy will maximize returns to taxpayers by ensuring complete results are widely available in a timely fashion. It will also minimize duplication of efforts in depositing literature that is coming from various publishing houses/publishers. Policies should integrate articles with grants managements systems, so that there is better agency accountability and transparency between the type of science that was funded, the level at

which the research was funded, and ultimate publication outcomes. The policies should create opportunities to create and maximize productivity management tools for federal and internal reporting.

Comment 7: In my opinion as basic scientist, if an investigator's laboratory research is federally funded, that means that most of his/her activities and those of the lab are supported by taxpayer dollars. Thus, honing an area of expertise over years/funding periods is afforded by federal funds. This expertise is called upon for writing book chapters and it make sense that encyclopedia entries or book chapters should also be publicly available. In fact, reviews and book chapters will be more intellectually accessible to the public since they are usually written broadly and with the intent to educate.

Another topic that might be considered (perhaps in future discussions of public access and scientific accountability) is archiving of all data from federally funded projects. Publications resulting from federally funded projects only account for a fraction of the actual data generated from the grant or fellowship. Many experiments result in unexpected or unexplainable findings or even failures that are not suitable for publication but could be useful to the broader research committee. In other words, these experiments, while considered "scientific dead ends" or failed experiments within a specific line of research, the data could aid another researcher who is approaching a similar question or using the same technology. The benefit derived by a subsequent researcher could be in the form of "don't make the same mistakes we did in this experiment" or "we attempted this protocol/technique multiple times by changing a, b and/or c". Since ALL of the research performed under a specific grant is federally funded, it might be useful to the broader scientific community to have a searchable repository of experimental findings that were too brief or incomplete for a proper journal publication. It will increase accountability and also illuminate to lawmakers and the general public the amount of trial-and-error that is involved in basic science research. Of course, this endeavor would need to be well managed, would require enormous amounts of storage, and benefit from proper curation of experimental data/results to prevent such a database from becoming a waste heap for useless information.

Comment 8: Immediate access to publications funded by federal resources is the ideal option and will optimize scientific and commercial utility of information contained in these articles. However, there should be an option for investigators to select a longer embargo period, particularly if the research is directly linked to a patent application.

Subject: Response to RFI in Federal Register Vol. 76, No. 214 Friday, November 4, 2011 regarding ACRA reauthorization

Date: December 27, 2011 2:49:38 PM EST

Carol L Hodes, PhD Independent consultant and researcher State College, PA 16801

Mr Wackler, As an independent researcher, I conduct literature searches to support the submission of grant proposals and articles for publication. State and federally funded libraries have discontinued many of their hard-copy journal subscription in favor of online subscriptions. The absence of hard-copy journals curtails research efforts because many of the full-text online journals require subscriptions, either institutional or individual, a practice that makes research difficult due to the cost of subscriptions. Since a good proportion of published research is publicly funded, it should be publicly available. Answers to specific 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? Yes. Journal articles that contain results of research studies supported with public funds should be publicly available as full text documents, either on the journal's website and/or when displayed in search results from databases (e.g., Proquest).
2. Regarding protection of 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. The purpose of publication is to disseminate information and introduce new ideas or results into the scientific scholarly community. Access to new results and ideas improves and expands the knowledge base of emerging concepts. Perhaps federally funded research should be exempt from copyright law.
3. The pros and cons of centralized and decentralized approaches to managing public access to peer-reviewed scholarly publications that result from federally funded research. I know little of centralized versus decentralized databases. Access from the online home page of a journal and the search results from databases such as Proquest are both important. The important thing is to enable access through a variety of databases and search engines.
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. Many publishers of scholarly journals, such as Sage or Routledge, for example, limit access to many journals to subscribers. Articles describing research that is government funded should be

available as full text from the journal home page, as well as a government database.

5. 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 grant number should be used as a identifier to link articles to the funding source.

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? Keywords and the ability to search for literal text strings can provide access from the online home page of journals and databases.

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? Conference proceedings often contain timely content similar to that of journal articles and should be treated the same way. 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? Due to the rapid pace of change in knowledge there might not be an appropriate embargo period.

Carol Hodes, PhD



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

December 26, 2011

Dear Task Force Members:

I write to offer comment on the question of open access to the scientific literature. My views are based on five years of service as the Editor-in-chief of the Proceedings of the National Academy of Sciences, and now as the Editor-in-Chief of a new open access journal, eLife, sponsored by the Howard Hughes Medical Institute (HHMI), the Wellcome Trust (WT) and the Max Planck Society (MPG). As an active investigator with over 35 years of experience in life science publishing and editing, I wish to express my strongest support for an official federal policy in support of open access to the literature. For most of my time in science, I would have argued that the scientific community was well-served by a mix of commercial and academic publishers, each satisfying a particular need of the community for work of general or specific interest. However, with the advent of online publication and the explosion of information from a burgeoning scientific enterprise, the need for rapid and complete access to the literature has grown to a point where a policy of open access must be adopted for all journals of record.

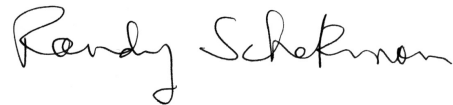
In my experience at PNAS and before that as an Associate Editor at the Journal of Cell Biology and the ASCB journal, Molecular Biology of the Cell, I found that a policy of free and open access two to six months post-publication was perfectly compatible with a business model that was revenue neutral (PNAS and MBoC) or which generated a small profit for the publisher (JCB). Unfortunately, commercial presses that rely on a subscription model of profit have not developed a business plan to accommodate the need for open access. Instead, through license fees that oblige the purchase of bundles of titles, the commercial presses have achieved unreasonable profit on the public investment in science and have forced cash-strapped university libraries to limit content to federally funded investigators.

In response to this challenge from the most prestigious publishers of life science research, the HHMI, WT and MPG have joined forces to start a new high-end life science journal that will be free to authors and all readers. These organizations accept publication expenses as part of their mandate and are unwilling to have publication access

limited to those institutions that can afford a site license. For some years now, investigators at HHMI and WT have been encouraged to publish in open access journals but this has not prompted the most selective journals to move in that direction. As a result, this new initiative, eLife, is intended to provide investigators and the entire life science community a vehicle for free exchange of work at the highest level of impact. In polling the attitudes of the investigators supported by HHMI, WT and MPG, we learned that a sizable majority of respondents (922/1052) consider open access from the moment of publication to be fairly or very important in their decision to submit a paper to the new journal.

Given the trends in publication and the explosion of knowledge in the life science community, I would argue that a federal policy to mandate publication in open access (immediate or rapid – two to six months – access) journals is in the best interests of science and the American public.

Sincerely,

A handwritten signature in black ink that reads "Randy Schekman". The signature is written in a cursive, flowing style.

Randy Schekman, Professor
Molecular and Cell Biology, and
Investigator, Howard Hughes
Medical Institute

Subject: Peer-Reviewed Scholarly Publications

Date: December 28, 2011 11:36:53 AM EST

Dear Sir/Madam,

I am the Editor in Chief of Combustion and Flame, the premier Combustion journal, and below please find my comments.

1. Authors, editors, publishers are committed to providing the broadest possible access to our publications. The goal is to expand access to the latest breakthroughs in the scientific research and developments in academic thought.
2. The main purpose of a scholarly journal is to report on original research or experimentation in order to make such information available to the global research community.
3. It requires considerable investment to maintain the highest standards for peer-reviewed scientific publication and sustainable mechanisms are required to enable this system to cooperate. This is threatened by access policies that do not take these costs into account, and it is critically important that any new policies do not damage the publishing institutions on which the Federal Government and science depend.
4. Overall, access to research articles is good, particularly in the United States, where 97% of researchers are happy with access to journal articles.
5. Publishers are committed to indentifying any gaps in access and have developed a number of mechanisms to close gaps in sustainable ways. This includes initiatives to broaden access for researchers in developing countries, patients, and the public.
6. Any effort by government to establish policies for peer-reviewed research should be done in consultation with all affected stakeholders, ensuring that such policies do not undermine the sustainability of the peer review publishing system which is necessary to ensure the quality and integrity of scientific research.

Sincerely,

Fokion N. Egolfopoulos

Professor

Editor in Chief, Combustion and Flame

Dept. of Aerospace & Mechanical Engineering

Viterbi School of Engineering

University of Southern California

Subject: Response to RFI: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

Date: December 28, 2011 1:57:20 PM EST

Dear Members of the National Science and Technology Council / Task Force on Public Access to Scholarly Publications:

The Association of Southeastern Research Libraries (ASERL) represents libraries at 40 leading research institutions in 11 states. Our members firmly support free and open public access to the scholarly publications and the use/re-use of data emanating from research that is funded with taxpayer dollars.

ASERL believes the American public and the American economy should be able to reap maximum benefits from investments made with our tax dollars. Additionally, the NIH has existing channels that can be leveraged so these data and publications can be made publicly available with relative ease and at minimal cost to the government. We believe it is important that these opportunities for economic growth, scientific innovation, and educational leadership be open to all. A federal policy requiring immediate, full, free access to research data and publications -- with the rights to re-use the data -- can significantly increase opportunities for innovation and economic growth. Such a requirement would also allow new research to be incorporated into teaching and learning as quickly as possible. And lastly, we believe it is the fair and right thing to do!

To state the converse, ASERL believes limitations on the products of publicly-funded research (e.g., embargoes or behind paywalls or other barriers) are against the public interest. Such opportunities should not be limited by the desires of private publishers or others who seek short-term private profits from long-term public investments.

In short, we urge you to create a policy that maximizes the public's return in its investments – full, free, immediate public access to the data and scholarly publications resulting from publicly-supported

research, with full rights to re-use the data. This can be accomplished via the Creative Commons CC-BY license model – a well-established, legally valid method for ensuring access and re-use while providing due credit to the original author(s). We believe anything less would reduce America’s return on our investments and limit opportunities for the future.

We ask OSTP to create policies that mandate public access to these publications and data without barriers. Many thanks for your consideration of our thoughts on these important matters.

--*jeb* John Burger, Executive Director **Association of Southeastern Research Libraries**

Subject: RFI Response

Date: December 28, 2011 2:09:02 PM EST

Constance J. Britton

United States Agricultural Information Network (USAIN) Wooster, OH I am writing on behalf of the United States Agricultural Information Network (USAIN) to respectfully respond to the Request for Information for recommendations related to public access to peer-reviewed scholarly publications resulting from federally funded research.

