AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

FEDERALLY FUNDED RESEARCH: EXAMINING PUBLIC ACCESS AND SCHOLARLY PUBLICATION INTERESTS

HEARING

BEFORE THE

SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HOUSE OF REPRESENTATIVES

ONE HUNDRED TWELFTH CONGRESS

SECOND SESSION

THURSDAY, MARCH 29, 2012

Serial No. 112–74

Printed for the use of the Committee on Science, Space, and Technology

Available via the World Wide Web: http://science.house.gov

U.S. GOVERNMENT PRINTING OFFICE

73–607PDF WASHINGTON : 2012

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2104 Mail: Stop IDCC, Washington, DC 20402–0001

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HON. RALPH M. HALL, Texas, *Chair*

F. JAMES SENSENBRENNER, JR., Wisconsin
LAMAR S. SMITH, Texas
DANA ROHRABACHER, California
ROSCOE G. BARTLETT, Maryland
FRANK D. LUCAS, Oklahoma
JUDY BIGGERT, Illinois
W. TODD AKIN, Missouri
RANDY NEUGEBAUER, Texas
MICHAEL T. McCAUL, Texas
PAUL C. BROUN, Georgia
SANDY ADAMS, Florida
BENJAMIN QUAYLE, Arizona
CHARLES J. "CHUCK" FLEISCHMANN, Tennessee
E. SCOTT RIGELL, Virginia
STEVEN M. PALAZZO, Mississippi
MO BROOKS, Alabama
ANDY HARRIS, Maryland
RANDY HULTGREN, Illinois
CHIP CRAVAACK, Minnesota
LARRY BUCSHON, Indiana
DAN BENISHEK, Michigan
VACANCY

EDDIE BERNICE JOHNSON, Texas
JERRY F. COSTELLO, Illinois
LYNN C. WOOLSEY, California
ZOE LOFGREN, California
BRAD MILLER, North Carolina
DANIEL LIPINSKI, Illinois
DONNA F. EDWARDS, Maryland
BEN R. LUJÁN, New Mexico
PAUL D. TONKO, New York
JERRY McNERNEY, California
TERRI A. SEWELL, Alabama
FREDERICA S. WILSON, Florida
HANSEN CLARKE, Michigan
SUZANNE BONAMICI, Oregon
VACANCY
VACANCY
VACANCY

SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT

HON. PAUL C. BROUN, Georgia, *Chair*

F. JAMES SENSENBRENNER, JR., Wisconsin
SANDY ADAMS, Florida
RANDY HULTGREN, Illinois
LARRY BUCSHON, Indiana
DAN BENISHEK, Michigan
VACANCY
RALPH M. HALL, Texas

PAUL D. TONKO, New York
ZOE LOFGREN, California
BRAD MILLER, North Carolina
JERRY McNERNEY, California
EDDIE BERNICE JOHNSON, Texas

CONTENTS

Thursday, March 29, 2012

	Page
Witness List	2
Hearing Charter	3

Opening Statements

Statement by Representative Paul C. Broun, Chairman, Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, U.S. House of Representatives	18
Written Statement	19
Statement by Representative Paul Tonko, Ranking Minority Member, Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, U.S. House of Representatives	20
Written Statement	21

Witnesses:

Mr. H. Frederick Dylla, Executive Director and Chief Executive Officer, American Institute of Physics	
Oral Statement	23
Written Statement	26
Mr. Elliot Maxwell, Project Director for the Digital Connections Council, Committee on Economic Development	
Oral Statement	33
Written Statement	36
Dr. Crispin Taylor, Executive Director, American Society of Plant Biologists	
Oral Statement	44
Written Statement	47
Mr. Stuart Shieber, Director, Office for Scholarly Communications, Harvard University	
Oral Statement	54
Written Statement	57
Mr. Scott Plutchak, Director, Lister Hill Library at University of Alabama at Birmingham	
Oral Statement	68
Written Statement	70
Discussion	76

Appendix I: Answers to Post-Hearing Questions

Mr. H. Frederick Dylla, Executive Director and Chief Executive Officer, American Institute of Physics	92
Mr. Elliot Maxwell, Project Director for the Digital Connections Council, Committee on Economic Development	103
Dr. Crispin Taylor, Executive Director, American Society of Plant Biologists	108
Mr. Stuart Shieber, Director, Office for Scholarly Communications, Harvard University	120
Mr. Scott Plutchak, Director, Lister Hill Library at University of Alabama at Birmingham	129

Page

Appendix II:

Letter submitted by Representative Zoe Lofgren, Member, Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, U.S. House of Representatives 138
Response to Federal Register Notice Request for Information submitted by Dr. H. Frederick Dylla 142
Response to Federal Register Notice Request for Information submitted by Dr. Crispin Taylor 166
ALPSP Response to OSTP Request for Information 179

FEDERALLY FUNDED RESEARCH: EXAMINING PUBLIC ACCESS AND SCHOLARLY PUBLICATION INTERESTS

THURSDAY, MARCH 29, 2012

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT,
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY,
Washington, D.C.

The Subcommittee met, pursuant to call, at 9:35 a.m., in Room 2318 of the Rayburn House Office Building, Hon. Paul Broun [Chairman of the Subcommittee] presiding.

RALPH M. HALL, TEXAS
CHAIRMAN

EDDIE BERNICE JOHNSON, TEXAS
RANKING MEMBER

U.S. HOUSE OF REPRESENTATIVES
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
2321 RAYBURN HOUSE OFFICE BUILDING
WASHINGTON, DC 20515-6301
(202) 225-6371
www.science.house.gov

Subcommittee on Investigations & Oversight Hearing

Federally Funded Research:
Examining Public Access and Scholarly Publication Interests

Thursday, March 29, 2012
9:30 a.m. to 11:30 a.m.
2318 Rayburn House Office Building

Witnesses

Dr. Frederick Dylla, Executive Director and Chief Executive Officer, American Institute of Physics.

Mr. Elliot Maxwell, Project Director for the Digital Connection Council, Committee on Economic Development.

Dr. Crispin Taylor, Executive Director, American Society of Plant Biologists.

Mr. Stuart Shieber, Director, Office for Scholarly Communications, Harvard University.

Mr. Scott Plutchak, Director, Lister Hill Library, University of Alabama at Birmingham.

U.S. House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Investigations & Oversight

HEARING CHARTER

"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

Thursday, March 29, 2012
9:30 a.m. – 11:30 a.m.
2318 Rayburn House Office Building

Purpose

On Thursday March 29, 2012, the Subcommittee on Investigations and Oversight will hold an oversight hearing to examine various models for disseminating federally funded research and their corresponding effects on the scientific process. Federally funded research is accessed through an increasing variety of methods beyond the traditional scholarly journal maintained by a scientific society that is made available only through a paid subscription. Some of the push towards greater public access stems from increasing complaints about the widely varying subscription costs of journals.[1]

Advocates have urged Congress to require federally funded research to be made available online to the public with limited or no access restrictions. Some longstanding journals have switched to this model while other journals and their publishers prefer their existing business models. They fear that switching to a business model with greater public access will not generate enough income to replace existing subscription revenue and will threaten their viability. Some highly specialized journals have already switched to business models with greater public access utilizing long-term financial commitments by their institutional subscribers to replace subscription revenue. New journals based upon minimal or no access restrictions have grown in number during the past decade.

Background

The federal government funds 31% of all R&D conducted in the U.S., compared to 62% by business entities and the remaining 7% by universities, local and state governments, and other non-profit organizations.[2] This federally funded research is conducted by federal employees, private companies, and public institutions such as universities and research centers. With the exception of biomedical and defense research, most federally funded research is funded by federal agencies under the Committee's primary jurisdiction.

---

[1] The upper range of journal pricing appears to be around $30,000 per year as charged by Wiley & Sons for some versions of annual online access to The Journal of Comparative Neurology.
[2] "Science and Engineering Indicators 2012." *NSF.gov*. Web. 27 Mar. 2012. <http://www.nsf.gov/statistics/seind12/start.htm>. See Appendices B and C.

One of the primary goals of federal research and development funding is the wide dissemination of robust research in a variety of fields.[3] When made available to others through research articles and other means, this research can then be used as a building block for future research efforts or be commercialized in some way. Prior to the Internet, access to, and dissemination of, research papers was accomplished through publication in a research journal available to subscribers of that journal. With the advent of widespread Internet access, the traditional journal publishing model began to change to focus less on printed access and more on online access that still requires an access charge.

In recent years, a movement has developed to allow anyone to access federally funded research for no additional cost. The rationale being that free access should be permitted because federal taxpayers have already paid to fund the research. Others point out that any effort to force specific models of access, such as free or low cost access, will threaten important components of the existing publishing system, especially the financial health of smaller journals whose revenues fund the operation of affiliated scientific societies. The intellectual property rights of journals to control copying and distribution of their copyrighted works may also be harmed by federal open access mandates.

Parties interested in how federally funded research is accessed include:

- Researchers that use federal funds to conduct research and want their articles published, not only to add to the scientific body, but also to demonstrate their skills and knowledge in the topic area.
- Journals and their societies that are responsible for the publication and distribution of research papers that have been peer reviewed and edited by the journal.
- Policy advocates and commercial entities who want to understand the science used as a basis for proposed federal action.
- Taxpayers who want access to the research funded by their tax dollars for reasons that range from their own investigation into medical issues that impact them or their families to small businesses that do not have large research and development resources.

The Roles and Interests of the Federally Funded Researcher

Researchers interested in using federal funds to undertake research apply for research grants from federal agencies such as the National Science Foundation. According to the NSF, the agency receives 40,000 funding applications per year of which 11,000 are funded. Rules applicable to most research grant recipients require the creation of a research paper based upon the research.[4]

Researchers want to be published for several reasons beyond their desire to expand scientific knowledge. For example, university tenure often depends upon the volume of publishing. Publishing can attract additional grants based upon the initial research. There is also the prestige factor in publishing articles in highly regarded journals that are then cited by other researchers.

---

[3] For example, Section 203(a)(3) of the Space Act of 1958, as amended, directs NASA to "provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

[4] Proposal and Award Policies and Procedures Guide, NSF 11-1, October 2010.

The Roles and Interests of a Journal and its Governing Society
Journals and their societies vary significantly in size and budget, generally driven by the number of researchers in that field. Societies are advocates for science in their field of study, hosting annual conferences and identifying grant and career opportunities for researchers in that field. Their journals are usually considered the most authoritative in that particular field of study.

Journals and their editors can serve as gatekeepers to particular areas of science through their own efforts and the decisions of peer review panels. This is viewed positively by traditional societies because poorly written or duplicative scientific papers can be weeded out prior to entering the peer review system that would waste the time of other researchers. In contrast, some advocates of greater public access argue that unconventional approaches to science may not be recognized by conservative journal editors so that the best option is simply to make all research available to let the reader decide what is important, accurate, and relevant.

Based upon a Carnegie Commission definition, OMB in 2004 defined peer review in its *Final Information Quality Bulletin for Peer Review* as "a form of deliberation involving an exchange of judgments about the appropriateness of methods and the strength of the author's inferences."[5] In essence, the quality of the review process depends upon the skills and interest of these "peers." The peer review process has been in existence since 1731 when the Royal Society of Edinburgh published Medical Essays and Observation.[6] At the time, the reviewers of these articles were individuals considered by the editor to be the "most versed in these matters."

The vast majority of journals still use the peer review process, but it is not a perfect system. The peer review process provides:

- An initial elimination by the journal editor of articles that do not meet commonly accepted scientific principles.
- An in-depth review of a journal topic conducted by those already knowledgeable about the topic.
- An ability to identify some research fraud, misconduct, and integrity issues.

Potential problems with the peer review process include:

- Choices of peer reviewers by the journal editor that are not sufficiently knowledgeable about the topic of that journal article.
- Any bias or lack of objectivity by one or more individuals on the peer review panel.
- No guarantees that all scientific fraud, misconduct, or integrity issues have been identified.
- Publication delays caused by the peer review process itself.[7]

Alternative peer review models have been tested by some journals, but it is unclear at this time whether these new models will replace or merely supplement the traditional peer review process.[8]

---

[5] Carnegie Commission on Science, Technology, and Government, Risk and the Environment: Improving Regulatory Decision Making, Carnegie Commission, New York, 1993: 75.
[6] The Ups and Downs of Peer Review, Benos et al. *Advances in Physiology Education,* 31: 145–152, 2007, p.145.
[7] Sieber, Joan (2006). Quality and value: How can we research peer review?, *Nature*. doi:10.1038/nature05006

The Roles and Interests of Commercial Journal Publishers
Societies that do not have the resources or interest in maintaining in-house publishing staff for their journals that may only be published a few times per year can outsource the work to larger commercial publishers. Under this system, a percentage of revenues or a flat fee is paid to the society. This system effectively merges the interests of both parties.

Although many commercial publishers state that they are not opposed to greater public access, they are concerned about, and are often in opposition to, government mandates that would encourage or force a specific business model to be followed. They are also concerned with government efforts to force the publication of the version of the article that is edited and peer reviewed by the journal, the "version of record", instead of the version initially submitted by the researcher, the "accepted manuscript."[9] Unlike a researcher who may have agreed to publish his work in a particular manner as a condition for receiving federal funds, journal publishers have not entered into any binding contracts or other agreements with federal agencies concerning their version of the article. Publishers believe that mandates requiring the version of record to be available for free either immediately or within a specific amount of time after initial publication in their journal is a violation of their intellectual property rights and threatens their basic business model.[10]

The Roles and Interests of Universities and Their Libraries
Although some universities and their libraries have taken the lead in supporting open access models as part of their mission to be a repository for knowledge, universities may have several interests related to access to federally funded research. Researchers want greater access to federal funding so that more research can be undertaken; university libraries want their users to have the widest possible access to research; and university based publishers want to ensure continued income from their publications to fund their operations. In response to overall budgetary concerns, libraries have looked to their journal subscription costs as one area to reduce spending. With the increasing growth of free online journals, this pressure to further reduce spending on paid journal subscriptions is likely to increase further.

Advocates for greater access point to the high subscription costs and profits of large commercial publishers as one reason to embrace open access models. For example, the largest for-profit journal publisher, Elsevier, publishes approximately 2,000 journals, some in partnership with scientific societies that outsource publication to Elsevier and share revenue. On annual revenue of 2.058 billion pounds (approximately $3.3 billion) in 2011, Elsevier earned a profit of 768 million pounds (approximately $1.2 billion) for a profit margin of 37%.[11] Commercial publishers see this profit as earned from work done by the publisher in marketing and creating tools for societies to use to publish their journals, rather than a profit off of the research itself.

---

[8] Overview: Nature's peer review trial. *Nature* (2006) | doi:10.1038/nature05535
[9] The definition of these terms is based upon the "Recommendations of the NISO/ALPSP Working Group on Versions of Journal Articles" available at http://www.niso.org/apps/group_public/download.php/48/Recommendations_TechnicalWG.pdf.
[10] See testimony of Allan Adler, American Association of Publishers, before the House Oversight and Government Reform Committee, July 29, 2010. Available at http://oversight.house.gov/wp-content/uploads/2012/01/20100729Adler.pdf.
[11] Reed Elsevier 2011 Annual Report. Half of this revenue is from the North American market. During the past five years, Elsevier revenue has increased by 37% while profits have increased by 61%.

The Taxpayer Interests
With taxpayer funding responsible for approximately one third of all research and development in the nation, taxpayers have a vested interest in how their research dollars are used and how the results of research can be accessed. [12] Duplicative research wastes taxpayer resources that could be used for other purposes. Research that is duplicative or hard to locate is less beneficial than unique research that can be quickly used to accelerate scientific progress. Follow-on research continues expanding the scientific record and validates previous research in that area. This in turn bolsters confidence in the validity of the conclusions.

The Foundations of Open Access
In 1991, an online repository of physics articles was created in a service called "ArXiv", a database that has now grown to over 700,000 articles.[13] Noting the success of ArXiv and similar databases, a group of interested researchers came together in Budapest, Hungary in February 2002 under the auspices of the Soros Foundation and released the Budapest Open Access Initiative defining open access as:

> " ... we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."[14]

The European Union released a report on scientific publishing in 2005 that highlighted the societal and scientific benefits of open access models.[15] Today, there are over 7,500 open access journals accessible online.[16]

Gold Open Access (Journals)
Gold open access is the term given to the publishing model by which the costs to publish an article are recouped by means other than subscription or access charges imposed upon the reader.[17] Although revenue from advertisements in the journal may be used to offset costs, the most common source of funding for this publishing model are fees collected in advance from the author or his sponsoring institution. These fees, sometimes called "page charges," "printing charges," or "publication charges," are used by the publishing journal to pay for arranging the peer review, final editing, and publication. Under this model, the researcher knows the

---

[12] National Science Foundation, *Science and Engineering Indicators 2012* (NSB 12-01), supra.
[13] "The First Free Research-Sharing Site, ArXiv, Turns 20 With an Uncertain Future." - *Wired Campus*. Web. 27 Mar. 2012. <http://chronicle.com/blogs/wiredcampus/the-first-free-research-sharing-site-arxiv-turns-20/32778?sid=wc>.
[14] "Budapest Open Access Initiative." *Budapest Open Access Initiative*. Web. 27 Mar. 2012. <http://www.soros.org/openaccess/read>.
[15] "Digital Broadband Content: Scientific Publishing" OECD. DSTI/ICCP/IE(2004)11/FINAL. 02 Sep. 2005.
[16] "Directory of Open Access Journals." *Directory of Open Access Journals*. Web. 27 Mar. 2012. <http://www.doaj.org/>.
[17] "Peter Suber, Open Access Overview." *Open Access Overview*. Web. 27 Mar. 2012. <http://www.earlham.edu/~peters/fos/overview.htm>.

applicable fees for gold open access journals before submitting a paper for possible publication. A review of current charges for gold open access journal finds a typical cost of several thousand dollars in charges to the author of the article. Charges in excess of five thousand dollars per paper appear to be rare.[18]

Federal guidelines permit the payment of such charges. Revised in May 2004, paragraph 34 of attachment B to OMB Circular A-87 states that: "Page charges for professional journal publications are allowable as a necessary part of research costs where: (1) The research papers report work supported by the Federal Government; and (2) The charges are levied impartially on all research papers published by the journal, whether or not by federally sponsored authors."

OMB and individual federal agencies do not currently require publication in an open access journal although there is anecdotal evidence that there is pressure from funding agencies for authors to seek out journals with the lowest page charges, rather than journals of the author's choosing. For example, the library of CERN, Europe's Organization for Nuclear Research, that receives federal funds to support operations of the Large Hadron Collider states that it "... encourages authors to publish in Open Access journals." [19]

Green Open Access (repositories)
Green open access is the term given to publicly accessible self-archiving efforts by article authors or various host institutions. Under the green open access model, once an article is considered ready for public release, the author or his host institution deposits a copy of the article in a publicly accessible online database. The author may not have asked others to review the article prior to publication and the article may have already been published elsewhere. If peer review has not occurred, the "strength" of the article as viewed by others may not be as high as that of peer reviewed articles. The Digital Access to Scholarship at Harvard (DASH) is one example of such a repository.[20]

Current Federal Agency Efforts Concerning Public Access of Federally Funded Research
The most significant role undertaken by the federal government relating to public access has been at the National Institutes of Health (NIH). NIH operates PUBMED Central, a centralized, publicly accessible database containing 2.4 million journal articles that have been submitted by 3,000 journals who deposit some or all of the articles in their journals to the database.[21]

Another effort by federal agencies to enable greater public access to research funded in part by federal agencies, albeit much smaller in scale than PUBMED Central, has been a series of affirmative steps by the libraries of federal energy labs to participate in international collectives to support the move of a small group of selected physics journals to an open access model. The

---

[18] A comparison of charges from several publications can be found at http://www.biomedcentral.com/about/apccomparison.
[19] "OA and low cost journals: where to publish?" *OA and low cost journals: where to publish?* Web. 27 Mar. 2012. <http://library.web.cern.ch/library/OpenAccess/Journals.html/>
[20] "Digital Access to Scholarship at Harvard: Opening Harvard Research." *Digital Access to Scholarship at Harvard: Opening Harvard Research*. Web. 27 Mar. 2012. <http://dash.harvard.edu/>.
[21] "PUBMED Central." *National Center for Biotechnology Information*. U.S. National Library of Medicine. Web. 27 Mar. 2012. <http://www.ncbi.nlm.nih.gov/pmc/>.

Sponsoring Consortium for Open Access Publishing in Particle Physics ($SCOAP^3$) is a worldwide consortium of:

- High-energy physics funding agencies
- High-energy physics laboratories
- Leading national and international libraries and library consortia.[22]

$SCOAP^3$ began with a 2007 proposal of interested groups including physics scientists to move a small group of journals focusing on high-energy physics and related fields to an open access model. The goal was to replace each journal's income from publications and access charges with guaranteed sources of annual dues from the coalition members in return for the chosen journals to move to an open access model. In the U.S., $SCOAP^3$ partners are primarily universities, but the following Department of Energy labs are also members through their respective libraries:

- Argonne National Laboratory
- Fermilab
- Los Alamos National Laboratory
- Lawrence Berkeley Laboratory - University of California, Berkeley
- Lawrence Livermore National Laboratory
- Pacific Northwest National Laboratory
- Savannah River National Laboratory
- Stanford Linear Accelerator Center
- Thomas Jefferson Laboratory[23]

Efforts to convert several physics journals to an open access model are now underway. Dues from $SCOAP^3$ partners will be calculated based upon the nationality of the articles published in the covered journals during the year 2005. For U.S. libraries and national labs, the U.S. share of the overall cost would be 24% although the financial contribution of DOE labs is far less than 1% of the overall cost.[24]

Previous Committee Activity

In June 2009, then Science, Space, and Technology Committee Chairman Bart Gordon asked interested parties to meet under the auspices of the Committee to discuss scholarly publishing issues and develop, to the greatest extent possible, recommendations for how public access to journals with content derived from federally funded research could be increased.[25] An ad-hoc group called the Scholarly Publishing Roundtable was formed. The participants included representatives from academia, research librarians, journal publishers, and researchers in the area of library and information science. Committee and OSTP staff also joined the participants.

In January 2010, the Scholarly Publishing Roundtable released a 25-page report containing eight recommendations as follows:

1. Agencies should work in full and open consultation with all stakeholders, and with OSTP, to develop their public access policies

[22] "$SCOAP^3$." *$SCOAP^3$*. Web. 27 Mar. 2012. <http://scoap3.org >
[23] "$SCOAP^3$ US." *$SCOAP^3$*. Web. 27 Mar. 2012. <http://scoap3.org/scoap3us_alpha.html>
[24] *Ibid.*
[25] The charge to the group can be found at http://www.aau.edu/WorkArea/DownloadAsset.aspx?id=9666.

2. Agencies should establish specific embargo periods between publication and public access
3. Policies should be guided by the need to foster interoperability
4. Every effort should be made to have the version of record as the version to which free access is provided
5. Government agencies should extend the reach of their access policies through voluntary collaborations with nongovernmental stakeholders
6. Policies should foster innovation in the research and educational use of scholarly publications
7. Government public access policies should address the need to resolve the challenges of long-term digital preservation
8. OSTP should establish a public access advisory committee

These recommendations were supported by 13 of the 15 roundtable participants. The two dissenters from the roundtable recommendations were the representatives from Elsevier and PLoS, the Public Library of Science, an open access publisher.

Following the release of these recommendations, the America Competes Reauthorization Act of 2010 tasked the Office of Science and Technology Policy (OSTP) to "establish a working group under the National Science and Technology Council with the responsibility to coordinate Federal science agency research and policies related to the dissemination and long-term stewardship of the results of unclassified research, including digital data and peer-reviewed scholarly publications, supported wholly, or in part, by funding from the Federal science agencies."[26] The working group was required to submit a report to Congress within one year after enactment. OSTP collected 378 public comments for this report that is expected to be submitted to Congress within the next few weeks.[27]

Related Legislation

Three pieces of legislation that focus on this issue have been introduced in the 112th Congress. On December 16, 2011, H.R. 3699, the Research Works Act of 2011 was introduced and referred to the House Oversight and Government Reform Committee.[28] On February 12th, 2012, the Federal Research Public Access Act of 2012 was introduced in the House and Senate as H.R. 4004 and S. 2096.[29] H.R. 4004 was referred to the House Oversight and Government Reform Committee and S. 2096 was referred to the Senate Homeland Security and Government Affairs Committee. No hearings or other legislative action have occurred on any of the legislation.

H.R. 3699 and H.R. 4004 / S. 2096 take opposite approaches to public access of federally funded research. H.R. 3699 effectively prohibits federal agencies from adopting open access mandates. In contrast, H.R. 4004 / S. 2096 requires federal agencies with extramural research expenditures of over $100 million to adopt specific policies that result in free public online access to peer-

[26] Section 103 of P.L. 111-358.

[27] "Request for Information: Public Access to Digital Data Resulting From Federally Funded Scientific Research," 76 FR 68517, November 4, 2011. Submitted comments can be found at http://www.whitehouse.gov/administration/eop/ostp/library/publicaccess.

[28] The sponsor of H.R. 3699 is Mr. Issa of California.

[29] The sponsor of H.R. 4004 is Mr. Doyle of Pennsylvania. The sponsor of the Senate companion bill is Senator Cornyn of Texas.

reviewed articles not later than six months after publication in peer-reviewed journals, shortening the current NIH requirement.

Issues

Intellectual Property

Article I, Section 8 of the U.S. Constitution gives 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." All publishers, journal or otherwise, rely upon U.S. copyright law, codified at Title 17 of the U.S. Code, as the basis for an ownership right that can be enforced in federal court. Although U.S. copyright law is detailed, for the purposes of this issue, the most relevant provisions of copyright law are the statutory rights to control uses of copyrighted works and defenses to infringement of those rights.[30]

All versions of a research paper and various versions of the related journal article are protected by U.S. copyright law. Under the traditional journal system, authors typically continue to own the copyright to their original version, but assign a non-exclusive right to the journal to reproduce the article, either in its initial form or, most commonly, after changes made due to peer review and journal editing and formatting. The journal then typically owns the copyright in the peer-reviewed, journal-edited version. Determining which version should be made available to the public for free is a major issue of concern for all parties since the version of record is considered the most authoritative.

Commercial publishers depend upon strong intellectual property protections to protect their publications.[31] Efforts by the government to force journals to retroactively make available their copyrighted articles on the Internet for public access could potentially run afoul of the takings clause of the U.S. Constitution. However, should a federal agency prospectively require recipients of federal funds to ensure that the publication of their work was made in an open access manner as a condition for receipt of federal funding for the research, a takings complaint would likely not be successful. The latter approach is the one taken so far by NIH.[32] However, this effectively forces journals to yield their intellectual property rights in the peer reviewed

---

[30] 17 U.S.C. 106 identifies six rights of copyright owners, the first three of which are directly relevant:
(1) to reproduce the copyrighted work in copies or phonorecords;
(2) to prepare derivative works based upon the copyrighted work;
(3) to distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending;
17 U.S.C. 107 identifies limitations of these rights, "fair use", and uses a four part balancing test to determine if the defense can be used:
(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
(2) the nature of the copyrighted work;
(3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
(4) the effect of the use upon the potential market for or value of the copyrighted work.
There are other limitations contained within Title 17 including some focused on education and library users.

[31] See testimony of Ralph Oman, former Register of Copyrights, before the House Oversight and Government Reform Committee, July 29, 2010. Available at http://oversight.house.gov/wp-content/uploads/2012/01/20100729Oman.pdf.

[32] NIH open access policies, current and past, are available at http://publicaccess.nih.gov/policy.htm.

article to the NIH requirement or not accept an article funded by NIH for publication. In an attempt to address publisher concerns that these articles will be resold by others, NIH does prohibit the mass downloading of PUBMED Central articles.[33]

Data Access
Printed research journals rarely include all of the supporting data used to support a research article due to space and printing costs, but in practice, scientists will often make the data available to other researchers upon request. Online access eliminates many publishing costs and some have argued that data created by federally funded research should also be made available so long as it does not conflict with other federal laws on privacy and confidentiality. This data could then be more easily used by other researchers and interested parties seeking to validate the data. There appears to be differing levels of support for greater access to the underlying data than for greater access to research journal articles.

Time Delays
Open access journals typically operate under a no delay system. Once an article is deemed ready for publication by the journal's editors, it is made available online immediately. Current law requires that "The Director of the National Institutes of Health shall require that all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication."[34] This delay allows publishers to continue to sell subscriptions to those who want immediate access to newly published research without free competition from PUBMED Central.

The need for a delay and the exact amount of its duration is a subject of debate with some advocates seeking immediate public release on the Internet while others seek to maintain a twelve month delay. Some have suggested alternative timeframes of a six or nine month delay in free access.

Witnesses

Dr. H. Frederick Dylla **
Executive Director and CEO, American Institute of Physics

Mr. Elliot Maxwell
Project Director for the Digital Connections Council, Committee on Economic Development

Dr. Crispin Taylor **
Executive Director, American Society of Plant Biologists

Mr. Stuart Shieber
Director, Office for Scholarly Communications, Harvard University

---

[33] "PMC Copyright Notice." *National Center for Biotechnology Information.* U.S. National Library of Medicine. Web. 27 Mar. 2012. <http://www.ncbi.nlm.nih.gov/pmc/about/copyright/>.
[34] P.L. 110-161, Section 218.

Mr. Scott Plutchak **
Director, Lister Hill Library at University of Alabama at Birmingham

** Dylla, Plutchak, and Taylor were members of the 2010 Scholarly Publishing Roundtable organized under the auspices of the Committee.

Appendix A

Section 103 of the America Competes Act of 2010

SEC. 103. INTERAGENCY PUBLIC ACCESS COMMITTEE.

(a) Establishment- The Director shall establish a working group under the National Science and Technology Council with the responsibility to coordinate Federal science agency research and policies related to the dissemination and long-term stewardship of the results of unclassified research, including digital data and peer-reviewed scholarly publications, supported wholly, or in part, by funding from the Federal science agencies.

(b) Responsibilities- The working group shall--

(1) identify the specific objectives and public interests that need to be addressed by any policies coordinated under (a);

(2) take into account inherent variability among Federal science agencies and scientific disciplines in the nature of research, types of data, and dissemination models;

(3) coordinate the development or designation of standards for research data, the structure of full text and metadata, navigation tools, and other applications to maximize interoperability across Federal science agencies, across science and engineering disciplines, and between research data and scholarly publications, taking into account existing consensus standards, including international standards;

(4) coordinate Federal science agency programs and activities that support research and education on tools and systems required to ensure preservation and stewardship of all forms of digital research data, including scholarly publications;

(5) work with international science and technology counterparts to maximize interoperability between United States based unclassified research databases and international databases and repositories;

(6) solicit input and recommendations from, and collaborate with, non-Federal stakeholders, including the public, universities, nonprofit and for-profit publishers, libraries, federally funded and non federally funded research scientists, and other organizations and institutions with a stake in long term preservation and access to the results of federally funded research;

(7) establish priorities for coordinating the development of any Federal science agency policies related to public access to the results of federally funded research

to maximize the benefits of such policies with respect to their potential economic or other impact on the science and engineering enterprise and the stakeholders thereof;

(8) take into consideration the distinction between scholarly publications and digital data;

(9) take into consideration the role that scientific publishers play in the peer review process in ensuring the integrity of the record of scientific research, including the investments and added value that they make; and

(10) examine Federal agency practices and procedures for providing research reports to the agencies charged with locating and preserving unclassified research.