USAIN (usain.org) is an organization of over 150 agricultural information professionals that provides a forum for discussion of agricultural issues, takes a leadership role in the formation of a national information policy as related to agriculture, makes recommendations to the National Agricultural Library (NAL) on agricultural information matters, and promotes collaboration and communication among its members. USAIN has testified before Congress, played an advisory role in the National Agricultural Text Digitizing Project, written a national agricultural literature preservation plan, served on blue ribbon panels to review NAL services, and participated in the selection process for new NAL Directors. Our members are skilled librarians with knowledge of the modern theories, principles, practices, techniques, and policy issues pertinent to the current practice of librarianship and information science. Many of our members work at Land Grant institutions with extensive federally-funded research programs and are experienced in acquiring, organizing, and preserving scientific and agricultural data. The USAIN Executive Council is privileged to provide the following input related to this important topic of public access to information.

Comment 1. Access to published research information is a fundamental requirement by scientists, students, innovators, entrepreneurs, and other interested citizens in order to further the process of the discovery of knowledge ultimately leading to the creation of new products, jobs, and the growth of the

economy. Published research information, created at the public's expense, but that sits behind a subscription pay-wall, may not be equitably available to all who might benefit from its content. Even funded researchers at public universities may have limited access to published literature due to the lack of a library or personal subscription. A system that requires that the publication output of federally-funded research be deposited in one or more publicly-accessible repositories would achieve a necessary level of equitable access.

Comment 2. Nothing in a public-access system should threaten the protection of intellectual property as covered by copyright. In fact, greater access to research information and data will ensure greater visibility and recognition of an author's intellectual achievements.

Comment 3. The national libraries are the obvious leaders in creating a system that supports public access to federally-funded research output. Libraries currently possess the knowledge, expertise, if not the resources, to develop and implement a repository system. In the case of research funded by the United States Department of Agriculture, the National Agricultural Library (NAL) has a key role to play in the acquisition and preservation of research publications and data resulting from USDA funding. NAL librarians and staff possess the knowledge, skills and expertise to manage paper and digital collections, and by applying metadata, to facilitate the discovery of stored content. The NAL is a demonstrated leader in the development of the metadata and the technical standards needed to insure interoperability of these systems. However, in creating a repository system, it is essential that new funding be appropriated to engage sufficient staff and provide the technical infrastructure necessary to insure success of this program. The best model of a centralized approach is that provided by the NIH-mandated deposit of peer-reviewed research articles in PubMed Central, where the deposit requirement and sufficient program funding have made this repository successful. A comparable repository for USDA-funded research could be managed by the NAL as an expansion of the existing NAL Digital Collections (NALDC). The advantages of a centralized repository are

better control of the deposit process, author compliance, and consistent metadata applications. Funding agencies managing a smaller grant portfolio may have a more difficult time supporting a separate repository, so centralization would benefit these agencies. Centralization also minimizes issues of interoperability, consistency and redundancy. Many universities maintain an institutional repository and could help facilitate required deposits within the institutional site or a centralized repository. Even with clearly articulated standards, achieving full interoperability across many repositories may be a challenging goal.

Comment 4. No comment.

Comment 5. Again, the national libraries have the requisite skills, experience and mandate to define and implement the standards that must be put in place to create an interoperable repository system. The minimum metadata elements for describing bibliographic information are currently well-defined by the Dublin Core metadata standard. These elements can be readily derived from publisher data and incorporated as part of the deposit. Adherence to this standard will facilitate the sharing of data from multiple repositories and lead to discovery by the public. Metadata standards are critical for describing publications and data within a repository, but institutions are also faced with the added challenge of increasing access to those resources. Resources must be highly discoverable and understood within a larger context of scientific data and research. For that to happen, several things must occur: 1) the advanced support of author disambiguation initiatives, such as ORCID, which "aims to solve the author/contributor name ambiguity problem in scholarly communications" 2) a general mandate requiring federally funded authors to identify their funding source when submitting publications to a repository and 3) the development and support of Semantic Web technologies that allow for the re-purposing, reuse, and analysis of publication and other data. By design, Semantic Web technologies are machine-readable; continuing to encourage the development and accessibility of these technologies would allow for flexible re-purposing of data, regardless of the model - centralized, decentralized,

or mixed-model - chosen by Federal agencies.

Comment 6. The richest benefit with the least burden will be gained by utilizing the expertise that resides with the national library system, including not only our national libraries, but also the university libraries. Libraries already possess extensive knowledge about the development and management of publication repositories. Key to this approach, however, is sufficient funding to implement this program.

Comment 7. While, ideally, the public would have access to all information created as a result of federally-funded research, the peer-reviewed journal articles are the critical body of work to be captured in a repository system. This literature represents the publishable results of the funded research, whereas conference proceedings often report preliminary results and book chapters may synthesize research from multiple sources and often come much later in the scholarly communication process. The journal literature is the most useful report of research outcomes to make available to the public.

Comment 8. No comment.

Constance J. Britton

Librarian

Ohio Agricultural Research and Development Center

The Ohio State University

Subject: public access

Date: December 28, 2011 6:37:49 PM EST

Regarding the issue of public access, I believe that publications resulting from government sponsored research should be freely available to the public. Data being collected prior to publication need not be made publicly available because the researchers have priority in publishing their own results, but after publication the papers and supporting datasets should be publicly accessible.

Gary D. Stormo, Ph.D.
Joseph Erlanger Professor
Department of Genetics
Center for Genome Sciences
Washington University Medical School

Subject: RFI response on Public Access to Peer-Reviewed....

Date: December 28, 2011 8:10:13 PM EST

Dear Task Force on Public Access to Scholarly Publications:

Thank you for providing the opportunity to comment on the several questions posed below (in snippet form)--

Comment 1. Are there steps....

All articles resulting from publicly funded (or publicly subsidized) research should be immediately and freely accessible ("Open Access" hereafter), so that the public can fully use them without commercial restriction. This accessibility maximizes public good from taxpayer investments.

Open Access:

Increases scientific productivity and generates new uses and applications for this research

Allows everybody to stay on top of cutting-edge ideas and discoveries

Increases the usage of each research article as established with bibliometric measures (the number of reads, citations, and uses of the article).

Expands the potential for interdisciplinary research, and encourages contributions by "unexpected" colleagues

Allows scientists to extract useful information, by automated sifting of text, images, tables, and contained data. The "machine readable" form of articles (the XML version) inherent in Open Access means that data mining and searching capabilities can be applied to large numbers of articles impossible by manual methods.

Federal agencies will have an improved accounting on the outcomes of their funded research: Agencies will have better information to assess the value of existing expenditures and target funding on the most promising research

A public-access policy is cost effective by leveraging existing infrastructure from the NIH's PubMed and PubMed Central.

Under the current system, the taxpayer pays for the research and then must pay again to read the results! A public access policy eliminates this "double charge".

It is critical to note that much of the money paid for subscriptions or access to commercial journals goes to major publishing firms headquartered OUTSIDE the US, contributing to their substantial profit margins and draining US coffers (federal and private) and contributing to the current highly unfavorable balance of payments. Major research libraries around the country pay exorbitant subscription fees, as they have no alternative, excepting open access repositories.

What type of access is needed? To derive the maximum benefit from federal research paid for by the taxpayer, full Open Access is absolutely necessary: free immediate access coupled with full rights to re-use fully in digital environment.

"Full Use" includes not only access to the content but also distribution, reuse, text mining, data mining, computation, creation of derivative works, IN PERPETUITY. Such use allows researchers to unlock additional value from the research investment and build on the results of others rather than repeating what has already been done.

Comment 2. What specific steps can be taken to protect the intellectual property...

It is important to clearly enunciate that commercial publishers do nothing to generate the new intellectual property in a scholarly research article. The authors write the article. The authors write the article. The authors write the article. Not the publisher! The publisher does the barest minimum in terms of service and added value that they can get away with.

Publishers claim that they run the peer review system, but nowadays that is farmed out to a journal management service that charges about \$20 per paper to handle submission and review. The Senior Editors or Editor in Chief, almost always uncompensated (and almost always faculty at leading universities), chose the reviewers. The journal management systems are automated enough so that little manual assistance is required on the part of the publisher. The publishers claim that they serve to grade the quality of the article and confer approval and prestige to the work. Excepting a handful of journals, this "prestige" depends on the quality and insightfulness of the Editors, Editorial Board, and peer reviewers more than the reputation of the journal itself.

New statistics on article usage now largely supplant the qualitative importance of publishing in such and such a journal, because the cumulative use and impact of each article can now be directly measured. This eliminates the highly subjective importance of publishing in Science versus Nature versus Cell or New England Journal of Medicine and so on. Each article now stands on its own, with its

individual quantitative statistics of impact, regardless of where it was published. "Journal prestige" is becoming an outmoded concept, a remnant of the past.

Sometimes the authors pay for copyediting before an article is accepted for publication. The Scientific Editors and the Peer Reviewers of most journals provide their services for free (to emphasize, the publisher does NOT pay them anything). The copyediting and proofreading of an accepted article are paid for by the authors with processing or handling fees or minimum reprint orders charged by the publisher. Commercial copy editing services are about \$60 per article for a typical scientific article on the current fee-for-service free market. Many publishers have invented artificial "styles" for each specific journal that they manage to exclude competition in the free market for copy editing services. The publisher often breaks at least a few standard rules of scientific and technical editing guidelines set forth in classic style guides c.f., the Council of Science Editors guide entitled, "Scientific Style and Format". This trick allows the publisher to re-edit and re-proof read to meet a given journal's deliberately contrived quirky "style". Thus, the publisher claims credit for the entire work having done next to nothing except a little copyediting.

In other words, the authors give the research results and carefully crafted text away for free, along with all the copyrights to the publisher. Why would authors do that? It seems insane, but authors were and are desperate to publish their studies, and they saw no alternative to giving away their works to a publisher who demanded all copyrights, despite doing nothing to improve the quality of the science.

Librarians were in a pickle too, with faculty (now the readers) clamoring for access and subscriptions to many expensive journals, and publishers escalating prices at several fold more than the rate of inflation. The publishers had everyone over a barrel, except their stockholders and management who demanded higher profits by increasing prices and reducing services. The librarians bought the high priced journal subscriptions up to the point that they lost most staff and their Deans and Provosts finally said no more dough. Nobody was happy, but the US taxpayer was taking it on the chin more than anybody by paying for the research and then paying again for access to it. This was the state of scientific publishing until the Open Access movement began. Open Access means that the US taxpayer directly, fully, and immediately benefits from publicly funded research as much as possible. Anything else is suboptimal and wrong.

Mechanisms to enable full use in perpetuity (i.e. distribution, reuse, text mining, computation, etc.) should be part of a government-wide public access policy.

The public also needs full use of these articles sooner than the current term of copyright allows. This can be accomplished by implementing appropriate

licenses--such as Creative Commons CC-BY licenses, which are enforceable under current copyright law.