(c) Patent or Copyright Law- Nothing in this section shall be construed to undermine any right under the provisions of title 17 or 35, United States Code.

(d) Application with Existing Law- Nothing defined in section (b) shall be construed to affect existing law with respect to Federal science agencies' policies related to public access.

(e) Report to Congress- Not later than 1 year after the date of enactment of this Act, the Director shall transmit a report to Congress describing--

(1) the specific objectives and public interest identified under (b)(1);

(2) any priorities established under subsection (b)(7);

(3) the impact the policies described under (a) have had on the science and engineering enterprise and the stakeholders, including the financial impact on research budgets;

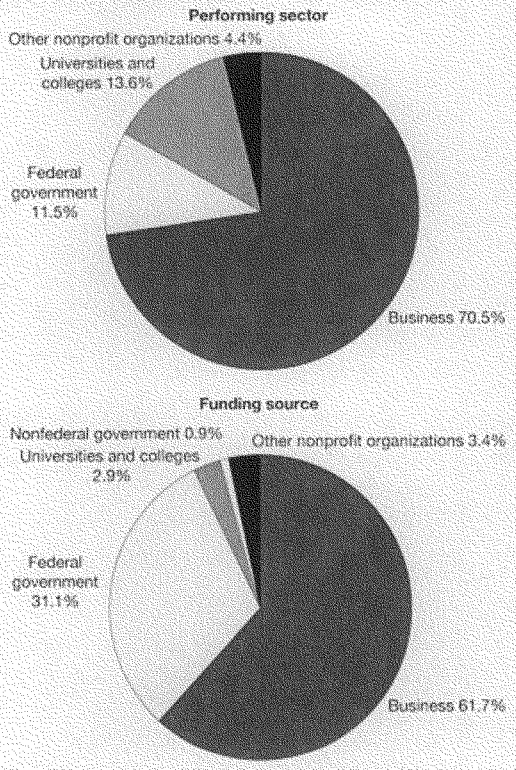
(4) the status of any Federal science agency policies related to public access to the results of federally funded research; and

(5) how any policies developed or being developed by Federal science agencies, as described in subsection (a), incorporate input from the non-Federal stakeholders described in subsection (b)(6).

(f) Federal Science Agency Defined- For the purposes of this section, the term `Federal science agency' means any Federal agency with an annual extramural research expenditure of over $100,000,000.

Appendix B

Figure 4-3
Shares of U.S. total R&D expenditures, by performing sector and funding source: 2009



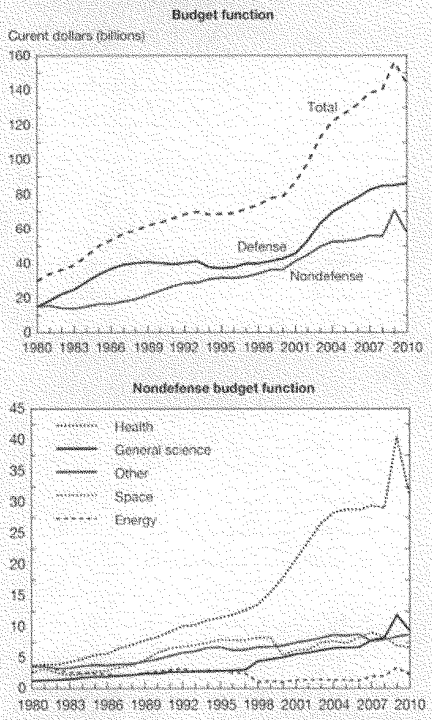
NOTES: Some figures involve estimates and may later be revised. National R&D expenditures are estimated to be $400.5 billion in 2009. Federal performing sector includes federal agencies and federally funded research and development centers. State and local government support to business is included in business support for business performance.

SOURCE: National Science Foundation, National Center for Science and Engineering Statistics, National Patterns of R&D Resources (annual series). See appendix tables 4-3 and 4-7.

Science and Engineering Indicators 2012

Appendix C

Figure 4-10
Federal budget authority for R&D, by budget function: FY 1980–2010



NOTES: Data for FY 2010 are preliminary. Data for FY 2009 include the additional federal funding for R&D appropriated by the American Recovery and Reinvestment Act of 2009. Other includes all nondefense functions not separately graphed, such as agriculture and transportation.1998 increase in general science and decrease in energy, and 2000 decrease in space were results of reclassification.

SOURCE: National Science Foundation, National Center for Science and Engineering Statistics, Federal R&D Funding by Budget Function (FY 2009–11). See appendix table 4-28.

Science and Engineering Indicators 2012

Chairman BROUN. The Subcommittee on Investigations and Oversight will now come to order.

Good morning, everyone. Welcome to today's hearing titled, "Federally Funded Research: Examining Public Access and Scholarly Publication Interests." You will find in front of you packets containing our witness panel's written testimony, their biographies, and Truth-in-Testimony disclosures. I now recognize myself for five minutes for an opening statement.

The formal review and communication of research finding dates back several centuries. Over this time, it has certainly served society very well. Scholars can argue over whether the existing structure, including peer review, is sufficient or if more can be done to ensure quality; but one thing is certain—society has greatly benefited from it.

This structure and process, however, is organic and ever-changing. As we progress through the digital age, expectations of access to scientific findings are increasing, specifically research funded by taxpayers. Just as the Internet has challenged entrenched interests in other mediums such as news and music, so too has it affected scholarly publishing. The academic community and scholarly publishing interests must be flexible enough to adapt to our ever-changing times.

Society's expectations of transparency are clearly increasing. Couple this trend with the fact that taxpayers rightfully expect access to research that they have funded, and you quickly realize that we all must work together to ensure that the various interests involved are treated fairly, and that ultimately, science and research are not harmed.

This is no small matter. There are more than 25,000 peer-reviewed journals produced by over 2,000 publishers. These journals publish more than 1.5 million articles a year and earn revenues between $8 and $10 billion from their subscribers. This revenue funds over 100,000 jobs worldwide and 30,000 jobs in the United States alone.

I have a lot of questions about how we should meet the challenges of expanding access to research without compromising the quality of the product or the rights of those involved in the process.

The National Institute of Health public access policy provides a good opportunity for Congress to review the effectiveness of increased transparency on certain fields of research, as well as its impact on publishers. We must be mindful, however, that what works or does not work for NIH and biomedical research may not be appropriate for other agencies and scientific fields. Is there a one-size-fits-all policy that can cover the entire Federal Government, or do specific agencies and disciplines require different approaches? What does the taxpayer have a right to access—the manuscript that is produced by the researcher or the final product that is peer-reviewed? How does copyright law affect each of these? How long after release of a paper should the public have access? Immediately? Six months? A year? Does this challenge—does this change depending on the discipline or the agency? Should all information and data associated with the research be made public? Is this reasonable or even possible? Should limitations be placed on

access to prevent misuse? How should that be decided or who should decide it? I can go on and on with question after question.

I am also curious about how public access has impacted the quality of the research. Has increased access impacted the number of citations and references, and is this even a valuable metric to determine effectiveness? Has greater access spurred additional inquiries or novel research? Has the public access affected innovation and commercialization? How do varying funding models for peer-review impact how researchers and agencies fund research? Are journals capable of adapting to meet new challenges presented by the digital age, transparency demands, and competitor models?

As you can see I have lots of questions and I still could on, but one thing is absolutely crystal clear. Any effort to fundamentally change the way in which federal research is reviewed, vetted, transmitted, and communicated should benefit from the Science, Space, and Technology Committee's input. We have been involved in investigating issues surrounding public access for a number of years and are uniquely qualified to evaluate the impacts on research, as well as on federal agencies.

Representative Gordon, the former full Committee Chairman, brought together a number of stakeholders in 2009 in order to find a common ground, and in 2010, the "Scholarly Publishing Roundtable" issued a report containing several recommendations. We also tasked the Office of Science and Technology Policy to address the issue in the America COMPETES Reauthorization Act of 2010 and expect their results very soon. Today's hearing is an extension of this longstanding engagement.

I look forward to working with all of the interested parties, and I want to thank our all of our witnesses for appearing here today.

[The prepared statement of Dr. Broun follows:]

PREPARED STATEMENT OF SUBCOMMITTEE CHAIRMAN PAUL BROUN

The formal review and communication of research findings dates back several centuries. Over this time, it has certainly served society well. Scholars can argue over whether the existing structure, including peer review, is sufficient or if more can be done to ensure quality, but one thing is certain, society has greatly benefited from it.

This structure and process, however, is organic and ever changing. As we progress through the digital age, expectations of access to scientific findings are increasing, specifically research funded by taxpayers. Just as the Internet has challenged entrenched interests in other mediums such as news and music, so too has it affected scholarly publishing. The academic community and scholarly publishing interests must be flexible enough to adapt to our ever-changing times.

Society's expectations of transparency are clearly increasing. Couple this trend with the fact that taxpayers rightfully expect access to research they have funded, and you quickly realize that we all must work together to ensure that the various interests involved are treated fairly, and that ultimately science and research are not harmed.

This is no small matter. There are more than 25,000 peer-reviewed journals, produced by over 2,000 publishers. These journals publish more than 1.5 million articles a year, and earn revenues between $8 and $10 billion dollars from their subscribers. This revenue funds over 100,000 jobs worldwide—30,000 in the U.S. alone.

I have a lot of questions about how we should meet the challenges of expanding access to research without compromising the quality of the product, or the rights of those involved in the process. The National Institute of Health (NIH) public access policy provides a good opportunity for Congress to review the effectiveness of increased transparency on certain fields of research, as well as its impact on publishers. We must be mindful, however, that what works or does not work for NIH and biomedical research, may not be appropriate for other agencies and scientific fields. Is there a one-size-fits-all policy that can cover the entire federal government,

or do specific agencies and disciplines require different approaches? What does the taxpayer have a right to access—the manuscript that is produced by the researcher, or the final product that is peer reviewed? How does copyright law affect each of these? How long after release of a paper should the public have access? Immediately? Six months? A year? Does this change depending on the discipline or the agency? Should all information and data associated with that research be made public? Is this reasonable or even possible? Should limitations be placed on access to prevent misuse? How should that be decided, and who should decide it?

I'm also curious about how public access has impacted the quality of research. Has increased access impacted the number of citations and references, and is this even a valuable metric to determine effectiveness? Has greater access spurred additional inquiries or novel research? How has public access affected innovation and commercialization? How do varying funding models for peer-review impact how researchers and agencies fund research? Are journals capable of adapting to meet new challenges presented by the digital age, transparency demands, and competitor models?

As you can see I have a lot of questions, but I think one thing is clear. Any effort to fundamentally change the way in which federal research is reviewed, vetted, transmitted, and communicated should benefit from the Science, Space, and Technology Committee's input. We have been involved in investigating issues surrounding public access for a number of years, and are uniquely qualified to evaluate the impacts on research and federal agencies. Rep. Gordon, the former Committee Chairman, brought together a number of stakeholders in 2009 in order to find common ground, and in 2010 the "Scholarly Publishing Roundtable" issued a report containing several recommendations. We also tasked the Office of Science and Technology Policy to address this issue in the America COMPETES Reauthorization Act of 2010 and expect their results soon. Today's hearing is an extension of this longstanding engagement.

I look forward to working with all of the interested parties, and want to thank our witnesses for appearing today.

Chairman BROUN. I now recognize my good friend, the Ranking Member from New York, Mr. Tonko, for his opening statement.

Mr. Tonko, you are recognized for five minutes.

Mr. TONKO. Thank you, Mr. Chair.

And I want to thank all of our witnesses for coming today to testify before our subcommittee. And further, let me thank you in advance for your patience. The schedule is going to be a little awkward today.

In 2010, this committee adopted language that set the stage for enhancing public access to federally funded research. Since that time, two legislative proposals have emerged and have received the scientific community's attention. At this point, I believe the language this committee included in the reauthorization of the *America COMPETES Act* remains the best path forward.

We have two competing public interests that play in the open access discussion. On the one hand, the taxpayers who provide support for research through grants provided by federal science agencies have an interest in having the research they fund deliver maximum public benefit. On the other hand, the public is not only interested in quantity; they want quality.

The scientific publishing enterprise, working with the research community, academia, and the government traditionally has had an important role in ensuring that quality through the management of the peer-review process. We are on the cusp of an information revolution. We hear about the impacts of this revolution on print media every day. Digital technologies are making it easier to produce, distribute, search, access, and archive information. Scholars are embracing open-access journals for their articles and we have seen a series of schools follow the Harvard faculty's model in

establishing publicly accessible faculty archives of their publications.

More journals are looking to move to a business model based on author charges to remain viable with the move to open access. The NIH's policy of making final submitted copies of manuscripts that result from NIH-funded research available on PubMed within 12 months of publication. Is also driving change amongst publishers. The National Science Foundation and the Department of Energy, significant funders in all nonmedical fields of science are both working on pilot programs that embrace an open access policy. With all of these changes underway, it is hard to see how traditional publishers will be able to survive without significantly rethinking their business model. This appears to be true for both the for-profit and not-for-profit publishers.

So what is the proper role for the government in this evolutionary process? Clearly, we need to ensure that vital public interests in the United States research enterprise are indeed served. Any new system that emerges needs to facilitate data sharing and interoperability across fields and archives. In addition, it must provide for the long-term stewardship of the scientific record. We need to consider the implications for federal grants of moving to new publishing business models, including author-paid publication.

Currently, federal grants help to support journal subscriptions through indirect costs on grants. Can we rely on the current policy path coupled to changing technology, emerging competition, and social norms among scholars to drive us steadily toward broader access to research results? It is too soon to tell. The landscape is dynamic and federal agencies have not yet completed their policy reviews and revisions.

We are still waiting for the report we requested from OSTP. If we proceed to a legislative approach, we may end by creating more problems than we indeed solved. An abrupt end to the current system could drive some publishers out of existence. It could result in the loss of established journals and weaken professional scientific societies. These outcomes would be counterproductive to the goal of having high-quality research widely published and disseminated. How we produce, share, and preserve knowledge is on the edge of the greatest change in four centuries. There are many new opportunities for improvement and this is indeed an exciting time. Transitions are always unsettling but they offer a period for constructive experimentation.

I believe we should take the time to hear from all interested parties, encourage the federal science agencies to move their efforts forward and refrain at this time from prejudging the best outcome through prescriptive legislation. We want to see the strongest possible system for sharing and preserving knowledge emerge from this transition period, and I look forward to hearing the perspectives and the concerns of our witnesses today and as this process moves forward.

Thank you very much, Mr. Chair.

[The prepared statement of Mr. Tonko follows:]

PREPARED STATEMENT OF SUBCOMMITTEE RANKING MEMBER PAUAL D. TONKO

Thank you, Mr. Chairman.

I want to thank all of our witnesses for coming today to testify before the Subcommittee.

In 2010, this Committee adopted language that set the stage for enhancing public access to Federally-funded research. Since that time, two legislative proposals have emerged and have received the scientific communities' attention.

At this point, I believe the language this Committee included in the reauthorization of the American Competes Act remains the best path forward.

We have two competing public interests at play in the open access discussion.

On the one hand, the taxpayers who provide support for research through grants provided by Federal science agencies have an interest in having the research they fund deliver maximum public benefit.

On the other hand, the public is not only interested in quantity. They want quality. The scientific publishing enterprise--working with the research community, academia, and the government traditionally has had an important role in ensuring that quality-through the management of the peer review process.

We are on the cusp of an information revolution. We hear about the impacts of this revolution on print media every day. Digital technologies are making it easier to produce, distribute, search, access, and archive information.

Scholars are embracing open access journals for their articles, and we have seen a series of schools follow Harvard faculty in establishing publically-accessible faculty archives of their publications.

More journals are looking to move to a business model based on author charges to remain viable with the move to open access. NIH's policy of making final submitted copies of manuscripts that result from NIH-funded research available on PubMed within 12 months of publication is also driving change among publishers.

The National Science Foundation and the Department of Energy—significant funders in all non-medical fields of science-are both working on pilot programs that embrace an open access policy.

With all these changes underway, it is hard to see how traditional publishers will be able to survive without significantly re-thinking their business model. This appears to be true for both for-profit and not-for-profit publishers.

So, what is the proper role for the government in this evolutionary process? Clearly, we need to ensure that the vital public interests in the U.S. research enterprise are served.

Any new system that emerges needs to facilitate data sharing and interoperability across fields and archives. In addition, it must provide for the long-term stewardship of the scientific record.

We need to consider the implications for federal grants of moving to new publishing business models including author-paid publication. Currently, federal grants help to support journal subscriptions through indirect costs on grants.

Can we rely on the current policy path, coupled to changing technology, emerging competition, and social norms among scholars to drive us steadily toward broader access to research results?

It is too soon to tell. The landscape is dynamic and Federal agencies have not yet completed their policy reviews and revisions. We are still waiting for the report we requested from OSTP.

If we proceed to a legislative approach, we may end by creating more problems than we solve. An abrupt end to the current system could drive some publishers out of existence. It could result in the loss of established journals and weaken professional scientific societies. These outcomes would be counterproductive to the goal of having high quality research widely published and disseminated.

How we produce, share, and preserve knowledge is on the edge of the greatest change in four centuries. There are many new opportunities for improvement and this is an exciting time. Transitions are always unsettling, but they offer a period for constructive experimentation.

I believe we should take the time to hear from all interested parties, encourage the Federal science agencies to move their efforts forward, and refrain at this time from prejudging the best outcome through prescriptive legislation.

We want to see the strongest possible system for sharing and preserving knowledge emerge from this transition period. I look forward to hearing the perspectives and concerns of our witnesses today and as this process moves forward.

Chairman BROUN. Thank you, Mr. Tonko.

If there are any other Members who wish to submit additional opening statements, your statements will be added to the record at this point.

Chairman BROUN. At this time, I would like to introduce our panel of witnesses—Dr. H. Frederick Dylla, the Executive Director and CEO of the American Institute of Physics; Mr. Elliot Maxwell, Project Director for the Digital Connections Council on the Committee on Economic Development. Boy, that is a mouthful. Dr. Crispin Taylor, the Executive Director of the American Society of Plant Biologists; Mr. Stuart Shieber, the Director of the Office for Scholarly Communication at Harvard University; and Mr. Scott Plutchak, the Director of Lister Hill Library at the University of Alabama in Birmingham.

As our witnesses should know, spoken testimony is limited to five minutes each, after which the members of the committee will be—will have five minutes each to ask questions. Your written testimony will be included in the record of the hearing.

It is the practice of this Subcommittee on Investigations and Oversight to receive testimony under oath. Do any of you have an objection to taking an oath? Okay. I see everybody's head shake side-to-side indicating "no." So let the record reflect that all witnesses are willing to take oath.

You may be represented by counsel. Do any of you have counsel here today? Also everyone indicated that they do not have counsel, so let the record reflect as such.

If all of you would now please stand and raise your right hand. Do you solemnly swear or affirm and tell the whole truth and nothing but the truth, so help you, God?

Thank you. You may be seated.

Let the record reflect that all the witnesses participating have taken the oath.

Now, I recognize our first witness, Dr. Dylla.

STATEMENT OF H. FREDERICK DYLLA,
EXECUTIVE DIRECTOR AND
CHIEF EXECUTIVE OFFICER,
AMERICAN INSTITUTE OF PHYSICS

Dr. DYLLA. Chairman Broun, Ranking Member Tonko, and members of the committee, good morning. I am Fred Dylla, the Executive Director and CEO of the American Institute of Physics. AIP is a not-for-profit publisher of scholarly research and an umbrella organization of over ten scientific societies that includes over 135,000 scientists. Before taking on this job five years ago, I was a practicing physicist, so I am keenly aware that innovation and discovery depend on access to research.

Today's hearing is addressing the essential question of the appropriate role of government in facilitating public access to taxpayer-funded research. Recently, there has been intense debate over public access, which is the ability for anyone to readily access scholarly research. All parties in this debate agree that enhanced public access is a good thing, is good for the economy; it is good for science.

My main message to you today is that the current system of scholarly communications is working. More journals are reaching more readers in more usable formats than ever before. Stakeholders have been able to make significant strides toward increasing access even while recognizing there are no quick fixes where one size fits all models.

As the public demand for free electronic information grows, however, the debate on how to pursue public access has intensified. Financial concerns lie at the heart of this issue. The argument is if the taxpayers fund the research, shouldn't the public have access to it? And the answer is of course. But in the world of scholarly publishing, freely available does not necessarily mean free of cost. It takes significant investments for publishers to give value to that research and communicate it in a meaningful way.

For AIP this amounts to about $2,000 per article, which includes the rigorous peer-review process that we manage, the editing, the production, permanently archiving, and much more. All the parties agree that these components add essential value and enable scientific progress. This committee has weighed in on this issue and your solution is working. Last year, legislation was enacted, Section 103 of the COMPETES Act, that calls for all stakeholders, including publishers, to work together to increase public access to scholarly publications, grantee reports and associated data.

Since that time, important collaborations have formed and we have made significant progress. Partnerships with the National Science Foundation and the Department of Energy have produced pilot programs that link grantee reports to publications, support data mining across agency and publisher platforms, and provide new tools and methodology for identifying publicly funded work. Publishers have the expertise to drive these projects, saving the funding agencies money. Innovative publication technology, efficient business practices, and connectivity to the international publishing community enable the government to show results quickly, increasing taxpayer access to publicly funded work.

Due to the complex and nuanced nature of this issue, proposals that significantly distort the current system of scholarly communications could even be harmful. Recently, the two most touted models—one-size-fits-all and mandated open access—divert the progress that has been made over the past few years and could even jeopardize the ultimate goal of enhanced public access.

How are these proposals harmful? First, a single business model that mandates free public access does not take into account the rich diversity in both the science and scientific publishing. A successful approach in one scientific discipline may unfairly hinder another. This could jeopardize many sectors of the $3 billion-a-year component of the U.S. publishing industry, which employs 30,000 people in this country. But secondly, a blanket approach that imposes open access in all journals would significantly impede our approach and other publishers' abilities to ensure quality and the value of published scholarly research.

Fortunately, the current system is working. The natural pressures of the marketplace continue to foster innovative products and dissemination methods. Within the scholarly publishing realm, new publishers, new journals, new business models are continually emerging signaling a healthy, competitive marketplace. These innovations date back to our Nation's birth with Ben Franklin—a scientist, a publisher, an entrepreneur who exemplifies our country's tradition of scholarly publishing and free enterprise. Today, the government has given us the tools to continue this tradition so we can pursue innovation and foster scientific discovery.

Within the current framework, I am confident that public access to all the results of scientific research will continue to flourish and grow.

Thank you. I am honored to be among such a distinguished committee and panel of witnesses.

[The prepared statement of Mr. Dylla follows:]

Federally Funded Research:
Examining Public Access and Scholarly Publication Interests

Written testimony submitted to the
House Committee on Science, Space, and Technology,
Subcommittee on Investigations and Oversight

March 29, 2012
Rayburn House Office Building, Room 2318
10 a.m. EDT

H. Frederick Dylla, Ph.D.
Executive Director and CEO
American Institute of Physics

Executive Summary

The American Institute of Physics (AIP) appreciates the opportunity to submit testimony as part of the House Committee on Science, Space and Technology, Subcommittee on Investigations and Oversight hearing on "Federally Funded Research: Examining Public Access and Scholarly Publication Interests."

AIP strongly supports the broad dissemination of scholarly research, which includes public access to journal articles, data, and related information. It is our position that the America COMPETES Reauthorization Act of 2010 (P.L. 111-358) contains the best and most effective framework to broaden public access. Indeed, recent collaborative efforts by AIP, other scientific societies, the White House Office of Science and Technology Policy, the Department of Energy, and the National Science Foundation have made significant progress toward these goals under the comprehensive structure outlined in America COMPETES.

Due to the complex and highly nuanced nature of this issue, new legislation or government policies that force free public access through a single dissemination business model would harm the collaborative process that stakeholders have worked so hard to achieve. Moreover, a blanket approach to public access would diminish the quality and value of published scholarly research and actually detract from achieving the goal of increasing access to scholarly literature.

Notes on Written Testimony Structure: The bulk of this written testimony is attached in the form of AIP's December 2011 responses to two Requests for Information (RFIs) from the White House Office of Science and Technology Policy (OSTP), which capture comprehensively AIP's positions on the key policy and technical issues in the public access debate.

About AIP and Issue Relevance

AIP is a 501(c)(3) not-for-profit membership corporation founded in 1931 for the purpose of promoting the advancement and diffusion of the knowledge of physics and its application to human welfare. AIP collaborates with its ten Member Societies to provide resources for activities such as scholarly publishing and outreach to the science community and the general public. Publishing scientific journals is the primary means by which scientific societies communicate advances in research to the community. Publishing is also AIP's primary source of revenue, supporting its outreach activities, which serve the broad physics community and the general public.

AIP is an umbrella organization of ten Member Societies that collectively represent a broad cross-section of more than 135,000 scientists, engineers, and educators in the global physical science community. With an extensive catalog of top-cited journals, AIP is one of the world's leading publishers in the physical sciences. AIP publishes 13 journals; two magazines, including its flagship publication *Physics Today*; and the AIP Conference Proceedings series. In keeping with its goal to increase access to and use of its journals, AIP reinvests its journal revenue in innovative electronic publishing technologies for scholarly journals and offers full-solution publishing services for many of its Member Societies.

In 2011, AIP published more than 15,000 scholarly articles in its journals. AIP also published more than 5,000 additional articles for its Member Societies for which it is the publisher of record (American Association of Physics Teachers; Acoustical Society of America; American Association of Physicists in Medicine; AVS: Science & Technology of Materials, Interfaces, and Processing; and The Society of Rheology).

Four AIP Member Societies (American Astronomical Society, American Geophysical Union, American Physical Society, and The Optical Society) manage their own publications, which collectively published an additional 39,000 articles in 2011. AIP and its Member Societies published approximately one quarter of the articles in the discipline of physics (approximately 240,000) in 2011.

AIP's journal revenues finance its entire publishing process, which includes highly skilled editorial management employing over 130 international scientific editors; end-to-end manuscript oversight with authors; coordination of the essential peer review process (which ensures scientific integrity); translation of the text and figures into the form visible on the published page; final publication (both in the traditional print journal and online and mobile versions); the development, dissemination, and maintenance of searchable and accessible journal databases; and preservation of the digital version throughout the coming decades. This enterprise requires extensive human and capital resources: for its archival journals, AIP expends more than $40 million annually for the entire operation, including editorial, production, bibliographic tagging, printing, online-hosting, and archiving tasks. AIP employs more than 200 people at our facilities in Melville, N.Y., and College Park, Md., who manage and support our publishing operations.

AIP uses the net revenues from its journal publishing operations to support its Physics Resources Center. This Center provides a variety of outreach services for the scientific community and the general public, including: media services for translating summaries of journal articles for dissemination in lay language media channels; operation of the Niels Bohr Library and Archive for preserving the history of the physical sciences; the Statistical Research Center, which tracks education and workforce statistics for physical scientists; a comprehensive science news service for educating the general public about advances in research; and administration of the Society of Physics Students on more than 700 university and college campuses [see AIP's Annual Reports for details: http://www.aip.org/aip/reports.html]. The $21 million in operating expenses to run these programs is partially offset by net revenues from AIP's journal publishing operations.

About Dr. Dylla

Since 2007, H. Frederick Dylla has served as Executive Director and CEO of AIP. Previously, Dylla served as Chief Technology Officer and as Associate Director at the Department of Energy's Jefferson Lab, where he spearheaded the Free Electron Laser (FEL) program. From 1975 to 1990, he held various positions at the Department of Energy's Princeton Plasma Physics Laboratory, where he helped develop technology for nuclear fusion reactors, particle accelerators, and materials processing for the microelectronics industry. He received his Ph.D. in physics from MIT (1975), is a Past President and Fellow of the American Vacuum Society

(AVS), and a Fellow of the American Physical Society and the American Association for the Advancement of Science. During his scientific career, Dylla has been an author of more than 190 peer-reviewed publications in scholarly literature and has served as an editor and on the editorial board of several journals and monograph series. He presently serves on the boards of the Professional and Scholarly Publishing Division of the Association of American Publishers and the International Association of Scientific, Medical and Technical Publishers. In 2009, Dylla helped organize and participated in the Scholarly Publishing Roundtable under the aegis of the U.S. House Science and Technology Committee. The Roundtable developed consensus recommendations for the development of public access policies for scholarly data and publications.

Access to Journal Content

Scholars and research scientists access AIP published content in very large numbers through several different channels, including subscription or license fees which are generally paid by their institutional libraries. AIP's subscription prices are competitively priced within the physics publishing market. With more than 13 million full-text downloads in 2011, the subscription-based cost per download for AIP content is in the $2-3 range, representing outstanding value for journals that are in the top rank of their class in terms of scholarly impact. AIP has also responded to budget challenges faced by the library market by offering other access models such as a low-price article rental program that has since been adopted by more than 40 other scholarly publishers. Overall, AIP provides a number of cost-effective and efficient means to access high quality peer reviewed content. Some of these access options are detailed further in the attachments.

AIP has joined a diverse group of journal publishers that make their articles freely available to academics and others in 100 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.

Notably, one of AIP's Member Societies, the American Physical Society, spearheaded an initiative that allows public access to all of their journals by making them available at no charge through public libraries and high schools around the country. So far, more than 600 libraries have signed up for this service.

Many librarians have become advocates of open access in response to cost pressures induced by the rapid growth in journals, proliferation of new journals in niche subjects, and the high relative price of some journals. Since library subscribers are AIP's most important customers, AIP and its Member Societies are very sensitive to their concerns and believe that the inclusive framework established under the COMPETES law can help address their concerns without threatening the quality and essential services to the scholarly community provided by scholarly publishing.

Rapidly Changing and Vibrant Marketplace
Within the scholarly publishing realm, new publishers, journals, and business models are continually emerging, signaling a healthy, competitive marketplace. It is AIP's belief that the government should support and encourage this diversity through its actions and policies via mutually beneficial partnerships with publishers, which would contribute to the U.S. economy and maximize the productivity of the scientific enterprise. This ability of scientific publishers to experiment with different publication, business, and access models is essential and assures the vitality, diversity, and effectiveness of the scholarly communication marketplace, leading to scientific and technological advances. This tradition of innovation in communications in the free market is a hallmark of the intersection of research, entrepreneurialism, and publishing going back to the earliest days of our nation.

Freely Available, But Not Free of Cost
AIP understands the enthusiasm for open access for the obvious reason that it increases access to research, which is at the core of our mission. AIP has been trying to build awareness among all affected parties that while open access may mean freely available, the costs to assure the quality, rigor, discovery, and production value of scientific publishing are not zero.