A policy that results in a “read-only” database must be avoided.

Comment 3. What are the pros and cons of centralized

The federal government is the best 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 usable, and are permanently preserved.

Any public-access policies that are developed must give the federal government adequate rights to archive and distribute publicly funded articles.

A federal public-access policy could involve multiple repositories as long as repositories support access and use conditions that allow all interested parties to build on them.

Simply providing the government with a copy to put in a "dark archive" is silly; without regular access/use, archival veracity cannot be ensured.

Current market attempts at archives are not adequate. For the government to maintain an archive is not a duplication of effort, but a necessity to ensure that the public's investment in research is protected and leveraged.

Comment 4. Are there models or new ideas for public-private partnerships

The best examples are PubMed and PubMed Central. These are analogous to the relationships of NCBI with the entire scientific community who routinely upload DNA, RNA, amino acid sequence data, and 3D structural information to the GenBank database and the like. NCBI shares data worldwide with several databases and vice versa. These sharing protocols work extremely well, and each serves as a backup for the others.

A great example of the burst of utility was in the migration of Medline into PubMed. There was a huge surge of activity and utility once the end-users (the academic and scholarly faculties) were able to search PubMed directly on the web, unfettered by intermediaries or nickel and dime charges. We have seen bursts of increased utility and value now that full articles are available unfettered online.

The major research universities and their libraries have extensive experience and

existing archive infrastructure, and should be actively encouraged to partner with federal agencies.

Another example: the HathiTrust

Comment 5. What steps can be taken....to encourage interoperable

Metadata should be viewed as a means for enabling specific actions, rather than simply item description. It should facilitate use, reuse, and analysis of published works.

The NLM DTDs and Schemas are well developed and ought to form a basis for future standards. XML valid against these DTDs needs to be thought of as the original, definitive, archival, and permanent source of an article.

Each publication and its supporting data need to be linked in thoughtful ways so that analysis of published text as “data objects” can be correctly supported.

Comment 6. How can Federal agencies that fund science maximize

Public access policies should ensure consistency of requirements. Broadly adopt the NLM DTDs and Schemas. IE, to reduce the complexity and cost, and increase the rate of compliance, use the NLM journal publishing DTD model.

Policies can also create opportunities to create/enhance productivity management tools, such as:

The creation and enhancement of CV's, bibliographies, and Principal Investigator (PI) profiles. Advanced analytics might help a PI decide who they ought to work with more.

Opportunity for universities to better measure research output.

Comment 7. Besides scholarly journal articles.....

Research book chapters, monographs, reports, texts, and conference proceedings that result from publicly funded research should be made openly and immediately accessible to the public just like publication of research articles in journals.

However, if authors write about work other than taxpayer funded research, and

are written outside the scope of research duties (c.f., educational materials, textbooks, or book chapters for instruction or entertainment) and paid by a commercial press, these are different circumstances, and policies should reflect these differences.

Comment 8. What is the appropriate embargo period

Immediate access is the ideal time to maximize scientific and commercial utility of information contained in these articles. Anything that blocks or slows access means a loss of opportunities, reduced productivity, job losses, and so on.

No data have been provided by any publisher that a short embargo or immediate access has harmed them.

At least one biomedical journal has published completely open and immediate access online without charging subscription fees or any author fees AT ALL. This journal is fully peer reviewed and edited. It published 380 primary science articles in 2011. It is run and managed strictly by volunteers (academic scientists). It serves a substantial fraction of the publication needs for ophthalmology and vision sciences researchers. This journal has been running highly successfully for 17 years. It proves the principle that any journal can always be fully and immediately open access. This is an absolutely crystal clear case establishing that no author fees or subscriptions are necessary to be successful and respected long-term. See www.molvis.org.

Best regards,
John Nickerson

John M. Nickerson, PhD
Professor
Ophthalmology Department
Emory University
B5602
1365B Clifton Road, NE
Atlanta, GA 30322

This e-mail message (including any attachments) is for the sole use of the intended recipient(s) and may contain confidential and privileged information. If the reader of this message is not the intended

recipient, you are hereby notified that any dissemination, distribution or copying of this message (including any attachments) is strictly prohibited.

If you have received this message in error, please contact the sender by reply e-mail message and destroy all copies of the original message (including attachments).

**Subject: RFI - Public Access to Peer Reviewed Scholarly Publications
Resulting from Federally Funded Research**

Date: December 29, 2011 10:46:02 AM EST

1) Expanding access to publications resulting from federally funded research seems a positive step towards possible advancement across science. Certainly, protecting intellectual property and copyright should be a priority. However, currently information is shared across a variety of academic fields through the use of electronic databases housed by University libraries or professional organizations and paid for by users. Researchers, academics, and practitioners in science understand the guidelines for respect of intellectual property, and universities and other public entities currently involved in the dissemination and storage of such information are well-practiced in how to protect such works. Steps that could be taken to broaden access include expanding the ability of universities and other public institutions to disseminate this information to the wider public - for example, there could be the creation of an electronic database (such as ProQuest, currently used widely for the dissemination of doctoral dissertations and theses) that would provide access to federally funded research similarly to the way that certain databases cater to a particular scientific discipline, but without the user fee. This database could be housed and maintained by universities similarly to how other databases are currently managed, but without the restrictions currently placed on pay-for-use databases. The wider availability of this information should encourage increased productivity by researchers in both the public and private sectors, and thus, would fuel economic growth due to innovations and advancements in the field. There would also be the opportunity for employment of librarians, computer programmers, and other IT professionals to maintain the database.

2) There should be a way to register with the database - for example, giving an email with one's name and address as well as affiliation, and perhaps there could be a way to track who downloaded and/or accessed what content. There also could be some kind of user agreement that must be affirmed prior to accessing the database so that it is understood that intellectual property must be respected. Certainly, there could also be embargo procedures for sensitive information similar to what are currently in place.

3) Decentralized approaches (such as via universities or other agencies)

could be an advantage, as there would be a way to distribute the workload and compensation that could be gained by maintaining the databases, and databases could be specialized to particular disciplines rather than managed by one huge agency without the expertise of the discipline. What about a combination of a centralized agency with whom all work is filed, but within which various disciplines are maintained by more decentralized agencies?

4) There are a vast variety of databases currently in use and managed by both private and public agencies in almost every scientific discipline. Encouraging communication and coordination among these groups seems key in putting this initiative into place. One partnership with which I'm familiar is the What Works Clearinghouse, which provides access to scholarly research around effective instructional practices and interventions. Public and private school systems and educators can access this work. Perhaps this agency could serve as a model or provide input into what a comprehensive database would comprise.

6) Perhaps only parts of the article (such as the abstract or an executive summary) or a more public-friendly document could be submitted as part of the submission for a more traditional journal article, that then allows for easier consumption by the public, and may omit some of the more tedious, technical, or sensitive content. To minimize burden on stakeholders, it would seem to be important to streamline the process of submission. Perhaps, as part of submitting to a journal, it is simply one additional step to submit the abstract, executive summary, etc, to the federal database?

7) Not necessarily. Book chapters are already available through libraries. Conference proceedings often lead into scholarly work that creates meaningful output in the form of journal articles. For right now, the focus should stay on journal articles as these are the most challenging to access because of the need to pay for access to such journals.

Jessica Oeth Schuttler, Ed.S.
Pre-Doctoral Psychology Fellow
Center for Child Health and Development
KU Med Center

Subject: public comment on access to publications from government funded research

Date: December 29, 2011 4:37:58 PM EST

Comment 1 Openly accessible federally funded scientific research broadens the ability for everyone involved in the scientific enterprise to build upon emerging thinking rather than limiting access to those scientists working for organizations with the means to pay for access to current information. This levels the playing field enabling start-ups and smaller scientific organization to contribute new products and grow the economy with the reinforcing effect of improving productivity by facilitating the exchange of information. A fully informed scientific community is then able to work faster and more efficiently to bring new products to market and more quickly. This can be achieved at a relatively low cost to benefit ratio given the infrastructure already in place at NIH with PubMedCentral. In the field of academic research, for example, immediate access is currently limited to those resources an institution can afford and choices are constantly made between supporting one academic discipline or another. Open access solves that problem by opening up information from all fields of research supported by the federal government and encourages interdisciplinary research by making resources equally available regardless of discipline.

Comment 2 At the most basic level, federally-funded scientific research belongs to the American public. The current system has been perverted in a way that enables commercial publishers to own the copyright to work done by federally-funded academic scientific researchers. Universities then have to purchase the work done by their own researchers that was funded by taxpayers. Universities have to convince their researchers to insist that their peer-reviewed work belongs to them and not the publisher of their work. Public access policies don't undermine intellectual property rights if they are carried out within the existing copyright framework.

Comment 3 – It makes a lot of sense for the federal government to create one approach for managing publicly access to peer-reviewed scholarly publications that result from federally-funded research. It reduces costs by “solving the problem” once and the federal government is the most stable institution in terms of ensuring that the research it has funded remains permanently accessible and usable. Even if the government required a private entity to ensure permanent access in reality commercial entities go out of business and the government would be left to “solve the problem” again.

Comment 4 – Public-private partnerships might be used to encourage innovation, but it should be clear that ultimately the contents of the archive are government property is the research was federally funded. Archives such as LOCKSS and Portico developed by non-profits are a more likely partner in archiving.

Comment 5- The metadata standards in effect at the time should be utilized to encourage interoperable search, discovery, and analysis capacity across disciplines and archives. The public should be able to access publications using commonly practiced search techniques and access points. Federal agencies should require authors to include acceptable metadata at the same time their work is deposited in the archive.

Comment 6- Burden and cost can be minimized by developing uniform standards and requirements across all federal funding agencies. Benefits will also be maximized by providing the public uniform searching and access tools to facilitate use of the peer-reviewed literature.

Comment 7 - Book chapters and conference proceedings that derive from federally-funded projects should be made publicly accessible.

Comment 8 - Philosophically, because federally funded research

belongs to the public, there should not be an embargo period. In the interests of those journal publishers who rely on subscription income, either a six or twelve month embargo seems to allow those publishers to maintain enough of their subscription base. I am not aware that any bioscience journal publishers have gone out of business since NIH instituted its public access policy, for example. Different embargo policies across different agencies don't make sense unless faster public access to information in a particular field can be demonstrated to stimulate the economy as a result. In our case, our journal inflation in FY 2011 averaged 5 % and our collections budget was reduced 1.5% and our inflationary increase was suspended. This translates to our faculty having access to about 6.5% less federally-funded literature in fields where open access is not a condition of the award.

Brinley Franklin

*Brinley Franklin
Vice Provost
University of Connecticut Libraries
369 Fairfield Way
Storrs, CT 06269-2005*

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

John Wiley & Sons (Wiley) 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 appreciate the opportunity to participate in the Administration's consultation with stakeholders in the scientific research enterprise.

Founded in 1807, Wiley is North America's oldest independent publisher, and has a distinguished history as a literary, scientific, technical, medical, and scholarly publisher, serving researchers and practitioners in the US and around the world. Today, we employ approximately 2,600 staff in the United States and 5,300 globally. We are one of the world's foremost academic and professional publishers. We publish over 1,500 scholarly peer-reviewed journals, and our online service Wiley Online Library (<http://onlinelibrary.wiley.com/>) provides electronic access to more than 5.5 million articles (as of December 2011) across these journals. Wiley-Blackwell, Wiley's scientific, technical, medical, and scholarly publishing group, is also the world's largest society publisher, working in partnership with over 800 learned and scholarly and professional societies which represent millions of members globally. These include the American Cancer Society (ACS), for which we publish *Cancer*, their flagship journal; the Sigma Theta Tau International Honor Society of Nursing, with more than 120,000 members; and the American Anthropological Association, for which we publish 23 journals. Many of the societies we and other publishers partner with depend to a significant extent on the revenues generated by publishers to support activities which benefit the communities those societies serve and the general societal good.

We publish subscription-based and open access journal titles. For instance, with the The American Heart Association (AHA) we recently announced an innovative venture to publish a new open access journal, *Journal of the American Heart Association*. This journal will serve as the first online-only open access journal for the AHA, and joins the AHA's prestigious portfolio of 11 peer-reviewed print and online subscription-based scientific journals.

As preamble to the RFI questions –

- The phrases "the results of research" and "peer reviewed scholarly publications" are used interchangeably in the RFI; they are not the same thing. We agree that the taxpayer should have full access to the results of federally-funded research. Any report which a federal (or any other) research funder chooses to commission from a grantee describing the results of the grantee's research and adds to a publicly-accessible database would be useful in promoting access to the results of research. Neither publishers, nor the learned societies for whom we publish, own the results of research or claim any ownership of any article until it is submitted to us for publication, when we begin to add value through peer review and other activities in the publication process. What we copyright and own is the value-added article; the article becomes our work product. Expropriation of these value-added articles by the government without compensation undermines copyright, intellectual property rights, jobs, and exports. Funding agencies receive regular reports from grantees that can be made publicly available, as the government deems appropriate, likely at minimal additional cost. Why would our government choose to take a publisher's or a society's work product, rather than commission a separate report from grantees? Because it recognizes the value of the publisher value-add but does not want to pay for it. Government funds much scientific research but that does not entitle it to access to and control of the journal articles arising from this research. Taxpayers may fund the research, but they do not fund the publication of this research and, therefore, should have no expectation of receiving free access to this material.

- The RFI also uses the phrase “long-term stewardship” repeatedly, implying that stewardship is an issue. Since the Royal Society launched the Philosophical Transactions in 1665, learned societies and journal publishers have been stewards of the literature. Indeed, since journal content delivery via the web was initiated in the mid-1990s, many publishers have invested not only in the technology necessary to deliver this content online, but also in the digital recovery of print material, often back to the very first issue of a title. This is true for all titles published by Wiley-Blackwell. By one estimate publishers have recovered 40 million articles previously available only in print and made them available online.
- The combination of investments in digital and online technology (by publishers as well as others), and the formation of library consortia (assisted by publishers in many cases) around the country and the world, has accelerated and broadened access to the peer-reviewed literature by orders of magnitude. There is more access to more content by more users now than ever before. Publisher innovation and investment over the past 15 years has made this possible and ensured “broad public access”. Wiley Online Library users now have access to 5.5 million journal articles. The platform will be visited by an estimated 100 million unique visitors in 2011, who will download 200 million articles; growth in usage and citations continues to be explosive. Usage of Wiley-Blackwell titles will grow, we project, by 60% in 2011 compared to 2010. Moreover, data also shows that, while academics rank information as 5th out of 15 factors important to their success, access to this information is ranked 13th out of 16 factors as a barrier to success. Most academics and researchers access the necessary literature via subscriptions or licenses maintained by their institutions. ([Access by UK Small and Medium-sized Enterprises to Professional and Academic Information](#), Mark Ware Consulting Ltd for Publishing Research Consortium – April 2009).
- Many research funders require research progress reports on all grants. Expanding this information by requiring the addition of a one-paragraph lay summary, and making both freely available, has more potential to enhance public understanding than does providing free access to scientific journals. Our strong preference would be that the federal government does not mandate deposit of journal manuscripts in a 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.

These reports are produced as part of each federally funded project and they are delivered to the government as a contract deliverable, thus there is a report for virtually every project. The project itself underwent peer review before being selected for funding, and the research results being reported on are solely that which the government funded. In short, these reports are the federally funded research results. Thus if the policy is to provide public access to federally funded research results, then these reports are the natural vehicle for doing so. The government already has them, so all it has to do is make them publicly available. Several federal science agencies already do this; no new system is required.

A great deal of federally funded research is of an applied nature, as opposed to being basic research. For example, in the Department of Energy applied programs account for a large fraction of the research budget. In many cases, applied research results are not published as journal articles, just in the project research reports. Yet this applied research is often the most suitable for technology transfer into the private sector, which is one of the major goals of public access policy. A focus on journal articles misses much applied research.

In addition, negative results are seldom published as journal articles, but are often useful and are described in research reports.

- Serious errors in manuscripts are frequently corrected after the peer review process. For example, in 2010 alone, we issued more than 1,000 corrigenda, errata, statements of concern and retractions with respect to published articles. We are extremely concerned that ~~using~~ making available any version other than the true “final” one (the “Version of Record”) will cause confusion, at a minimum, and could significantly compromise the scientific record. We take seriously our role as the stewards of the research literature and version control is an important component of that role.
- Mandating a single approach to public access will stifle innovation in what is now a rapidly changing environment, both by decreasing the amount that publishers are able to invest and by reducing our – and our society partners’ - incentive to develop new tools, delivery vehicles, and functionality.

In summary, publishing is a business underpinned by the copyright laws of the US and almost all other countries. However, in the vocabulary of many current anti-copyright activists, “public” is being conflated with “free.” There is already a robust public access model for the dissemination of the peer-reviewed results of taxpayer (and other) funded research – the global journal corpus. Agencies dispensing funds to support taxpayer-funded research may wish to collect and publish free of charge reports generated by the recipients of those funds. However, these agencies have no rights to the research articles written for and published by journals, nor is such a claim justified by any notional absence of access. There is no evidence that making the current broad public access to the journal literature free will improve research productivity or the public wealth. On the contrary, free access, like copyright piracy, is 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?

Growing existing and new markets is what private sector publishers do, both for our own journals and for those we publish on behalf of our society partners.

If “market” is used in the commercial sense, publishers are growing new markets around the world as developing economies industrialize. Making publishers’ and societies’ content freely available is likely to stifle those opportunities, as customers choose to access free versions of journal articles rather than pay for the Version of Record. As an example, having established a growing business in China, after decades of content pirating, we now see two thirds of PubMedCentral (PMC – the repository for mandated NIH grantees deposit) usage originating outside the US, much of it we assume from China. We also know that there are businesses in China reselling PMC content which, although made freely available from PMC, still carries a publisher’s or a society’s copyright. How do US taxpayers benefit from the transfer of intellectual property such as this, owned by US corporations and not-for-profit societies, to businesses and governments abroad?

There are no studies that support the notion that free access to the research literature will increase research productivity or economic growth. The fallacy is 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 huge investments; access to the research literature is not a constraint.

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 someone has to pay for it. The government pays only for the 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 or warranted. 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.

(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 could:

- Make funds available for the purchase of open access to published articles. Several research funders already do this (Howard Hughes Medical Institute, The Wellcome Trust, Max-Planck Institutes). These costs are a small fraction of the investment in the research itself.
- License content from publishers and learned societies and make it available to specific audiences. We license content to customers of many kinds, including government agencies, and would be pleased to enter into negotiations regarding access by specific communities to packages of content.
- Make the funder-collected and maintained outputs of taxpayer-funded research, including grant reports or research progress reports, freely available to the public. Work with private sector publishers to make that content discoverable and link it to the journal literature.

What it should not do is to take accepted or published articles from publishers or learned societies (directly or via a mandate placed on grantees) and make them freely 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?

See the Preamble for our response to “long term stewardship”.

Scholarly journal articles have been published for several hundred years by a combination of society, not-for-profit, and for-profit publishers. Together, we have provided access to the literature for scholars and researchers. A multiplicity of publishers has not prevented broad public access. In fact, one could argue that it has been an advantage in promoting competition which has, in recent years, driven development of increasingly sophisticated platforms to deliver this content.

Publishers over the past decade have developed the Digital Object Identifier (DOI) 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, including Wiley, in 2002, 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.

Is the government really a credible provider of these kinds of services? Given government budget constraints why would the government consider using taxpayer dollars to duplicate an existing, well-functioning service? PubMed Central, the repository for mandated NIH grantees, 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?

(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.

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. The technical details of implementation will be worked out through consultations with CrossRef and appropriate publisher, society, and agency representatives. Our goal is to launch this feature in 2012.

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 (enabled

through the DOI) to the publisher's platform, where article abstracts are freely available and the Version of Record (VoR) (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. We would be 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 could 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 (ORCID) project (www.orcid.org). ORCID is a newly established non-profit organization whose goal is 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. Publishers including Wiley have contributed heavily to ORCID's initial funding.

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 (OSI). ORCID is governed by representatives from a broad cross-section of stakeholders including publishers, societies, libraries, and other institutions (see <http://orcid.org/board-of-directors>).

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. We propose to 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.

The Publishing Research Consortium recently completed a study on article-level content mining based on a broad survey of ongoing or planned activities among nearly 30 STM publishers or associations, <http://www.publishingresearch.net/documents/PRCSmitJAMreport20June2011VersionofRecord.pdf>

Linking to/from research reports

We propose a collaboration 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 us 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?

See (4) above under 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?

Publishers would argue that those who can benefit from access to the peer-reviewed journal literature already have access; researcher surveys bear out this assertion. We would also argue that we own and have copyrighted journal articles which are published in the journal titles we publish.

Publishers and learned societies are committed to the wide dissemination of our 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?

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 free access to that content is simply an expropriation of that content. The Federal Government might as well demand free access to *Time* or *Newsweek*.

(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 are no "appropriate" embargo periods. Any embargo period is a dramatic shortening of the period of copyright protection afforded all publishers.

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, Wiley believes that publishers – and learned societies – themselves should 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 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.