As Maria Leptin, the Director of the European Molecular Biology Organization, wrote in a March 16, 2012 editorial in the journal *Science*: *"Any transition to open access on a large scale will require a clear understanding of the financial challenges that will be faced. Put simply, publishing costs money, and open access does not mean 'for free' – someone must foot the bill."*

Currently, more than 25,000 scholarly journals are being published worldwide, and institutional subscriptions generate income for 90 percent of these titles. For most of the remaining 10 percent, authors or sponsoring agencies pay an upfront fee per article. These articles are posted on the web without subscription barriers as soon as they are published.

Potential Legislative and Policy Impacts
Current policy efforts to increase public access have focused on two approaches: (a) accelerate the transition from the subscription model to the open access model or (b) mandate the release of subscription content after a specified embargo period. The NIH Public Access Mandate, which was introduced in 2008, requires scholarly articles to be posted on NIH's PubMedCentral website 12 months after publication, if any of the authors had NIH funding for any portion of the underlying research reported on in the articles. The proposed Federal Research Public Access Act (FRPAA), introduced in both houses of Congress last month would extend the release mandate to all disciplines represented by the eleven federal agencies that fund research and decrease the timeframe to 6 months.

It is AIP's position that neither of these one-size-fits-all approaches is an appropriate solution for the diverse array of journals published across all the disciplines represented by federally funded research efforts. The open access model is growing at a reasonable rate for fields where such a model is appropriate (i.e., well-funded or fast-moving disciplines). Delayed-release models are not viable for fields where articles have citation lifetimes of years, such as mathematics, theoretical physics, and the social sciences. Additionally, the scholarly community

should determine the methods of dissemination that are appropriate to their fields. Forcing the adoption of either model would likely cause significant harm to the enterprise of scholarly publishing. Furthermore, forcing the adoption of these models is not necessary, given the natural pressures of the marketplace that continually drive the industry to evolve and innovate a wide array of products and dissemination methods.

Agency/Publisher Pilot Projects Launched

The most appropriate role for the federal government is to encourage federal agency/publisher partnerships, examples of which have arisen as a direct result of COMPETES. 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.

Specifically, the DOE Office of Scientific and Technical Information (OSTI) is collaborating with journal publishers to improve DOE's ability to demonstrate the outcomes of the research it funds. This involves engaging with publishers to identify and broaden access to the journal articles reporting on research funded by DOE. To this end, OSTI has embarked on a pilot project to enhance journal article full-text searching, with the intent to make citations of DOE-funded journal articles available in the search and retrieval applications operated by OSTI.

In this pilot project, the scholarly publisher Wiley provides citations to OSTI, including abstracts and hyperlinks to a landing page for the publisher's version of the article. Wiley provides the full text of the article for use in OSTI's archive, which improves search precision and recall. Through this existing infrastructure OSTI would make the journal publisher's full text searchable.

One of the most important tasks that the publishing community has undertaken (funded by subscription and other revenue) is ensuring the proper tagging of articles, verification and disaggregation of author names, and references. An initiative in this vein is the FundRef project—a collaboration involving CrossRef (a Digital Object Identifier registration agency), several funding agencies and publishers to establish protocols for identifying the funding agencies associated with journal articles. OSTI is working with CENDI (a federal scientific and technical information managers group) members to have a standard list of agencies so there will be no ambiguity for agency names.

CrossRef has agreed to add two new metadata elements to the CrossRef database, which would apply to each journal article: funding agency identity and grant number. FundRef pilot demonstrations with at least five major publishers including AIP will be deployed next year.

Similarly, officials at NSF are assuming a leadership role in initiating one-to-two-year pilot projects on expanded public access to research results; these involve universal identifiers for better search results and linking between NSF and publisher databases.

The NSF Directorate for Math and Physical Sciences is in discussions with AIP and the American Astronomical Society to establish a pilot project to link the data behind figures and tables with publications.

Based on these valuable experiences, other agencies and publishers can gain valuable insight on how future collaborations might be structured to promulgate the success of these initial agency-publisher partnerships.

The America COMPETES Act requires the U.S. federal agencies that fund scientific research to develop policies for access to and interoperability among databases, and archiving for data and publications that are derived from public funding. Publishers have valuable expertise that can help in this process, but only if publishing continues to be sustainable. Such collaboration between publishers and the government is already happening in ways that will increase public access to reports, data, and publications derived from federally funded research. These collaborative initiatives create efficiencies and cost savings for the funding agencies.

Conclusion

AIP believes that a uniform access policy or mandate for scholarly publications would be an ineffective approach. An overarching government-wide policy that would simply mandate a short publication embargo period would fail to take into account such key factors as the specific needs of any given agency, the rapidly changing marketplace and nature of scholarly publishing, and the unique considerations of the various fields of science and the journals that serve them.

The creative and thoughtful discussions that have been spurred by the existing America COMPETES law, organic market forces, and collaborative efforts already underway between publishers and several federal agencies, offer a pragmatic and productive route to success in broadening public access to the all the products of federally funded research: grantee reports, associated data, and the resulting peer-reviewed publications. These efforts reinforce the view of many in the scholarly publishing community that new legislation is not needed at this time.

I believe—and the evidence from the post-COMPETES partnerships shows—that we are making real progress on the interrelated issues of access and interoperability among public and private information platforms and databases.

Chairman BROUN. Thank you, Dr. Dylla.

Mr. Maxwell, you are recognized for five minutes.

STATEMENT OF MR. ELLIOT E. MAXWELL, PROJECT DIRECTOR, DIGITAL CONNECTIONS COUNCIL, COMMITTEE ON ECONOMIC DEVELOPMENT

Mr. MAXWELL. Thank you, Mr. Chairman, Ranking Member Tonko, Members of the Subcommittee. First, let me express my appreciation for the opportunity to testify today. My testimony today is based on a report, "The Future of Taxpayer Funded Research: Who Will Control Access to the Results?" which I wrote under the auspices of the Committee for Economic Development with generous support from the Ewing Marion Kauffman Foundation. That report addresses the cost and benefits of the public access policies of the National Institutes of Health, as well as proposals to overturn or extend that policy. I would ask that a copy of the report be inserted into the record.

Chairman BROUN. Certainly. Thank you. So ordered.

[Visit http://www.emaxwell.net/linked/DCCReport—Final—Feb2012.pdf to view the report.]

Mr. MAXWELL. For the last several years, I have served as Project Director of CED's Digital Connections Council, and during this time, the Council has made several reports on how greater openness is enabled by the Internet and by the digitization of information. This kind of openness is in fact exemplified by the NIH public access policy. It makes it possible and it benefits from the availability of the articles that are made available.

The policy has been in effect for nearly four years, but there are disagreements about its effect. Supporters have argued that it increased public access, substantial positive impacts on the progress of science, innovation, and economic growth and should be extended. Opponents are primarily, but not entirely, publishers of proprietary journals. And I want to make sure that people understand that there are publishers that are not proprietary but are open-access publishers. And so when people talk about publishers as if they were a single body, it is not an accurate way of depicting what the situation is.

Primarily, publishers of proprietary journals have argued that increased public access will or have damaged the subscription-supported publishing business and will undercut peer review and may force journals to close or reduce the number of outlets. CED initiated this report because we wanted to try to get what the facts were, what the effects were, and rather than speculate about these things, to look at what is on the record after four years. And we think we have a kind of disinterested view of this because we are not in favor of one or another kind of dissemination. But the central question of the report was, how does the NIH policy and proposals to extend it or to overturn it affect the production and dissemination of high-quality science research. That seems to be the single most important thing to remember, that it is about the progress of science, innovation, and economic growth. It is not about the effect on any particular mode of dissemination. And so we wanted to look at it from that standpoint.

Now, we were concerned about the effects on different ways of disseminating research because only to the extent that they affect the production and dissemination of high-quality research. And quality, as Mr.—as the Chairman and Ranking Member have indicated, is an important issue but it is not necessarily equated with proprietary publishing. It is related to peer review and the different means of assuring quality. And we want to underline that.

It is worth noting, as Fred has said, that all the parties agree about the importance of access. Proprietary publishers make a point in their arguments about their efforts to increase access to groups that are underserved. I would suggest that the question in some ways needs to be flipped around. What all parties also agree is that the NIH policy for public access has increased access. There is no dispute about that. So the question you face, I think, is to say we have a means and we know how to use it of increasing access. The issue now should be what do we need to do to limit that as opposed to how do we extend what we have now in the scholarly publishing journal to other federal agencies? Because we have both the means and the knowledge of how to increase access and we should start there and back off from that as opposed to saying how do we tweak what we have in those areas beyond the NIH?

So it is not, I think, a matter of dispute about the value of access or a matter of dispute about the success of NIH. It is really only about, I think, the details of how it might be worked out in other areas. But we start with the presumption of access as opposed to the presumption that the publishers control access. And because we have a time of an extraordinary pressure on our budgets, we need to be thinking about how to get the most from the very large scientific investment by taxpayers in support of research.

What we have found in looking at the literature was that increased public access accelerates progress in science by speeding up and broadening diffusion of knowledge, providing better access to more people—because as more people have access, more people can do something with it—increases the range and quality and variation of experiments that lead to better solutions. The more people have access, the different perspectives they bring more likely to have solutions than to narrow that. We have found that it has important benefits to authors because their work is more available and it increases the efficiency of the research enterprise because you can see what is being done and you can get as much as you can without making taxpayers pay twice for the research both in terms of the grants and in terms of subscriptions.

We have looked at all of the evidence that we have about the impact on the scholarly publishing community and we found no evidence that public access as provided by the NIH policy has harmed the subscription journals. We found no persuasive evidence that increased public access threatens the ability to peer review. We found no persuasive evidence of a significant reduction in traditional publishing. We found only a small effect on the rate of growth of profits on the proprietary publishers, and that was in the midst of the largest recession we have had in years and years and years. None of even the——

Chairman BROUN. Mr. Maxwell, if you would go ahead. You are a little over time now and we are——

Mr. MAXWELL. I am sorry.

Chairman BROUN. —pressed for our time because we are going to have votes very quickly and I just want to try to get through as much of this testimony. So if you would summarize very quickly, I would appreciate it.

Mr. MAXWELL. Absolutely. I think that we should recognize the very large benefits that exist and start from the premise of increasing access, and then, if necessary, dial that back. And we should think of the experience that people have had in the NIH and related biomedical world and start to move down to a shorter embargo, recognizing that there may be differences amongst the various disciplines.

[The prepared statement of Mr. Maxwell follows:]

Testimony of Elliot E. Maxwell

Director, Digital Connections Council
Committee for Economic Development

March 29, 2012

U.S. House of Representatives Committee on Science, Space and Technology
Subcommittee on Investigations and Oversight
2318 Rayburn House Office Building
Washington, DC 20515

TESTIMONY OF ELLIOT E. MAXWELL BEFORE THE SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT OF THE COMMITTEE ON SCIENCE, SPACE AND TECHNOLOGY THE UNITED STATES HOUSE OF REPRESENTATIVES MARCH 29, 2012

Mr. Chairman and members of the Subcommittee,

First let me express my appreciation for the opportunity to testify today. The interest of the House Science Committee in this area is longstanding and deserves commendation. I am honored to be invited to share in the efforts to ensure that the taxpayers of this country obtain the greatest possible return on their investment in federally funded research.

My testimony today is based on a report, "The Future of Taxpayer Funded Research: Who Will Control Access to the Results?" which I wrote under the auspices of the Committee for Economic Development (CED) with generous support from the Ewing Marion Kauffman Foundation. That report addresses the costs and benefits of the public access policies of the National Institutes of Health (NIH) as well as proposals to overturn or extend that policy; it does not analyze or make recommendations regarding particular legislative proposals. I would ask that a copy of this report be included in the hearing record.

I should make clear that I am testifying as the author of the report but do not speak for CED or any of its members or funders; the opinions I express today are my own. The Ewing Marion Kauffman Foundation exercised no control over the research or the findings, and neither CED nor I have any financial interest in the conclusions or recommendations.

SOME BACKGROUND ON CED AND THE DIGITAL CONNECTIONS COUNCIL (DCC)

CED is a non-profit, non-partisan business-led public policy organization. CED conducts research on major economic and social issues and actively informs and engages the business community in an effort to achieve policy reform for the good of the nation. Membership is made up of some 200 senior corporate executives and university leaders who lead CED's research and outreach efforts.

For the last several years I have served as project director of CED's Digital Connections Council (DCC), which is chaired by Paul Horn, former IBM Senior Vice President for Research and currently New York University Distinguished Scientist in Residence & Senior Vice Provost for Research. During this time CED has issued several DCC reports on how greater "openness" made possible by the digitization of information and the growth of the Internet can lead to increased benefits to society and how it can improve specific domains such as healthcare and higher education.

"Openness" as used in these reports can be thought of as a continuum from completely open--such as something posted to the World Wide Web and available to all without any restrictions--to completely closed--such as a formula written down but kept under a pillow and never shared. Openness has two aspects: accessibility and responsiveness. To the degree that information or processes are accessible--e.g. are available without need to pay a subscription or for the recipient to be at a particular place--they are more open. And to the extent that what is accessible is responsive--e.g. can be repurposed, and reused--they are more open. While achieving greater openness has many positive rewards, these reports all have stressed that careful thought should be given to determining the right degree of openness for the particular situation. Electronic health records, for example, should be open to all the medical personnel providing treatment to an individual, but not open to a landlord, and should be responsive to reports of lab test results but not alterable by anyone not authorized to do so.

The results of greater openness made possible by digitization and the growth of the Internet can be seen in the rise of open source software, the development of open educational courseware, the emergence of open innovation, the global scientific collaboration in the Human Genome Project (HGP) and the immediate announcement of its results, and, most importantly for today's hearing, the NIH public access policy which was the subject of the report on the "Future of Taxpayer-Funded Research."

SOME BACKGROUND ON THE NIH PUBLIC ACCESS POLICIES

Progress in science is built upon the work of those who came before. Demonstrations of such progress in research could be found in published scientific journals that for several hundred years have been among the most important vehicles for the dissemination of new scientific knowledge.

Until very recently modern scientific journals were funded almost entirely by institutional subscriptions; subscribers such as libraries and their users had access to new research results while others without subscription access had to wait for other means of knowledge diffusion to have access to this new knowledge. With the adoption of the NIH public access policies, an alternative and complementary model for access to NIH funded research results was given an enormous boost.

The U.S. National Institutes of Health is the largest single funder of biomedical research in the world with a budget of $31 billion that, through its grants making process, generates 90,000 articles each year. Since 2008 NIH' s public access policy has required that its grantees place a copy of their peer-reviewed manuscripts accepted for publication by a scientific journal in PubMed Central (PMC), an online digital repository open to all; the work would be available no later than 12 months after the version of record is published. In 2005, NIH had asked grantees to deposit their work voluntarily. By 2008 only a small percentage of grantees--7%--had done so and NIH made deposit mandatory, a policy decision based on instructions to NIH in a law passed by Congress and signed by the President.

PMC now includes more than 2.3 million full text manuscripts and articles and over a thousand journals now voluntarily deposit all of their articles into the database, whether or not they deal with NIH funded research. Over 500,000 unique visitors access PMC on a typical workday.

The policy has been in effect for nearly four years but there are disagreements about its impact. Supporters of the policy have argued that the increased public access has substantial positive impacts on the progress of science, innovation and economic growth, and should be extended to cover extramural research funded by other major federal funders of unclassified scientific, technical, and medical (STM) research. Opponents, primarily but not entirely, publishers of proprietary STM journals, have argued that the increased public access has or will damage their subscription supported publishing businesses and, by so doing, will undercut the peer-review system. They have stated that the financial pressure may force publishers to close, and as a result, reduce the amount and quality of research by reducing the number of outlets for research, while at the same time undercutting their copyright interests. Publishers also argue that the manuscripts authored by researchers are not the direct result of the NIH funding and that NIH should be making public the reports filed by grantees pursuant to federal regulations.