By Electronic Submission to:
publicaccess@ostp.gov

Ted Wackler, Chief of Staff
Office of Science and Technology Policy
Executive Office of the President
725 17th Street Room 5228
Washington, DC 20502
Phone: (202) 456-7116
Fax: (202) 456-6021

December 29, 2011

Dear Mr. Wackler:

The undersigned researchers of the Berkeley Digital Library Copyright Project¹ respectfully submit these comments in response to the Office of Science and Technology Policy's Requests for Information regarding public access to peer reviewed scholarly publications from federally funded scientific research ("RFI"), dated November 4, 2011.²

The Berkeley Digital Library Project is a grant-funded project that aims to investigate copyright obstacles faced by libraries and other like-minded organizations in their efforts to realize the full potential of present and future digital library initiatives. Our efforts are concentrated on both the obstacles themselves and the range of possible legal, technological, social, and market-based solutions to overcome them. Among other issues, we are specifically examining challenges with respect to orphan works, library privileges, digital lending, metadata ownership, licensing, and the ways that those issues impact the creation of a national digital public library. Principal investigators for the project are Berkeley Law professors Pamela Samuelson, Jason Schultz, and Jennifer Urban. David Hansen is the project's full-time research fellow.

The comments below explain the impact that a federal access policy would have on the creation of a national digital public library, and the broader legal issues that should be considered to ensure that such a policy aligns—or at a minimum, does not conflict with—a national digital public library initiative.

¹ UC Berkeley affiliation is provided for identification purposes only.

² 76 Fed. Reg. 68518 (Nov. 4, 2011).



I. A National Digital Public Library and Aggregated Access

Serious efforts are underway to develop a national digital public library. In October 2011 the Digital Public Library of America (“DPLA”) launched, from the National Archives in Washington, D.C., efforts to take “concrete steps toward the realization of a large-scale digital public library that will make the cultural and scientific record available to all.”³ That project has the ambitious goal of bringing together and further building upon the digital resources already made available by the likes of the Library of Congress, the HathiTrust, the Internet Archive, and the numerous universities and research organizations that currently maintain institutional repositories of digital copies of works produced from scientific research. These efforts are designed to provide access within the United States on a level comparable to that enabled by initiatives in other parts of the world, such as in the European Union (“EU”), which is already well on its way to developing an EU-wide digital public library (“Europeana”), that currently provides access to millions of items that represent Europe’s cultural and scientific record.⁴

In essence, the DPLA and related projects aim to provide widespread aggregated access to our nation’s creative and scientific output. Because federal funding plays a large role in the creation of that output, it is important to consider how a federally funded access policy might fit into an aggregated system of national digital access. In particular, because digital libraries are important in both the creation and dissemination of federally funded scientific research, they are stakeholders who have an intense interest in the intellectual property structure of such a policy. The comments below outline principles that should be considered in creating that policy. These comments are responsive to questions (2) and (3) of the RFI regarding:

- (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?; and
- (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?

II. Intellectual Property Principles For National Digital Access

A federal open access policy for peer reviewed scholarly publications would greatly further the movements described above toward widespread national digital access. The benefits of such access are myriad, and can be realized without jeopardizing the system of incentives upon which the current copyright system operates. These benefits and the ways that they complement (and not harm) the interests of scientists and researchers are well explained in the comments submitted by others (e.g., the many university libraries also replying to this RFI). Although almost any policy that increases online access to publications resulting from federally-funded research would be beneficial, a few principles can be discerned to extract the greatest value from such a policy by enabling dissemination through a national digital public library. These comments outline three principles in

³ *About*, DIGITAL PUBLIC LIBRARY OF AMERICA, <http://dp.la/about/> (last visited Dec. 8, 2011).

⁴ *See About Us*, EUROPEANA, <http://pro.europeana.eu/web/guest/about> (last visited Dec. 8, 2011).

particular—consistency, discoverability, and extensibility—that would focus the intellectual property framework of a federal policy toward a greater impact through aggregated online access via systems such as the DPLA.

a. Consistency

A federal policy should be consistent in terms of the intellectual property rights and licenses that it adopts. While such a policy need not espouse a rule of complete uniformity across all federally-funded projects (indeed, that may be unwise as incentives for creation differ among academic disciplines), it should adhere to a principle of consistency that will ensure that the rights in scholarly works produced from federal funding are subject to one of only a few licenses, and that those licenses remain consistent over time.

One of the greatest problems in aggregating digital access to current scholarly peer-reviewed publications is the wide range of rights models that exist among academic publishers. Indeed, one ongoing attempt to catalog the current rights situation by the University of Nottingham has resulted in a list of over 1000 different academic publisher contracts, each with varied terms which dictate the ways in which authors, their institutions, and others may use published works online.⁵ Those policies vary along several lines, including the version of the article that can be posted freely online, the embargo period before such a use may be made, and the types of subsequent uses to which the article may be put. At the academic institution level, this variety of rights has proved a challenge in developing institution-wide repositories for open access to research outputs.⁶ Developing a nationwide repository would only exacerbate that challenge, but a consistent federal policy could help simplify the situation considerably, and could allow projects like the DPLA to more efficiently mark items for inclusion in its digital collections.

In a similar vein, the terms upon which access to these works are made should be compatible, as far as is possible, with existing access-oriented licenses. CreativeCommons, for example, hosts a suite of oft-used licensing models—many of them applied under academic institutional access policies—which are suitable for many academic works.⁷ Adopting consistent license terms that are compatible with more broadly used language like these licenses would allow content aggregators (such as the DPLA) to more easily parse content for inclusion in its collections, to inform patrons of any use restrictions, and would allow future researchers and creators of derivative works to more easily make use of works produced under a federal open access policy.

⁵ See *Frequently Asked Questions*, SHERPA/ROMEO, <http://www.sherpa.ac.uk/romeo/faq.php?fidnum=|&mode=simple&la=en> (last visited Dec. 9, 2011).

⁶ This problem has existed for some time. For a review of these challenges from the University of Kansas—one of the first U.S. institutions to tackle this problem—see Holly Mercer & Ada Emmett, *RoMEO Green project at the University of Kansas: an experiment to encourage interest and participation among faculty and jumpstart populating the KU ScholarWorks Repository*, in *PROCEEDINGS OF THE 68TH ANNUAL MEETING OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY (ASIST)*, NEW ORLEANS, at 1433–41 (2005), available at <http://hdl.handle.net/1808/873>.

⁷ See *About the Licenses*, CREATIVECOMMONS, <http://creativecommons.org/licenses/> (last visited Dec. 20, 2011).

b. Discoverability

A federal access policy should also ensure that claimed intellectual property rights and related licenses that cover federally funded works are easily discoverable, both on the face of the article produced and in its machine-readable metadata. This principle compliments the consistency principle noted above, as consistency in licensing is of little use if the works covered by those various licenses are not easily discernible.

An access policy which mandates the creation and attachment of consistent intellectual property metadata would go a long way toward addressing some of the most vexing intellectual property issues that have confronted digital libraries over the past decade. The problem of digital access to works that are orphaned—that is, works whose owners cannot be located, and who are therefore unable to authorize online access—has troubled the Copyright Office, Congress, and a range of digital library initiatives as they strive to provide increased access to works whose copyright status is uncertain.⁸ A proactive federal access policy that preserves copyright information would prevent a similar problem from arising with respect to federally funded scholarly research publications.

For the maintenance and discoverability of this data on a larger scale, a centralized registry of metadata—including copyright and license status—on federally funded works covered by the access policy would be beneficial. Such a registry would ease the burden of libraries and users seeking to user particular works or even large groupings of works together by providing a single search point for rights analysis. An ideal system would also host, preserve, and enable access to these works, though it would be unwise to reserve that role exclusively for a federal repository, as many other digital repositories and digital libraries are willing and experienced with doing the same.

c. Extensibility

Finally, the metadata, digital copies of articles, and associated data that are produced under a federal access policy should be structured so that their use is extensible beyond the initial repository or registry in which they deposited. Doing so would preserve the essentially public nature of these works. A federal access policy that practically restricts holding of research outputs to one agency or one private entity stifles the potential for creative new access solutions, and risks monopolizing the very works that an access policy would try to make available.

Following the principle of extensibility requires that works be made accessible and useable in formats that comport with open information data standards (e.g., linked open data).⁹ Extensibility also requires that a federal policy allow outside entities to host the materials in addition to any central repository. Although a centralized depository of these materials may be an important feature promoting access—useful for, for example, consistency over time, usage tracking and record maintenance—it is axiomatic that more, distributed copies ensure preservation better than a system which maintains only one centralized copy.¹⁰

⁸ See David Hansen, *Orphan Works: Definitional Issues* (Berkeley Digital Library Copyright Project, White Paper No. 1, 2011), <http://ssrn.com/abstract=1974614>.

⁹ See, e.g., Tim Berners-Lee, *Linked Data*, W3.ORG, <http://www.w3.org/DesignIssues/LinkedData.html> (explaining the basic structure and benefits of linked open data for connecting information on the worldwide web).

¹⁰ See, e.g., Stanford University Libraries, LOCKSS Program, <http://www.lockss.org/lockss/Home> (last visited December 9, 2011) (“LOCKSS (Lots of Copies Keep Stuff Safe), based at Stanford University

Most importantly, extensibility requires that licenses to the covered works be tied not to the particular entity hosting the works, but rather to the ultimate use of the work, which is available to all hosting entities (and users). License terms should also be permissive in the uses they allow, being careful to expand and not contract the uses to which libraries and end users may put the covered works. Existing exceptions to copyright, such as fair use or the library exceptions codified in Section 108 of the Copyright Act should not be limited by license terms. Contractual limitations on these exceptions have frustrated library efforts to increase online access for their patrons in the past,¹¹ and such limitations could similarly bind federally funded research in a way that prevents useful access and that curtails future productive uses of these works. These are precisely the type of activities that existing copyright exceptions (and in particular, fair use) are designed to foster.

Adopting a principle of extensibility would enable national digital access through an entity like the DPLA, but would also serve an important role beyond that initiative. While the DPLA currently rests on today's horizon, untold projects to improve access and usability of scientific research remain foreseeable for years to come. A federal access policy designed on a principle of extensibility would enable those, as well as current, access solution projects.

An access policy for peer-reviewed scholarly publications resulting from federally funded research that is built upon the principles outlined above would increase the reach and impact of federal research dollars. In particular, dissemination through channels such as the DPLA could maximize the impact of a federal policy by integrating that content into a platform designed to reach a broader segment of the American public.

Respectfully,



David Hansen
Digital Library Fellow

On behalf of David Hansen, Pamela Samuelson, Jason Schultz, and Jennifer Urban.

Libraries, is an international community initiative that provides libraries with digital preservation tools and support so that they can easily and inexpensively collect and preserve their own copies of authorized e-content.”).

¹¹ See, e.g., Stanley Wilder, *The Erosion of Fair Use Protections for Digital Scholarship*, CHARLESTON ADVISOR 57, 57–58 (2011), http://www.carli.illinois.edu/mem-serv/mem-train/1104ereslicensing/Erosion_of_fair_use.pdf.

Subject: Public access to federally funded research

Date: December 30, 2011 8:50:06 AM EST

To whom it may concern,

With respect to "ensuring long-term stewardship and broad public access to peer-reviewed scholarly publications that result from federally funded scientific research," we and other college libraries, representing the interests of our users, academic researchers, believe that unregulated, commercial publishers have done a disservice to the scholarly community in recent years. These publishers' unreasonably high book and subscription pricing have already severely limited access for scientists and others at all but the largest and richest institutions, with obvious, unfortunate consequences for academic practice in particular and the public welfare more broadly. Inasmuch as these publishers (Elsevier, Wiley, Springer, etc.) do rely heavily on federally subsidized research for their content, it seems only fair that they be required to enable broader access through reduced subscription prices and/or that alternative, cheaper sorts of publishing venues be encouraged by federal policy.

Let me briefly outline one less-recognized implication for libraries of this recent development in publishing. Even at a relatively wealthy institution like mine, Amherst College, our library holdings, and so our users' access to information, have been increasingly limited by the unconscionably huge price rises demanded year after year by commercial scientific publishers; journal pricing has been the most burdensome, but book pricing in the sciences is also out of line with publishing in other areas. We have had to pick and choose what we can afford to make available, and as we necessarily devote larger percentages of our budgets to purchase access to material our scientists consider critical for their research and teaching - which is to say scientific journal literature largely - we're shifting resources that heretofore supported other disciplines, in the humanities and social sciences, thereby shrinking access for scholars in those fields as well. In 1990, the Amherst College Library spent around 50% of our annual materials budget on serials, with approximately 60% of *that* going to the sciences; currently, we have to devote 70% of our total annual expenditures to journals, mostly in the sciences. So, not only are publishers unfairly abridging scholars' and the general public's entitlement to access federally-funded research, they've upset the delicate multidisciplinary ecology of academia, and the damage they're doing will only increase if the government doesn't step in.

Yours respectfully,

Michael Kasper

Collection Development Coordinator



Submitted Electronically to: publicaccess@ostp.gov

December 28, 2011

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

Dear Sir or Madam:

The American Society of Hematology (ASH) appreciates this opportunity to respond to the Office of Science and Technology Policy's November 3, 2011 Request for Information on Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research and share its recommendations on approaches for ensuring broad public access to peer-reviewed scholarly publications that emanate from federally funded scientific research.

ASH represents over 16,000 scientists and clinicians committed to the study and treatment of blood and blood-related diseases. These diseases include malignant hematologic disorders such as leukemia and lymphoma, non-malignant conditions including anemia and hemophilia, and congenital disorders such as sickle cell anemia and thalassemia. ASH members are active participants in federal programs, recipients of federal grants, and contributors to the federal government's research accomplishments. The Society publishes the premier scientific journal in hematology, *Blood*, and is committed to a collaborative relationship with the government to assure that important research findings are published and disseminated by print and electronic means to the public through rigorous independent peer review.

ASH fully supports the goal of broad public access to research publications. In fact, ASH supports free access to *Blood* on the broadest possible basis. Although ASH cannot adopt or support a publishing model that is not economically sustainable over the long run, certain sections of the journal are always free on-line: abstracts and tables of contents, *Inside Blood* commentaries, "How I treat" articles, and five research articles in every issue. *Blood* maintains a 12-month embargo for current articles, but content older than 12 months is free to all on-line. In addition, ASH and many other not-for-profit publishers allow free immediate access to selected articles with important public health or clinical significance and distribute free articles to scientists working in many developing nations. As a result, more scientific papers are available now to more people than at any time in history.

While federal funds may support – in whole or in part – the research reported in journal articles, it is extremely important to realize that the federal government does not pay for the very important processes that lead to the publication of that research. ASH and many other not-for-profit scientific societies provide important services that are necessary to ensure the publication of accurate scientific information: peer review, copyediting, formatting, printing for distribution, and publishing on-line. These services represent a substantial private sector investment that results in prompt access to research results and the reliable archiving of articles at no additional cost to the public. Mandating a specific

2012

President

Armand Keating, MD
Princess Margaret Hospital
610 University Avenue, Suite 5-303
Toronto, ON M5G 2M9
CANADA
phone 416-946-4595
fax 416-946-4530
armand.keating@uhn.on.ca

President-Elect

Janis L. Abkowitz, MD
University of Washington
Box 357710
Seattle, WA 98195-0001
phone 206-685-7877
fax 206-543-3560
janabl@u.washington.edu

Vice President

Linda J. Burns, MD
Division of Hematology, Oncology,
and Transplantation
420 Delaware Street, SE
MMC 480/Room 14-154A Moos Tower
Minneapolis, MN 55455-0341
phone 612-624-8144
fax 612-625-9988
burns019@umn.edu

Secretary

Charles S. Abrams, MD
University of Pennsylvania
School of Medicine
421 Curie Boulevard, #912
Philadelphia, PA 19104-6140
phone 215-573-3288
fax 215-573-7400
abrams@mail.med.upenn.edu

Treasurer

Richard A. Larson, MD
University of Chicago
5841 S. Maryland Avenue, MC-2115
Chicago, IL 60637-1470
phone 773-702-6783
fax 773-702-3002
rlarson@medicine.bsd.uchicago.edu

Councillors

Kenneth Anderson, MD
David Bodine, PhD
Michael A. Caligiuri, MD
Joseph Connors, MD
Marilyn Telen, MD
Alexis Thompson, MD, MPH
David Williams, MD
John Winkelmann, MD

Editors-In-Chief

Cynthia Dunbar, MD, *Blood*
Charles Parker, MD, *The Hematologist*

Executive Director

Martha L. Liggett, Esq.
mliggett@hematology.org

time for public release of manuscripts could be detrimental to not-for-profit scientific societies like ASH and jeopardize the crucial processes that are necessary to ensure that publications provide accurate scientific information.

There are several critical questions that must be addressed when Federal agencies consider developing a public access policy. These include: What will be the impact on scientific journal business models? What will be the impact on peer review? What will be the impact of expanded access on federal influence on research? What will be the impact of expanded access on federal funding of research? How will authors pay for expanded access and how will a public access policy impact limited research dollars? Below please find responses to several of the specific questions posed in the November 3 Request for Information that are relevant to the ASH membership:

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 productivity of the American scientific enterprise?

This question appears to focus on the role of federal agencies and public access policies in helping economic growth and improving productivity of the American scientific enterprise. ASH calls your attention to a November 10, 2011 inquiry to National Institutes of Health (NIH) Director Frances Collins by Representative Joseph R Pitts who chairs the Subcommittee on Health of the Energy and Commerce Committee in the U.S. House of Representatives. Chairman Pitts raises several similar questions. The Energy and Commerce Committee is engaged in a broad review of agencies under its jurisdiction to prevent federal regulation from undermining the ability of the private sector to create or sustain jobs. The Committee is concerned that the cost of regulation not exceed any net U.S. public benefit and that regulatory policy support strong economic recovery and boosting of job creation. Consequently, the Committee is examining the impact of the NIH Public Access Policy and PubMed Central database (PMC) on the science, technology and medical publishing field. The Committee is specifically concerned about the impact of the NIH Public Access Policy on the competitiveness of the science, technology and medical publishing industry.

Regarding what type of access to publications should be required, ASH recommends that the development of any government public access policy be designed so as not to jeopardize the business model of the journal, if it is to maximize economic growth. It is critical that any public access policy implemented by federal agencies preserve the viability of peer review and ensure the integrity of the scientific record. Various journals currently use different strategies to recover the costs of these operations: some charge subscription or access fees to readers; some charge article processing fees to authors; some are subsidized by scholarly societies, research institutions, or funding agencies; and many use a hybrid model combining various funding streams in their business models. Even without a government mandate, many not-for-profit publishers provide free access to their journals either immediately upon publication or after some period. The specifics of the access policy vary according to how a journal recovers costs. Consequently, ASH believes no one solution will solve all problems and recommends that federal agencies work cooperatively with all stakeholders to address any specific issues.

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?

Regarding intellectual property protections, ASH believes existing intellectual property and patent protections are adequate and the Society does not have further recommendations. Regarding the question of what policies should not be adopted with respect to public access to peer-reviewed scholarly publication, ASH strongly recommends that federal agencies steer away from certain policies. ASH opposes policies that impose too short of an embargo period that hamper a subscription-based business model. In addition, the Society opposes policies that impact the integrity of the scientific record by providing access to something other than the final version of an article.

Question 4: Are there models 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 2006, ASH has participated in the PMC(NIH Portfolio) Archive program as an alternative to the NIH Public Access Policy. The NIH Portfolio program works as follows: Participating publishers submit to NIH the final version of NIH funded research articles upon publication. NIH has internal use only of the articles during the journal's embargo period, which can be no longer than 12 months. During the embargo period, NIH can link to the journal website to provide access to NIH-funded research articles. Following the embargo period, NIH can provide links to the journal and can also distribute the articles directly through its PMC website. While the NIH Portfolio program is not ideal from the publisher perspective because it has implementation costs, ASH strongly believes it provides a better alternative for *Blood* and *Blood* authors than the NIH Public Access Policy.

Advantages of the NIH Portfolio program include:

- NIH obtains 100 percent compliance in its Public Access Policy by participating journals because the journals submit to NIH the final version of all NIH funded research articles upon publication on behalf of their authors.
- Authors of participating journals do not have to submit their manuscripts to NIH through the NIH Public Access Policy, but are fully compliant because the participating journals submit for them.
- NIH also has the ability to create a stable archive of peer-reviewed research publications resulting from NIH-funded research and a secure searchable compendium of these peer reviewed research publications that NIH can use to manage research portfolios and set research priorities.
- The program protects the integrity of journal articles by allowing the journal to submit the final article.
- The program also maintains journal business models by protecting the embargo period and the peer review system.
- The program allows expanded free access of science to researchers and the public.

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?

ASH respectfully recommends that first federal agencies explore the question of access and the extent of any problem. As noted above, ASH believes no one solution will fit all problems and it would be best to work cooperatively with all stakeholders in addressing specific issues. Again, while not ideal from the publisher perspective, ASH strongly believes the NIH Portfolio program, which was developed collaboratively between medical society publishers and the NIH, could serve as one paradigm.