SOME BACKGROUND ON THE "THE FUTURE OF TAXPAYER-FUNDED RESEARCH; WHO WILL CONTROL ACCESS TO THE RESULTS?"

CED initiated the report because the differing assertions about the impact of the NIH public access policy and the different proposals to extend or overturn it raised important public policy issues and were related to earlier efforts by the DCC. Unlike the debates over the policy around the time of its adoption, there were now nearly four years of experience with the policy on which to base conclusions.

I'd like to make one point absolutely clear. The report is focused on the question of how the NIH public access policy, and its potential extension to other federally funded extramural research--or its reversal--might affect the development and dissemination of high quality scientific research and its benefits to our society.

The impact of the policies on proprietary publishers (for profit or not for profit) or open access publishers (publishers that rely on author payments rather than subscriptions) digital repositories or any particular means of disseminating knowledge was important (for the purposes of this analysis) only in so far as the impact would affect the development and dissemination of high quality research. This focus was chosen because the impact of the policy on the production and dissemination of knowledge is the central issue and of high public importance given the very substantial taxpayer expenditures in support of research and the enormous public benefits that can be gained from that research.

The report does not dwell on the fundamental importance of research to innovation and economic growth; that case has been made well by generations of economists, scientists, and others and is supported by both proponents and opponents of the NIH public access

policy. The report does look at the costs and benefits of increased public access to research results through the lens of "openness" with a particular interest in how greater public access (and greater openness) affects progress in science, the productivity of the research enterprise, the process of innovation, the commercialization of research, and ultimately economic growth.

THE SHARED AGREEMENT ON THE ISSUE OF INCREASED PUBLIC ACCESS

It is worth noting that all the parties involved in the debate about the NIH public access policy support the concept of greater public access to the results of scientific research. Even opponents of the policy who believe that there is not a current problem with public access which would justify public policy intervention have argued that they are already making changes to increase public access where there might be a problem, for instance with regard to patients and others interested in the literature on a particular medical condition, or for scientists from less developed countries who could not afford to pay current subscription rates. So, increased public access is, by all accounts, of benefit to the society absent some compelling cost or other counter argument.

It is also worth stressing that no one disputes that the NIH public access policy has already substantially increased public access to the results of NIH funded research from what was previously available from traditional STM publishers.

THE BENEFITS OF INCREASED PUBLIC INTEREST

While all parties agree that increased access is of benefit, it is important to understand exactly why it is beneficial. The report examined this question and reviewed the most current literature to see what could be learned. The report found specific benefits in four major areas:

Increased public access accelerates progress in science by speeding up and broadening diffusion of knowledge. This was of some benefit to researchers in the field covered by a journal allowing them to get to the frontiers of knowledge more quickly. But these researchers in a field were more likely to have access through subscriptions as they constitute the target audience for the journal. But increased access benefited others who do not have equally good access to recent research results such as researchers in the field at less well resourced institutions, researchers in other fields, clinicians and patients, and those in the general public who are able to contribute to scientific and technological development. The report noted, in particular, the benefit of access to those in the private sector developing new goods and services who rely on access to scientific research and who report considerable difficulties obtaining access. (The report provides ample evidence of the lower levels of access experienced by such groups and the problems that they encounter without access such as provided by the NIH public access policy.)

Providing better access to this expanded group of readers has important benefits. As Fiona Murray of MIT and her colleagues have pointed out, expanding access increases the number of, and the diversity of, potential follow-on researchers.

This leads to the exploration of a larger number and a wider variety of research paths and experiments to find solutions, increasing the likelihood of success. Increased public access which leads to more and more varied follow-on research also leads to faster movement from basic research to applied research.

Heidi Williams, also of MIT has shown, in her study of the competition between Celera Corporation and the Human Genome Project to decode the human genome, that providing increased public access to research results--as practiced by the HGP--not only resulted in more follow on research but in faster commercialization of the research through new products and services. (The 30% gains in follow-on research and commercialization attributed to the openness of the HGP process persists even today.) More follow-on research and faster commercialization increases economic growth and creates new jobs.

The report includes substantial research demonstrating the positive economic benefits of increased public access under very conservative assumptions.

Research results which are made more publicly available generate more follow on research and more citations in future articles. This is an important benefit to the authors of the manuscripts that are made available. The processes for academic advancement--e.g. tenure and promotion decisions--recognize citations in follow- on research as indicators of the quality of the research. (The report did recommend that these same processes need to be reexamined in order to reward researchers who increase public access to their work by early disclosure of their findings or by sharing new tools and processes. It further recommended that federal agencies recognize such contributions to the progress of science in making grant decisions and in selecting grant panels, etc.).

Increasing access to research results also increases the efficiency of the research enterprise. It helps researchers to locate past research and avoid duplicative or dead end lines of inquiry. It also facilitates the continuing evaluation of research, helps promote accountability for funders, and better administration of the research enterprise allowing a sharper focus on research priorities.

In all of these ways increasing public access promotes the maximum return on the taxpayer's investment in research. Moreover, taxpayers are not asked to pay twice for the same research, first through government grants and then again to obtain access to the results through subscriptions.

IF ALL THE PARTIES AGREE THAT INCREASING PUBLIC ACCESS IS BENEFICIAL, AND IF THE REPORT AND RECENT RESEARCH DEMONSTRATE THE SIGNIFICANT VALUE RESULTING FROM INCREASED PUBLIC ACCESS, ARE THERE DEMONSTRABLE NEGATIVE EFFECTS WHICH OUTWEIGH THIS VALUE?

In doing the research that led to the report we reviewed all the public filings and testimony provided by those who opposed the NIH public access policy and who claimed

that it had or was likely to cause substantial damage. We were unable to find persuasive evidence of such damage even though the policy has now been in effect for nearly 4 years, having gone into effect on April 7, 2008.

We found no persuasive evidence that greater public access as provided by the NIH policy has substantially harmed the subscription supported STM publishers over the last four years or threatens the sustainability of such journals. While there have been subscription cancellations in the last 4 years, these have linked in surveys done on behalf of publishers with the impact of the recession that began in 2008. Academic institutions were hit hard by this recession and library budgets were reduced year after year. (Increased public access was cited by fewer than 5% of respondents as a contributing factor in cancellation decisions in one publisher survey of cancellations cited in the report.)

We found no persuasive evidence that increased public access substantially threatened the ability of STM publishers to fund peer review. In considering the impact of increased public access on peer review it should be noted that authors provide their manuscripts to publishers without being compensated by the publisher (other than through the fact of publication itself). Peer reviews are conducted on a voluntary basis by scholars as part of their obligations as scholars as is much of the editorial effort for STM journals.

We found no persuasive evidence of a significant reduction in traditional publishing outlets; in fact we found no evidence of any high impact proprietary journal ceasing publication for financial reasons. There are more STM journals being published today than there were when the NIH policy went into effect. Over 7500 Open Access journals have started over the last decade. There is no evidence of any shortage of outlets for the distribution of high quality research.

The best evidence of the financial status of the STM journals and of any impact from increased public access remains in the hands of the publishers. We searched public filings and testimony but we also turned to another source with strong incentives to correctly portray the financial conditions of traditional publishers--the financial analysts that cover the STM industry segment and who make recommendations to investors. The STM segment has been a traditional favorite of investors given their relatively high profit margins and their ability to continually raise prices faster than inflation. Financial analysts noted that the ***growth*** of profits for the STM publishers dropped from 6-7% in the first part of the last decade to roughly 4% during the recession. The general consensus over the last several years is for profit growth to increase to 5-6% as the economy recovers. Some financial analysts also reported that the STM publishers downplayed any financial threat due to policies designed to increase public access.

The report concluded that the benefits of increased public access were clear and demonstrable and that there was no persuasive evidence of costs that would outweigh the benefits and provide support for a reversal of the policy. The

cost/benefit analysis provides substantial supports for extending the policies to other federal funders of substantial extramural research.

The benefits of increased access are so great than any delay in availability of research results should be minimized to the extent practicable. A maximum six month delay, now employed by other government and private research funders has not been shown to have any negative impact in the biomedical field. Given the benefits of increased public access, those who seek delay or special conditions based on their disciplines being different from the biomedical arena should bear the burden of proof that the benefits of delay to the development and dissemination of high quality research outweigh the postponed benefits of greater public.

The NIH policy focuses on allowing users of PMC to access and read manuscripts setting out research results of NIH funded research. This policy has had substantial benefit. But the manuscript is not the only measure by which to judge increased access nor is reading an article the only goal of most researchers; in theory researchers should be able to access the manuscript and its subparts--underlying data, protocols, tools utilized for analysis etc. The return on investment in government research would, in theory, be increased to the extent the manuscript and its subparts are machine readable, subject to text and data mining and computable, capable of being displayed, linked and translated into other languages, and subject to analysis with tools chosen by the reader. The challenge will be to find the appropriate degree of openness including access to the manuscript and its subparts and being able to use what is available.

But going beyond the manuscript raises many new questions. Major issues particularly regarding unlimited use and reuse and access to data, remain. Some of these can be addressed by various stakeholders working together; others might be referred to the National Academies. There have been many initiatives in this area which should be supported in an effort to reach consensus which will enable the appropriate degree of increased openness to be determined.

Digital depositories and other mechanisms for dissemination of knowledge provide high returns on investment and should receive greater funding.

Government should work with stakeholders on standards for metadata to enable search and discovery, and standards to ensure interoperability and rules for access among repositories to guarantee access.

Government should also minimize differences among public access rules for federal agencies to promote access and decrease the cost of compliance for both public and private sector entities, particularly those entities that receive funding from multiple federal agencies.

CONCLUSION

Chairman BROUN. Thank you, Mr. Maxwell.

Dr. Taylor, you are recognized for five minutes. And please if you could keep it within five minutes. Your whole written testimony will be part of the record. Thank you.

STATEMENT OF DR. CRISPIN TAYLOR,
EXECUTIVE DIRECTOR,
AMERICAN SOCIETY OF PLANT BIOLOGISTS

Dr. TAYLOR. Thank you very much, Chairman Broun, Ranking Member Tonko, and members of the subcommittee. I am very grateful for the opportunity to come testify today. My name is Crispin Taylor and my title is Executive Director of the American Society of Plant Biologists, a 5,000-member professional society that is based just up the road in Rockville, Maryland. I am also a former Chair of the North American chapter of an organization called the Association of Learned and Professional Society Publishers, which goes by the acronym ALPSP. ALPSP is a global trade association whose membership primarily comprises university presses and professional society publishers. Most of ALPSP's organizational members, including the one I direct, publish one or at most a handful of journals. These journals tend to be among the most highly regarded in their respective disciplines.

It is these two top-ranked journals called "Plant Physiology" and "The Plant Cell," are integrally involved in developing, validating, communicating, disseminating, and ultimately advancing fundamental knowledge about plant biology. This is what the journals and indeed the society are all about. To publish these two journals, the society expends millions of dollars annually on peer review, editorial management, production, printing, shipping, and hosting the online version of the journals on a fully digital, highly reliable platform.

In your invitation, you asked me about the degree to which each of these journals are dependent upon subscriptions. More than half of the Society's $6 million in annual operating revenues derives from subscription payments from some 2,000 institutions worldwide. Despite our strenuous efforts and the fact that we price our two-journal package extremely competitively, we are already finding it increasingly challenging to maintain this customer base.

Your invitation also asked me how scholarly societies might be affected by public access policies promulgated by the government. Well, frankly, that depends on the policies. For example, the policy that mandates a one-size-fits-all embargo period, especially one as short as 6 months, is in my opinion likely to have profoundly negative impacts, especially for smaller professional societies like ASPB. This is not only an assertion. We know from the usage data for "Plant Physiology" that more than half of the article downloads and thus arguably half of the value of the journal to the subscriber take place after the first 6 months. Moreover, librarians have told us in informal conversations that they would be inclined to cancel their subscriptions and wait for release of content at 6 months if that is what it came to.

And that is just ASPB. ALPSP's members publish journals across a wide range of scholarly disciplines, including math, the social sciences, and humanities. For journals in these and other areas of

scholarship, half-lives for article downloads are typically longer than one year, and so I would expect the impacts of even a 12-month mandated embargo to be even more detrimental in those fields.

Although my concerns regarding the adverse impact of mandated embargos are serious, I expect it may take a while for this scenario to play out in the form of failed journals or shuttered Societies. I have two more pressing worries regarding mandated embargos. First, the subscription revenues shrink. The capacity for smaller publishers like ASPB to innovate will be closed off. We will be unable to further improve the utility and impact of our journals, and we will not have time to launch and monetize new products and services like the mobile app on my cell phone through which I can access content for "Plant Phys" and "Plant Cell." Both new services will allow us to diversify our revenue streams and move away from the current business model.

Second, if those mandates come with an obligation to deposit articles in a centrally operated government repository such as the PubMed Central one we have heard about, then for many journals, downloads from those repositories will cut into usage by our own journal websites, further lowering the value of our journals to the subscribers.

So much for policies and regulations that would be harmful. As my testimony indicates, I do think that governments have a legitimate interest in scholarly communication, but I think that interest would be most effectively expressed by encouraging continued innovation rather than stifling it. As we have heard, this approach is already articulated in Section 103 of the COMPETES Act, which envisions collaborative and cooperative engagement of all stakeholders.

Among other things, helpful policies will encourage the continued development and adoption of industry-wide standards, building off the early implementation by almost the entire scholarly communication ecosystem is something called the Digital Object Identifier, or DOI, as a form of Social Security number if you like for journal articles and other pieces of information online. Such standards allow for evermore robust and useful interoperability of otherwise disparate information.

To avoid the distribution of incomplete or imperfect versions of articles, policies and practices should also aim toward providing access to the definitive version of an article, the so-called version of record. It is this version typically available on a publisher website that is actively stewarded and preserved for posterity and to which any corrections or amendments are immediately linked.

In concluding, you asked me whether there is common ground to be found. Despite the differences of opinion and perspectives you have already heard and that we will doubtless explore in greater detail in the coming couple of hours, my answer is a resounding yes. The process begun by the scholarly publishing roundtable was brought together by publishers, librarians, and university leaders is already playing out via the *America COMPETES Act* and the establishment of collaborative public-private projects and partnerships demonstrating that we will make much more progress together as a community of stakeholders than we would in isolation.

Thank you very much.
[The prepared statement of Dr. Taylor follows:]

American Society of Plant Biologists

Cultivating a better future through plant biology research

Official Written Testimony on Federally Funded Research:
Examining Public Access and Scholarly Publication Interests
Submitted to the House Science, Space, and Technology Committee
Subcommittee on Investigations and Oversight
U.S. House of Representatives
Washington, D.C.

Submitted by
Crispin Taylor, Executive Director, American Society of Plant Biologists

On behalf of the Association of Learned and Professional Society Publishers (ALPSP) and the American Society of Plant Biologists (ASPB), I submit this testimony for the official record to the House Science Space and Technology Committee, Subcommittee on Investigations and Oversight. I would like to thank Chairman Broun, Ranking Member Tonko and Subcommittee members very much for their consideration of this testimony regarding access to scholarly information. I have taken the liberty of attaching as appendices detailed comments on very similar topics that were submitted recently by ASPB and ALPSP to the Office of Science and Technology Policy in response to a request for information from that office.

The key points of my testimony are that **the government should adopt sensible, flexible, and cautious approaches to drafting and revising public access policies or regulations.** These approaches should **engage all concerned parties**, including federal agencies, scientists, university administrators, librarians, publishers, and the public, and they should **foster innovation and collaboration**. Policies should focus attention on providing **access to the definitive version of an article**, developing **robust metadata standards**, and on ensuring **increased interoperability among journal articles** and other valuable sources of information online. And they should **recognize and embrace the global nature of scientific research** and scholarly publishing. Although this testimony is not intended to address a particular piece of proposed legislation, it is important to point out that these attributes are largely spelled out in existing legislation – specifically in Section 103 of the America COMPETES Act of 2010 (Public Law No: 111-358[1]) – which itself incorporates many of the recommendations in the report of the Scholarly Publishing Roundtable[2].

[1] http://www.gpo.gov/fdsys/pkg/PLAW-111publ358/html/PLAW-111publ358.htm
[2] http://www.aau.edu/WorkArea/DownloadAsset.aspx?id=10044

Introduction – About ASPB, ALPSP, and Dr. Taylor

The scientific literature has been a part of my education and professional life since I was an undergraduate reading papers in my university library in the United Kingdom. As a graduate student in Michigan and as a postdoc in North Carolina, my appreciation for the literature and the role of publishers in filtering and disseminating it became rather more personal, as I began to see my own work getting published and I started attending scientific meetings at which the latest findings were presented and discussed. Even then, it was abundantly clear to me that **scholarly publishers** – and particularly society and university-based publishers – **are integral components of the communities they serve and support**. In many instances, including ASPB's, they were brought into existence by those communities, and in most cases, they continue to publish the most impactful research in their fields. This is something I have come to know well, because as ASPB's CEO I have become closely involved in a number of activities that have direct bearing on the Society's continued good health – and on the topic of this hearing. First, as chair of the North American chapter of ALPSP (and, as such, a member of that organization's governing council); and second, as an invited member of the Scholarly Publishing Roundtable, which was convened by the forerunner to this committee during the previous Congress. More fundamentally, as CEO, it is also my responsibility to work with my colleagues and Society's elected and appointed governance leaders to ensure that ASPB's long and strong track record as a publisher and a supporter of plant science and plant scientists has as illustrious a future as it does a past.

ASPB is a 501(c)(3) not-for-profit membership corporation created in 1926 and headquartered in Rockville, MD. Today, ASPB is an organization of approximately 5,000 professional plant biology researchers and educators with members in all 50 states and throughout the world. A strong voice for the global plant science community, the Society's mission—achieved through work in the realms of research, education, and public policy—is **to promote the growth and development of plant biology, to encourage and communicate research** in plant biology, and **to promote the interests and growth of plant scientists** in general.

As a large part of its mission to communicate plant science research, the Society self-publishes two of the most widely cited plant science research journals: *The Plant Cell*[3] and *Plant Physiology*[4]. Since 2002, ASPB has also published The Arabidopsis Book[5] (TAB), an innovative, free access, peer-reviewed publication that represents a new model for communicating up-to-date and comprehensive information about a broad range of topics in research on the model plant *Arabidopsis thaliana* and related species. New articles are published as novel research fields emerge, and older content is substantively revised on an ongoing basis so that it can be kept up to date.

Teaching Tools in Plant Biology[6] (TTPB), which is aimed at improving teaching and learning in plant biology, is another innovative product in which ASPB is currently investing. TTPB, an editorial innovation of *The Plant Cell*, combines up-to-date peer-reviewed research with flexible presentation components that can be used alone or integrated into teachers' lesson plans so that they can confidently present exciting plant biology topics in their classrooms. Each Teaching Tool incorporates a short essay introducing each topic, PowerPoint slides, suggested readings, and tips for engaging students in the material.

---

[3] http://www.plantcell.org/
[4] http://www.plantphysiol.org/
[5] http://my.aspb.org/members/group.asp?id=68456
[6] http://www.plantcell.org/site/teachingtools/teaching.xhtml

ASPB is a member of ALPSP, the international trade association representing scholarly and professional publishers across all academic disciplines. ALPSP has **a broad and diverse membership** of over 300 organizations in 37 countries who publish **over half the world's total active journals**, as well as books, databases and other products. **ALPSP's mission is to connect, train and inform** the scholarly and professional publishing community and to **play an active part in shaping the future** of academic and scholarly communication. In total, the scholarly publishing industry employs around 50,000 people and contributes roughly US$3.5 billion to the US balance of trade. In the US, ALPSP represents 60 organizations in 14 states employing an estimated 3,000 individuals.

Of Business Models, Mandates, Access, and Embargos

Scholarly publishing is an international enterprise, with around 1.5 million articles published annually[7]. US researchers dominate this output with a 29% share of the total. The majority of publishers (95%) are small, publishing one or two journals. At the other end of the scale, the 100 largest publishers account for 67% of the total number of journals. Collectively, scholarly publishers adopt a variety of business models to support publication of their journals, reflecting the diverse market in which they operate. No existing digital business model has demonstrated its viability to the satisfaction of all, and ASPB cautions against de facto government endorsement of any single approach.

ASPB and many of ALPSP's other members who publish just one or a handful of journals are, in effect, **small businesses**, and we behave as such. As components of a global information, research and development infrastructure, publishers and professional societies contribute richly and meaningfully in furthering the nation's competitiveness, and we believe that **the work we do supports scholarship, innovation, and economic growth**. We endeavor to foster a culture of innovation and, as do small businesses everywhere, to be prepared to overcome the challenges – and embrace the opportunities – inherent in operating in rapidly evolving business environment.

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

Whether an article is read online or in print, high-quality peer review, page composition (XML), copyediting, and the listing and linking of bibliographic and reference data must be managed, necessitating considerable human capital investment in staff, in addition to scores of editors around the world. **Our editors maintain the quality and reputation of our journals, utilizing the well-established system of peer review**, whereby independent experts review submitted articles. Accepted articles are those that pass muster based on established criteria, including novelty and significance of the research findings. Managing peer review for ASPB's journals is a complex undertaking. It requires sophisticated electronic resources, associated support personnel, and help from thousands of referees. Each year ASPB makes such necessary investments to fulfill its public nonprofit mission, generating both an intellectual and a financial return through the dissemination of scientific research.

But as our mission statements indicate, our motivation is not profit; it's the continued vitality of the disciplines we represent. **So ASPB funnels any surpluses generated from our publishing activities**

---

[7] http://www.stm-assoc.org/industry-statistics/the-stm-report/

into other aspects of our mission – scientific meetings, professional development, public engagement, and broadening participation in our professions. We also frequently chose to make the content we publish freely available, whether through discrete products like TAB (mentioned above), in developing countries, or on an article-by-article basis within the journals through an additional author payment.

Indeed, **ASPB has chosen to make the full content of both of its journals freely available 12 months after publication**, both via the journals' websites and via the NIH's repository, PubMed Central[8]. ASPB has been depositing content at PubMed Central for over a decade, motivated both by an interest in ensuring that high quality plant science research was available to biomedical researchers, and as a component of the Society's long term preservation and archiving strategy. Along with 150 other publishers, ASPB also participates in a suite of services collectively known as Research4Life[9] that makes the literature we publish freely available to researchers in dozens of developing countries. And the full text of both journals, as well as each TTPB, is available upon publication to all members of the Society.

As a result of these and similar publisher-driven initiatives across the industry – and strenuous efforts to develop the global subscription market – **accessing scholarly information is not a problem for scholars**. Indeed, a recent survey from the Publishing Research Consortium found that 97% of researchers in North America have very or fairly easy access to research journals[10].

For ASPB and the majority of other scholarly publishers, however, **embargos are necessary to preserve the initial value of the content we publish** – and to generate the income we use to continue to support our mission, including publishing. And although ASPB has determined that a 12-month embargo strikes an appropriate balance between its objective to disseminate content as broadly as possible and its need to make ends meet, for other journals in other disciplines, different embargo periods are required. **Thus, Federal agencies should not impose embargo periods on non-federally funded businesses.** Individual publisher business models are not arbitrary, but are carefully calibrated to meet the needs of the particular markets in which they operate.

In ASPB's case, the journals generate approximately 80% of the Society's $6 million in annual revenue. A little more than half of the total income derives from 2,000 institutional subscriptions, which we work very hard to sell to universities and corporations around the world, and another 20% from charges levied on hundreds of authors. By contrast, ASPB devotes about half of its operating budget to supporting the journals publishing operation, with the remainder devoted to advancing the broader scholarly missions of the society – missions that, because we are part of the same community, hew to the larger goals of the academic research and education endeavor.

Even so, ASPB is working toward shifting this balance by exploring **the development of new revenue-generating products and services** that meet the needs of our audience and broaden our markets while simultaneously addressing the Society's mission and embracing the opportunities inherent in a shift from print-based to digitally empowered forms of communication. ASPB also continues to experiment with and explore **novel approaches toward expanding access** to its journals' content, including very-low-cost article rental models and, with *Plant Physiology*, via a membership-based free access option.

---

[8] http://www.ncbi.nlm.nih.gov/pmc/
[9] http://www.research4life.org/
[10] http://www.publishingresearch.net/projects.htm Access vs. Importance

However, **it will take some time to establish new revenue streams**, and in the meantime we need the flexibility afforded by the 12 month embargo period we ourselves established for the journals over a decade ago to create the space and time in which to innovate. We know from informal conversations with university librarians and from our own assessment of data on the use of our journals online that **federal imposition of a 6-month embargo would lead to cancellation of subscriptions** and a further tightening of our bottom line. Put simply, shortened embargo periods would undermine the society's capacity to invest for the future.

ALPSP, too, **is not in favor of mandated deposit to centralized open repositories**. In addition to significant concerns about long-term sustainability and piracy, open repositories have deleterious effects on the publishing model; for example, NIH does not currently provide publishers with full, detailed usage statistics from PubMed Central, which means publishers are unable to supply libraries with the complete picture with regard to their institution's use of a wide range of journals. Such usage data are crucial in determining renewals and whilst this situation persists, subscriptions are being cancelled based on incomplete usage data. Furthermore, data from the National Institute of Health reports that more than half of all PubMed Central users are from outside the US. **This repository is therefore reducing the export market for the US publishing industry.**

A Sensible Approach

ASPB strongly supports approaches toward further improving public access – and toward enhancing the utility and value of scholarly information in general – that are inclusive, flexible, forward-looking, and factually based. One-size-fits-all mandates, and, indeed, mandates of any kind, are antithetic to such an approach.

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

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

ASPB and ALPSP are aware that since passage of the America COMPETES reauthorization a number of collaborative projects involving publishers and federal agencies – particularly the National Science

Foundation and the Department of Energy – have emerged. Both organizations strongly support such efforts, which include efforts to more accurately tie published research articles to particular funding opportunities (i.e., grants) and linking between research reports submitted by grantees and subsequent journal publications. These efforts fall on a continuum of vital public-private partnerships that have blossomed in the scholarly publishing arena since journals first started going online in the early 1990s.

Most notable among these is CrossRef [11], which was established by publishers initially to develop mechanisms for linking between the citations sections of journal articles – an important and impactful publisher-led, standards-based innovation that has become a fixture of the scholarly communication enterprise. More recent efforts – by CrossRef and others – focus on author name disambiguation, establishing standards-based unique ids for datasets and other information, and linking database entries to journal articles describing them.

Developing Standards to Foster Interoperability

A defining feature of the Internet is that information is dispersed and widely distributed. It is, nevertheless, readily discoverable. So, **the use of a centralized, government-controlled platform for a large corpus of scholarly content has many significant downsides**, not the least of which is increased and unnecessary costs to the government and **an unnecessary diversion of funds that would be better used to directly support research and discovery**. A centralized approach discourages innovation by driving traffic away from innovators, including publishers, thus minimizing scientific and commercial opportunities.

Indeed, publishers have gone to considerable lengths in developing tools to ensure interoperability between different access systems. For example the Digital Object Identifier (DOI[12]) system, to provide persistent identification of digital objects, CrossRef and its various ongoing projects aimed at connecting users with primary research content and the Open Research and Contributor ID (ORCID[13]) initiative, to solve author name ambiguity in scholarly communications and latterly resolving institutional naming ambiguity.

Publishers are continuing to invest in metadata standards, which improve the ease with which relevant articles can be discovered. With such excellent standards, search tools are all that is required to connect users with the most appropriate content for their needs, and importantly to the VoR. Such metadata standards include those developed by EDItEUR[14], IDEAlliance (PRISM)[15] and NISO[16]. In addition, the Dublin Core Metadata Initiative[17] provides key specifications and best practice regarding the use of metadata for the description of various digital resources (including books and journal articles). It enables interoperability of different applications and vocabularies and optimizes the metadata for searching. DataCite[18], which extends the CrossRef-promoted Digital Object Identifier (DOI) to datasets, is also noteworthy.

---

[11] http://www.crossref.org
[12] http://www.doi.org
[13] http://orcid.org
[14] http://www.editeur.org/
[15] http://www.idealliance.org/specifications/prism/
[16] http://www.niso.org/standards/
[17] http://dublincore.org/
[18] 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.

Within the plant biology domain, ASPB is seeking to collaborate with operators of a prominent knowledge base in plant biology that incorporates a rich array of genomic information from a wide variety of plant species to establish mechanisms for algorithmically connecting journal articles to database entries upon publication. Specifically, the collaborators propose to enable the retrieval of functional gene annotations and molecular annotations from ASPB journal articles using data-mining tools such as Textpresso[19] and BioCreative[20], both of which make use of Natural Language Processing and are organized around robust and highly structured ontologies – standardized dictionaries of terms. The collaborators plan to create a reference library that includes known and predicted gene names, symbols, functions, phenotypes, and pathway annotations in three target plant species. Together with the ontologies, which will play a key role in structuring data annotation, the library will also help establish data capture architectures that the ASPB journals would implement with their authors as manuscripts are being submitted, thereby directly, immediately, and algorithmically connecting published journal articles with the underlying datasets and knowledgebase. Both collaborators envision developing proof-of-concept data-mining methodologies that would be broadly applicable in other fields of research. **Such connections will markedly improve value and utility of scholarly works.**

Despite this community-led innovation in developing metadata standards, an important role for government in this arena is to drive and fund the interoperability standards that would facilitate and enable ever richer connections among journal articles and other types of scholarly information available online and promote the widespread adoption and use of such standards globally.

Version of Record

Studies have demonstrated that researchers prefer to access the publisher-created Version of Record (VoR) from a peer-reviewed journal as the authoritative, definitive version, over versions in subject or institutional repositories[21],[22]. It is therefore ASPB's contention that approaches toward further increasing public access to research articles should have as their primary objective to provide access to the VoR.

Conclusion

As I stated at the outset, it is my contention that the government should adopt sensible, flexible, and cautious approaches to drafting and revising policies or regulations aimed at further improving public access. These approaches should engage all concerned parties, including federal agencies, scientists, university administrators, librarians, publishers, and the public, and they should foster innovation and collaboration. Policies should focus attention on providing access to the VoR, developing robust metadata standards, and on ensuring increased interoperability among journal articles and other valuable sources of information online. And they should recognize and embrace the global nature of scientific research and scholarly publishing.

This concludes my testimony. Many thanks again for your time and attention.

Crispin Taylor, PhD
Executive Director
American Society of Plant Biologists

---

[19] http://www.textpresso.org/
[20] http://biocreative.sourceforge.net/
[21] http://www.peerproject.eu/reports/ D4.2 PEER Behavioural Research – Final Report
[22] http://www.publishingresearch.net/projects.htm Research Publication Characteristics and Their Relative Values

Chairman BROUN. Thank you, Dr. Taylor. You were just dead-on. Those buzzes that you just heard indicates that we have a vote on the Floor. If the two of you can be very quick and get through exactly five minutes each, we can get through both of you all and we will recess and come back. So I would like to try to do that if you all could please hold your comments to five minutes or if you could cut it 30 seconds or so short. I don't want to shortchange anybody, Mr. Maxwell, or anybody but Mr. Shieber, you are recognized for five minutes.