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 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?

ASH believes the embargo period needs to be suited to the needs of the stakeholders. The NIH Portfolio program allows for a 12-month embargo, which works for most not-for-profit publishers. This decision was made recognizing the important role journals play in the validation and dissemination of scientific information and that a shorter period would jeopardize the ability of the journals to sustain the peer review process should subscription revenues decline if the embargo period were reduced. However, different fields of science have different patterns of usage and citation. There appears to be no uniform optimal embargo period across all scientific disciplines. While a 12 month embargo might work for most journals in the research areas funded by the National Institutes of Health, it is unlikely that the same is true for research funded by other federal agencies.

Again, ASH appreciates the opportunity to submit these comments and the Society would be pleased to further discuss its public access policy. For more information, please contact ASH Senior Director of Government Relations, Practice and Scientific Affairs Mila Becker at mbecker@hematology.org or 202.776.0544.

Sincerely yours,



Armand Keating, MD
President

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

From
 Shirley Baker, Vice Chancellor for Scholarly Resources and Dean of University Libraries, Washington University in St. Louis

Washington University is a member of the Scholarly Publishing and Academic Resources Coalition (SPARC) and the Coalition of Open Access Policy Institutions (COAPI). Both who will be submitting extensive and heavily referenced responses. My comments will be less formal.

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 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:

“The Faculty of Washington University in St. Louis is committed to making its scholarship and creative works freely and easily available to the world community.” ([WUSTL Open Access Resolution](#)). Full, immediate open access to research articles drives scientific innovation and productivity, delivers more value than an environment in which access and use are restricted, and can be implemented in a cost-effective manner.

Type of Access:

- All articles resulting from publicly funded research should be made freely accessible and fully reusable without commercial restriction, with no embargo or the shortest possible embargo that can be tolerated by all interested parties.
- Public access policies can be successfully implemented within the current copyright framework through the use of appropriate licenses.
- Although the request here is for comments on federally funded scientific research, the same principles apply to humanities and social sciences research. Scholarship is increasingly interdisciplinary.

Benefits:

- Enabling broad/full reuse means that researchers will be able to find and build on the results of others, rather than repeating research that has already been done. This will extend the value of the taxpayers’ initial investment.
- There are over 25,000 peer-reviewed research journals across all scientific and scholarly fields worldwide. Most universities and their researchers can only afford access to a small and shrinking fraction of these journals. Open access to research from publicly-funded work will solve some of the barriers to access.
- Allows an increasingly number of diverse users to stay current on cutting-edge research.

- The use of new tools such as machine reading also opens up new research pathways, making new connections possible and allowing users to identify relevant articles, access them, and read more information faster.
- Encourages contributions by “unforeseen participants” – expanding the potential for innovation and interdisciplinary application of the research.
- Promotes commercialization of discoveries in the form of new products and services.
- Open access to the peer-reviewed scientific literature allows scientists to incorporate new findings into their thinking faster, thereby increasing the productivity and innovation of the scientific enterprise.
- Open access maximizes research impact and progress and thereby maximizes the return on U.S. taxpayers investment in research. See: Bibliography of Findings on the Open Access Impact Advantage <http://opcit.eprints.org/oacitation-biblio.html>
- May increase government agency accountability and support informed, transparent, and evidence-based budget and policy decision-making.
- Provides congressional budget drafters, appropriators, and authorizers with better information to assess the value of existing expenditures and better target funding.
- Provides direct access to federal, state and local government representatives who might need the research results themselves to more effectively accomplish their work.
- Provides the public direct access to federally funded research when the need arises and when there are no alternatives.

Costs:

- A public access policy would be a cost-corrective modification to the current system where the public pays for the research, then does not have access to it without paying again. The relative cost would be reduced (and an imbalance rebalanced) from a broader public access policy.
- Since the NIH public access policy, Washington University Libraries has cut many journal subscriptions due to budget and use considerations, but none because all or part of the journal was available in PubMed Central, with or without embargo. No data has been provided by any publisher that the public access policy (with 12 month embargo currently in use by NIH and numerous other funders around the world) has harmed them. Since embargo periods are more protective to publishers than the lack of such embargoes, it can be surmised that publishers have not been harmed. There may be costs to publishers but that cost would be minimal.
- The actual costs of permanent, interoperable, fully searchable storage are not really known. This depends on what mechanism(s) are used (see Comment 3).
 - The benefits of an open-access policy similar to the NIH policy are estimated to be approximately 8 times greater than the costs. Cost/benefit issues are summarized in Houghton, J.W. & Oppenheim, C. (2010) The Economic Implications of Alternative Publishing Models. Prometheus 26(1): 41-54 [DOI: 10.1080/08109021003676359](https://doi.org/10.1080/08109021003676359).
 - Washington University is committed to curating the scholarly article output of its researchers with a digital repository, hence another vision to consider would be multiple institutional and subject archives operating (independently) to standards which allow searching, finding and using from all. New tools and standards would certainly develop for enhance access and use of this increased amount of content from publicly-funded research articles wherever it is stored.

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?

Response:

Under the Copyright Act, copyright in a work resides with the author until the author shares or surrenders some or all his rights. The intellectual property interests of publishers should be honored to the extent that they are given the necessary rights or permissions to distribute publicly funded works; however, they should not be granted rights that they do not need to accomplish the objective of distributing the work. As such, a balance can be made where authors grant to publishers the rights needed for a given period of time while retaining enough rights themselves to meet their own research and teaching obligations, as well as any conditions required under institutional or funder agreements. Creative Commons licenses, such as CC-BY, have been used by several commercial and non-profit journals (BioMed Central and PLoS) for years.

The Faculty of Washington University in St. Louis is already committed to making all of their scholarship and creative works freely and easily available to the world community. Faculty members are encouraged to seek venues for their works that share this ideal. In particular, when consistent with their professional development, members of the Faculty are encouraged to amend copyright agreements in order to retain the right to use his or her own work and to deposit such work in the University's digital repository or another depository, which is freely accessible to 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?

Response:

Federal archive(s) may be the simplest and most cost-effective mechanism. The Washington University Libraries has some concerns for long-term stewardship by the federal government since it needs to maintain budget flexibility and agencies can die or be cut with little warning. Using federal libraries, such as the National Library of Medicine, the National Agricultural Library, the Library of Congress, and the National Archives may be more appropriate since they have a mission for long-term storage.

As long as the archives are secure for the long term and created to standards for interoperability, a decentralized approach might be just as effective. Challenges for a decentralized approach include accountability to the federal agencies and confusion for funded researchers especially if different agencies have different policies.

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?

Response:

The Washington University Libraries see ample opportunities for public private partnerships that could allow for innovation in access to information without closing access to academic scholarship.

Publishers represent one type of entity that might be encouraged to participate in public/private partnership by providing repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles. However, the following should be noted:

- Under no condition should a publisher site – or any other stakeholder site - be the single point of access for publicly-funded articles. Publishers regularly buy and sell journals and their archives and occasionally drop journals and their archives entirely.
- With their extensive experience and existing archive infrastructure, universities and libraries have the opportunity to play an expanded role, and should be actively encouraged to partner with federal agencies.
- None of the 50+ research funders who currently have public access policies are using publisher sites as permanent archives. However, some funders have partnered with academic and research institutions in this role.

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?

Response:

Interoperable search, discovery and analysis capacity across disciplines and archives is already being developed and will certainly grow when more open access content becomes available.

Standard metadata services include support for interoperability with traditional library technologies as well as with general web operations. Repository systems should also provide support for:

- Use, reuse, and analysis of published works.
- Machine-readability and machine-interoperability (full text indexing and searching by web crawlers and commercial search engines).
- Open URL-based service.
- Meaningful and machine-navigable bridges between publications and the underlying data that support them. Examples include vocabularies for semantic relationships and unique identifiers.

Minimum standards will need to be suggested and eventually required.

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?

Response:

Any successful public-access policy must be consistent in its requirements and mandates because:

- Researchers often hold concurrent grants from multiple agencies.
- Uniform requirements and procedures regarding deposit of peer-reviewed literature will need to be established across all funding agencies involved.

Effective public-access policies can help maximize returns to taxpayers by ensuring that complete results are widely available in a timely manner. Policies should:

- Take advantage of existing protocols (e.g., SWORD) to facilitate automatic deposit of manuscripts to multiple repositories
- Encourage development of additional tools/services.
- Improve agency accountability policies in order to integrate articles with grants management systems (internal and external).

Public access policies will present opportunities to create/enhance productivity management tools for federal and internal reporting, through:

- Creation/enhancement of bibliographies and principal investigator profiles.
- Creation of opportunities for universities to better measure research output and impact.
- Use of University repositories as teaching tools (e.g., for teaching scholars about literature analytics, etc.)

It is important to note that such policy changes would only just begin to mitigate the burden lack of access to Federally funded research imposes on the public, that is, the lack of access to the research it funds.

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?

Response:

Educational materials (such as such as book chapters, texts, conference proceedings) that result from publicly funded research should be made readily accessible to the public when possible:

- The immediate efforts should focus on the scholarly journal articles published as a result of federally funded research, to stay focused on the primary mechanisms that scientists use to communicate their results and the current limitations on access to such publications.
- Different conditions may apply to different types of materials (e.g., authors are generally not paid for journals articles, but may be paid for textbook chapters).
- Policies should reflect these differences; for example, the policies under which educational materials are made accessible may need to differ from those directed at journal articles.

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?

Response:

No embargo period/immediate access would be ideal to optimize scientific and commercial utility of information contained in these articles.

However, to accommodate journal publishers who rely on subscription income, an author-determined embargo period of 0-12 months has proven to be acceptable across multiple disciplines.

- No publisher has indicated that the above embargo period policy (currently in use by NIH and numerous other funders around the world) has harmed them.
- Embargos of 12 months or less are the norm in research funder policies around the globe (see <http://roarmap.eprints.org/>)
- Embargos of 12 months or less have been adopted by hundreds of journals.
- Even publishers who previously expressed concern that opening access to back content would result in loss of revenue have now changed their embargo practices.
- Since the inception of the NIH public access policy, the Washington University Libraries has cancelled many journal subscriptions due to budget and use considerations, but no cancellations occurred because all or part of the journal was available in PubMed Central, with or without embargo.

If subscription-based publishers contend that embargo periods negatively impact revenue (presumably through cancellations), then all factors that impact subscriptions should be accounted for, including:

- Growth of journals and papers in discipline (competitive analysis)
- Price – and pricing history – of journal and competitive titles
- Impact of required bundles vs. single journals in discipline
- Library budget numbers/trends
- Real revenue resulting from “long-tail” citation articles
- Percentage of articles derived from federal funds

Respectively submitted by:

Shirley K. Baker
 Vice Chancellor for Scholarly Resources and Dean of University Libraries
 Washington University in St. Louis

I am sending this message in response to the RFI issued by the Office of Science and Technology Policy concerning public access to peer-reviewed scholarly publications resulting from federally funded research (Federal Register, v. 74, issue 214, November 4, 2011).

I'll begin my response by saying that I strongly believe that information, especially journal articles, resulting from federally funded research should be made publically available immediately upon publication. As a tax-paying citizen, I think I should have timely access to the information generated with my tax dollars. My responses to specific questions in the RFI are as follows:

Question 1:

- Grow existing and new markets – Making the information in journal articles freely available and fully usable will enable people to develop new products and services based on that knowledge. Having the information available quickly, with little or no embargo, will help accelerate the development of these new products and services.
- Improving the productivity of the scientific enterprise – Much research has been done over the past two decades on how discoveries are made and what can be done to accelerate the research and discovery process. The research shows that freely sharing information can greatly accelerate discovery and increase research productivity. For example, in a paper entitled *The Value of Openness in Scientific Problem Solving*, Karim Lakhani from the Harvard Business School says that “Openness and unrestricted information sharing amongst scientists...are critical to scientific progress.” He promotes the value of sharing information freely across disciplines because discoveries are often made at the periphery of a discipline by people who look at a problem very differently from traditional researchers actively studying a particular topic (www.hbs.edu/research/pdf/07-050.pdf). Michael Nielsen also discusses the value of freely sharing information in his book *Reinventing Discovery: The New Era of Networked Science*. The NIH Public Access Policy increases the flow of information and the productivity of researchers not only in academic institutions but in commercial companies as well, and similar policies for other federal funding agencies would further increase research productivity and enable important discoveries to be made more quickly.
- Cost/Benefit of policies – Government policies similar to the NIH PA Policy would make the entire research process more efficient and effective and greatly increase taxpayers' return on investment in research. Policies could be implemented in ways similar to the NIH policy, the implementation for which only costs 1/100th of 1% of the overall NIH budget. Not only is the NIH policy a cost effective method of insuring that information is freely shared, but it also provides an effective way for NIH to track the productivity/outcomes of the research it funds.
- Type of access needed – The value of providing access to information can be increased significantly if users of the information are also allowed to fully re-use the information without barriers.

Question 2:

- Protection for intellectual property – Public access policies can work within the current copyright policies. I think that federally funded authors should not be forced to relinquish copyright to their works upon publication, as the copyright should really belong to the author, or be shared between the author and the institution and/or funding agency supporting the research. Government policies should allow full use (distribution, reuse, data mining, etc.) of

information resulting from federally funded research, as this increases the value of the information and helps researchers build on that information to make further discoveries.

Question 3:

- Pros and cons of centralized and decentralized approaches – I think a centralized approach to providing access to scholarly publications resulting from federal funding is best, and that government policies should ensure that the articles are preserved and made available in standard and usable formats. NIH's PubMed Central is a good example of how a centralized approach can work well. Having multiple repositories, especially if some of them are private, would be more complicated, but could be adequate if they follow standards for access and use, and are not dark archives.

Question 6

- Maximizing benefits of policies while minimizing burdens – Standardized and consistent policies can result in data compilation that can be used to help institutions as well as the federal agencies better assess research productivity. The burdens of complying with the policies can be minimized if the policy requirements are standardized across agencies and consistent over a significant period of time.

Question 7

- Other materials – Yes, other types of publications based on federally funded research should be covered by government policies so they, too, can be freely accessible. I realize these types of publications are different from journal articles for which authors normally do not get paid, so the policies for them might need to be different from those covering journal articles.

Question 8

- Embargos – Ideally, publications resulting from federally funded research should be available to the public immediately upon publication. The current NIH PA policy allows a twelve-month embargo to help publishers protect their subscription income, and some of the funding agencies in Europe allow only a six-month embargo. I can find no evidence that these embargos have hurt publishers' subscription revenues, so I recommend an embargo period no longer than six months. Even an embargo period six months long will have a significant impact on the speed with which new discoveries are made.
- I have eliminated 15 positions (12.5 FTE's, 25% of our entire staff) in my library in the past ten years to help pay for the increasing cost of journals and databases. Despite the shifting of some personnel money to the journal budget, my library has been forced to cut journal subscriptions nearly every year because of increased subscription costs, including cutting 780 subscriptions (17% of our total subscriptions) this year alone. The embargo period was not one of the major factors we considered in determining which journals to cut. We looked at usage statistics, price and price history (if journal costs keep rising significantly every year, it's not sustainable) faculty and student input, subject coverage (what else we subscribed to in the same disciplines), the impact that our cancellations will have on other libraries in our group purchases, whether journals are part of a package and/or can be purchased individually, etc. Embargo periods have

had little if any impact on the journal cancellations in my library, which is the only academic health sciences library in my state.

- Having different embargo periods for journals in different disciplines could get complicated and confusing for researchers.

In summary, I strongly support and applaud the government's efforts to ensure broad public access to, and preservation of, peer-reviewed information resulting from federally funded research. I fervently hope that the National Science and Technology Council's Task Force on Public Access to Scholarly Publications will support expanding these efforts to more federal agencies. Thank you for giving us the opportunity to comment on this issue.

Mary L. Ryan, MLS, MPH
Library Director
University of Arkansas for Medical Sciences
Little Rock AR

FR Doc. 2011-28623
Filed 11/3/2011
Graduate and Professional Student Association
gapsa@gmu.edu
George Mason University
Fairfax, VA 22030

Response to Question 1a: We would have to say that the ideal would be that the Federal government ensures that the complete collection of articles resulting from publicly funded research is made freely accessible to the public and that the public can fully use them without commercial restrictions. Students and scientists especially need full Open Access (immediate, free online availability coupled with rights to reuse fully), which will create the environment that will improve students' educations, maximize scientific productivity, and accelerate commercial innovation. Enabling full reuse of these articles allows innovative individuals and companies to construct new services and new products on this publicly-funded content, improving the ability for U.S. companies to compete successfully on the global stage (where we are beginning to fall behind in hard sciences research). Furthermore, providing faster access to this crucial information allows innovative individuals, entrepreneurs and businesses to incorporate ideas generated from this research into their development cycles more quickly, speeding the launch of new services and products into the marketplace. Lastly, acceleration of this kind of commercialization spurs economic growth, creating new job opportunities across broad sectors of the economy, important in the best of times, even more so in bad economic times.

Response to Q1b: Research articles are quite literally the building blocks of a student's education. Open Access allows research results to be quickly incorporated into the teaching and learning process – improving the quality of education quickly and cost-effectively. Providing American students with the most complete, up-to-date education possible boosts US economic competitiveness, especially in innovative, cutting-edge fields like biotechnology and alternative energy. Professors can only teach what they have access to. Open Access greatly improves students' ability to get a complete education. Open Access helps students get projects off the ground and build businesses around their research. Losing access to the relevant research literature is a significant barrier for students who might consider leaving the Academy to start a business around their research. Students' library cards expire at graduation, meaning that the day they 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 in the Academy or private sector. This cost is even greater in a weak economy where students may spend a significant amount of time in their job search. Strong public access policies help level the playing field for students outside of the wealthiest institutions, who are at a distinct disadvantage when it comes to building their education from the most up-to-date research. This means they're less well prepared to contribute when it's time to put their education to use in the private or public sector. Open Access to research articles is critical a driver of scientific innovation and productivity. Open Access leads to increased citations and follow-on research. Open Access promotes diversity of follow-on research, and increases the pursuit of new research pathways. Open Access lets scientists incorporate new findings into their research faster. Opening access to research articles allows

scientists to get to – and read more – information than they previously could. It also lets scientists use new tools to incorporate more articles into their research faster. Open Access can enable machines as a new category of readers and users – opening up vast, previously unobtainable new research pathways, and making new connections possible.

Response to Q1c: Compared to benefits outlined above, the costs are very minimal. The NIH Public Access Policy costs approximately \$3.5 million per year out of a \$30+ billion budget. This is 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.

Response to Q1d: Access must be free, immediate, and coupled with the rights to reuse the articles fully in a digital environment. Restrictions that limit how we can use the scientific research we paid for limits the return to us as taxpayers. Access without reuse delivers only a fraction of the value. Broad re-use allows researchers to unlock additional value from research investment – now, and for decades. Enabling full reuse means we can do more with less; we don't have to duplicate research to be able to build on results, and we can continue to extract value from our initial investment for years to come. Courses are typically only 3 to 4 months long. Access must be immediate to provide students with the most up-to-date education possible – anything less impedes students' educations and hurts US economic competitiveness.

Response to Q2: No Comment

Response to Q3: An effective federal public access policy involving multiple research funding agencies could potentially involve multiple public access repositories. Approved repositories that meet conditions for public accessibility, use rights, interoperability, and long-term preservation of articles could be maintained by third parties, encouraging innovative public/private partnerships. All repositories must support access and use conditions that allow *all* interested parties to build on them, and to create new services, products, etc. on top of this publicly funded information. The federal government is an 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. To ensure this, any public access policies that are developed must give the federal government adequate rights to archive and make distribute publicly funded articles.

Response to Q4: No Comment

Response to Q5: No Comment

Response to Q6: For any public access policy to be successful, there must be consistency of requirements and mandates. Currently, research universities have faculty members and researchers who hold grants from all federal funding agencies and some of them have grants from multiple agencies concurrently. To the extent practicable, uniform requirements and procedures regarding deposit of peer-reviewed literature should be established across all funding agencies covered. Uniformity of deposit requirements will reduce the complexity and cost while at the same time, increase the rate of compliance.

Response to Q7: Educational materials that result from publicly funded research should be made readily accessible to the public. Other materials, such as book chapters, texts, conference proceedings should be made accessible, although the policies under which they are made accessible may need to differ from those directed at journal articles. Different conditions apply to different types of material (i.e., authors are not paid for journals articles, but may be paid for text book chapters) and policies should reflect these differences. In no way should these policies serve to jeopardize a researcher's agreement with publishers of his/her book or the like.

Response to Q8: Students want and need access now. The U.S. should be educating our students with the best and most current information possible. It is unacceptable to ask our students to “make do” with old information. Furthermore, since courses are typically only 3 to 4 months long, an embargo necessarily means students are missing information that would provide them a more up-to-date education. Still, an embargo period where an author can determine a hard stop-date between 0-12 months has been proven effective across multiple disciplines and would be an acceptable compromise. No data has ever been provided by any publisher that I am aware of that this embargo period has harmed them or their bottom-line. An embargo of 12 months or less has been adopted by hundreds of journals.