STATEMENT OF MR. STUART M. SHIEBER, DIRECTOR, OFFICE FOR SCHOLARLY COMMUNICATIONS, HARVARD UNIVERSITY

Mr. SHIEBER. Okay. Thank you, Chairman Broun, Ranking Member Tonko, members of the committee. My name is Stuart Shieber. I am a Computer Science Professor at Harvard University. As a faculty member, I led the development of Harvard's Open Access Policies, and since 2008, I have also served as the Faculty Director of Harvard's Office for Scholarly Communication. I want to thank you for the opportunity to speak with you today.

Harvard's longstanding research policy calls for the idea that we as—calls for the ideas that we as faculty and researchers produce to receive "the widest possible dissemination." At one time, this was achieved by distributing scholarly articles in the form of printed issues of journals sent to the research libraries of the world for reading by their patrons and paid for by subscription fees. But internet-era technologies hold the promise of transforming this system, distributing and using knowledge in ways not previously imaginable. Ideally, this would lead to a universality of access to research results known as open access, truly achieving the widest possible dissemination of our research.

The benefits of open access are many. It eliminates barriers to reading scholarly literature and broadens access beyond just those who, like myself, are privileged to be within the orbit of a major research library. It expands access to those at the full range of schools, businesses large and small, and the general public. Open access makes practical the novel reuse of the literature through computer analysis of the entire corpus of research results. Economists have shown that broader public access to federally funded research would have positive impacts on the U.S. economy totaling billions of dollars, covering its cost many times over. But perhaps most importantly, open access to research is an intrinsic public good. As Thomas Jefferson wrote, "the most important bill in our whole code is that for the diffusion of knowledge among the people."

Unfortunately, this promise is not currently being realized because of systemic problems in scholarly publishing. Libraries have observed with alarm a long-term dramatic hyperinflation in subscription costs of journals which have increased at a steady rate of about seven percent per year. Other symptoms of dysfunction include huge price disparities among journals of similar quality and extraordinary profit margins. Even at Harvard, which holds the largest academic library in the world, we have had to curtail serial spending through a painful series of journal cancellations.

In 2008, the faculty of arts and sciences, my own faculty, enacted by unanimous vote an open-access policy. Faculty decided to grant a license to the university to openly distribute our scholarly articles and commit to providing copies of our manuscript articles for distribution through our online repository. This policy allows faculty to retain sufficient rights to provide open supplemental access to our scholarly articles. Since then, similar policies have been voted by faculty bodies across Harvard, as well as institutions as diverse as MIT, University of Kansas, Stanford, Princeton, Bucknell, Columbia, Oberlin, Duke, and many others. Because of the Harvard policies, we now provide online access to over 7,000 articles, representing 4,000 Harvard-affiliated authors. These articles have been downloaded almost three-quarters of a million times to people on every continent in the world, including, surprisingly, Antarctica.

This approach to the access problem is also seen in the extraordinarily successful public access policy of the National Institutes of Health. Today, NIH provides free online access to 2.4 million articles downloaded a million times per day by half a million individual users. These users come from universities, research labs, companies, and the general public showing the broad scope of the latent demand for these materials. The NIH model could be replicated at other funded agencies as envisioned in the recently reintroduced bipartisan Federal Research Public Access Act.

The standard objection to this kind of policy is that supplemental open access could harm the publishing industry, but as Mr. Maxwell's report for the CED concluded, after four years of the NIH policy and institutional policies like Harvard's, "there is no persuasive evidence that increased access threatens the sustainability of traditional subscription-supported journals or their ability to fund rigorous peer review."

In addition, the author gains value from publication of an article in a journal. Vetting, copyediting, typesetting, and most importantly, imprimatur of the journal, value that authors and their institutions and funders should be, would be, and are willing to pay for. Thus, in a hypothetical world of harm to subscriptions, journals can switch to a different revenue model charging a one-time publication fee to cover the costs of publishing the article. I state this as though the publication fee revenue model is itself hypothetical, but it is not. Peer-reviewed open-access journals already exist in the thousands, many using the publication fee revenue model, which by now is a proven mechanism used by commercial and nonprofit publishers alike, including even the most established journal publishers.

I opened my statement by quoting a mission of academics such as myself to provide the widest possible dissemination—open access—to the ideas and knowledge resulting from our research. Government, too, has an interest in the widest possible dissemination of government-funded research—to maximize the return on the taxpayers' enormous investment in that research, to bring economic benefits that far exceed the costs, to provide transparency, to inform the public. Providing open access to the publicly funded research literature, fulfilling Jefferson's call for diffusion of knowledge, will benefit researchers, taxpayers, and every person who

gains from new technologies, new medicines, new jobs, and new solutions to longstanding problems of every kind.

Thank you.

[The prepared statement of Mr. Shieber follows:]

Statement of Stuart M. Shieber before the
Committee on Science, Space and Technology
Subcommittee on Investigations and Oversight
U.S. House of Representatives

March 29, 2012

Chairman Broun and Members of the Subcommittee:

My name is Stuart Shieber. I am the James O. Welch, Jr. and Virginia B. Welch Professor of Computer Science at Harvard University. My primary field of research is computational linguistics, the study of human language from a computer science perspective, often with application to the engineering of useful computer systems that manipulate language. As a faculty member, I led the development and enactment of Harvard's open-access policies. Since October of 2008, I have served in the additional role as the faculty director of Harvard's Office for Scholarly Communication. Thank you for the opportunity to speak with you today about some of the actions that we have taken at Harvard to provide the broadest possible access to the results of our research.

THE POTENTIAL FOR OPEN ACCESS

The mission of the university is to create, preserve, and disseminate knowledge to the benefit of all. In Harvard's Faculty of Arts and Sciences (FAS), where I hold my faculty post, we codify this in the FAS Grey Book, which states that research policy "should encourage the notion that ideas or creative works produced at the University should be used for the greatest possible public benefit. This would normally mean the widest possible dissemination and use of such ideas or materials."

At one time, the widest possible dissemination was achieved by distributing the scholarly articles describing the fruits of research in the form of printed issues of peer-reviewed journals, sent to the research libraries of the world for reading by their patrons, and paid for by subscription fees. These fees covered the various services provided to the authors of the articles — management of the peer review process, copy-editing, typesetting, and other production processes — as well as the printing, binding, and shipping of the physical objects.

Thanks to the forward thinking of federal science funding agencies, including NSF, DARPA, NASA, and DOE, we now have available computing and networking technologies that hold the promise of transforming the mechanisms for disseminating and using knowledge in ways not imaginable even a few decades ago. The internet allows nearly instantaneous distribution of content for essentially zero marginal cost to a large and rapidly increasing proportion of humanity. Ideally, this would ramify in a universality of access to research results, thereby truly achieving the widest possible dissemination.

2

The benefits of such so-called *open access* are manifold. The signatories of the 2002 Budapest Open Access Initiative state that

> The public good [open access] make[s] possible is the world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds. Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge.

From a more pragmatic point of view, a large body of research has shown that public research has a large positive impact on economic growth, and that access to the scholarly literature is central to that impact. Martin and Tang's recent review of the literature concludes that "there have been numerous attempts to measure the economic impact of publicly funded research and development (R&D), all of which show a large positive contribution to economic growth."[1] It is therefore not surprising that Houghton's modeling of the effect of broader public access to federally funded research shows that the benefits to the US economy come to the billions of dollars and are eight times the costs.[2]

Opening access to the literature makes it available not only to human readers, but to computer processing as well. There are some million and a half scholarly articles published each year.[3] No human can read them all or even the tiny fraction in a particular subfield, but computers can, and computer analysis of the text, known as *text mining*, has the potential not only to extract high-quality structured data from article databases but even to generate new research hypotheses. My own field of research, computational linguistics, includes text mining. I have collaborated with colleagues in the East Asian Languages and Civilization department on text mining of tens of thousands of classical Chinese biographies and with colleagues in the History department on computational analysis of pre-modern Latin texts. Performing similar analyses on the current research literature, however, is encumbered by proscriptions of copyright and contract because the dominant publishing mechanisms are not open.

---

[1]Ben R. Martin and Puay Tang, The benefits from publicly funded research, SEWPS Paper No. 161, SPRU—Science and Technology Policy Research, University of Sussex, Brighton (2007). `http://www.sussex.ac.uk/spru/documents/sewp161`

[2]John Houghton, *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs* (July 2010). `http://www.arl.org/sparc/bm~doc/vufrpaa`

[3]Scholarly Publishing Roundtable, *Report and Recommendations from the Scholarly Publishing Roundtable* (January 2010). `http://www.aau.edu/WorkArea/DownloadAsset.aspx?id=10044`

In Harvard's response to the Office of Science and Technology Policy's request for information on public access,[4] Provost Alan Garber highlighted the economic potential for the kinds of reuse enabled by open access.

> Public access not only facilitates innovation in research-driven industries such as medicine and manufacturing. It stimulates the growth of a new industry adding value to the newly accessible research itself. This new industry includes search, current awareness, impact measurement, data integration, citation linking, text and data mining, translation, indexing, organizing, recommending, and summarizing. These new services not only create new jobs and pay taxes, but they make the underlying research itself more useful. Research funding agencies needn't take on the job of provide all these services themselves. As long as they ensure that the funded research is digital, online, free of charge, and free for reuse, they can rely on an after-market of motivated developers and entrepreneurs to bring it to users in the forms in which it will be most useful. Indeed, scholarly publishers are themselves in a good position to provide many of these value-added services, which could provide an additional revenue source for the industry.

Finally, free and open access to the scholarly literature is an intrinsic good. It is in the interest of the researchers generating the research and those who might build upon it, the public who take interest in the research, the press who help interpret the results, and the government who funds these efforts. All things being equal, open access to the research literature ought to be the standard.

SYSTEMIC PROBLEMS IN THE JOURNAL PUBLISHING SYSTEM

Unfortunately, over the last several years, it has become increasingly clear to many that this goal of the "widest possible dissemination" was in jeopardy because of systemic problems in the current mechanisms of scholarly communication, which are not able to take full advantage of the new technologies to maximize the access to research and therefore its potential for social good.

By way of background, I should review the standard process for disseminating research results. Scholars and researchers — often with government funding — perform research and write up their results in the form of articles, which are submitted to journals that are under the editorial control of the editor-in-chief and editorial boards made up of other scholars. These editors find appropriate reviewers, also scholars, to read and provide detailed reviews of the articles, which authors use to improve the quality of the articles. Reviewers also provide advice to the editors on whether the articles are appropriate for publication in the journal, the final decisions being

[4] Alan Garber, Harvard response to the White House RFI on public access to research (January 2012). `http://osc.hul.harvard.edu/stp-rfi-response-january-2012`

made by the editors. Participants in these aspects of the publishing process are overwhelmingly volunteers, scholars who provide their time freely as a necessary part of their engagement in the research enterprise. The management of this process, handling the logistics, is typically performed by the journal's publisher, who receives the copyright in the article from the author for its services. The publisher also handles any further production process such as copy-editing and typesetting of accepted articles and their distribution to subscribers through print issue or more commonly these days through online access. This access is provided to researchers by their institutional libraries, which pay for annual subscriptions to the journals.

Libraries have observed with alarm a long-term dramatic rise in subscription costs of journals. The Association of Research Libraries, whose members represent the leading research libraries of the United States and Canada, have tracked serials expenditures for over three decades. From 1986 through 2010 (the most recent year with available data), expenditures in ARL libraries have increased by a factor of almost 5. Even discounting for inflation, the increase is almost 2.5 times. These increases correspond to an annualized rate of almost 7% per year, during a period in which inflation has averaged less than 3%.[5]

Another diagnostic of the market dysfunction in the journal publishing system is the huge disparity in subscription costs between different journals. Bergstrom and Bergstrom showed that even within a single field of research, commercial journals are *on average* five times more expensive per page than non-profit journals.[6] When compared by cost per citation, which controls better for journal quality, the disparity becomes even greater, a factor of 10 times. Odylzko notes that "The great disparity in costs among journals is a sign of an industry that has not had to worry about efficiency."[7] Finally, the extraordinary profit margins, increasing even over the last few years while research libraries' budgets were under tremendous pressure, provide yet another signal of the absence of a functioning competitive market.

The Harvard library system is the largest academic library in the world, and the fifth largest library of any sort. In attempting to provide access to research results to our faculty and students, the university subscribes to tens of thousands of serials at a cost of about 9 million dollars per year. Nonetheless, we too have been buffeted by the tremendous growth in journal costs over the last decades, with Harvard's serials expenditures growing by a factor of 3 between 1986 and

[5]Association of Research Libraries, Monograph and Serial Costs in ARL Libraries, 1986-2010 (2010). `http://www.arl.org/bm~doc/t2_monser10.xls`

[6]Carl T. Bergstrom and Theodore C. Bergstrom, The costs and benefits of library site licenses to academic journals, *Proceedings of the National Academy of Sciences*, volume 101, number 3 (20 January 2004). `http://dx.doi.org/10.1073/pnas.0305628101`

[7]Andrew Odlyzko, The Economics of Electronic Journals, *First Monday*, volume 2, number 8 (4 August 1997). `http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/542/463`

2004.[8] Such geometric increases in expenditures could not be sustained indefinitely. Over the years since 2004 our journal expenditure increases have been curtailed through an aggressive effort at deduplication, elimination of print subscriptions, and a painful series of journal cancellations. As a researcher, I know that Harvard does not subscribe to all of the journals that I would like access to for my own research, and if Harvard, with its scale, cannot provide optimal subscription access, other universities without our resources are in an even more restricted position.

Correspondingly, the articles that we ourselves generate as authors are not able to be accessed as broadly as we would like. We write articles not for direct financial gain — we are not paid for the articles and receive no royalties — but rather so that others can read them and make use of the discoveries they describe. To the extent that access is limited, those goals are thwarted.

The economic causes of these observed phenomena are quite understandable. Journal access is a monopolistic good. Libraries can buy access to a journal's articles only from the publisher of that journal, by virtue of the monopoly character of copyright. In addition, the high prices of journals are hidden from the "consumers" of the journals, the researchers reading the articles, because an intermediary, the library, pays the subscriptions on their behalf. The market therefore embeds a moral hazard. Under such conditions, market failure is not surprising; one would expect inelasticity of demand, hyperinflation, and inefficiency in the market, and that is what we observe. Prices inflate, leading to some libraries canceling journals, leading to further price increases to recoup revenue — a spiral that ends in higher and higher prices paid by fewer and fewer libraries. The market is structured to provide institutions a Hobson's choice between unsustainable expenditures or reduced access.

The unfortunate side effect of this market dysfunction has been that as fewer libraries can afford the journals, access to the research results they contain is diminished. In 2005, then Provost of Harvard Steven Hyman appointed an ad hoc committee, which I chaired, to examine these issues and make recommendations as to what measures Harvard might pursue to mitigate this problem of access to our writings. Since then, we have been pursuing a variety of approaches to maximize access to the writings of Harvard researchers.

ADDRESSING INSUFFICIENT ACCESS THROUGH AN OPEN-ACCESS POLICY

One of these approaches involves the self-imposition by faculty of an open-access policy according to which faculty grant a license to the university to distribute our scholarly articles and commit to providing copies of our manuscript articles for such distribution. By virtue of this kind of policy, the problem of access limitation is mitigated by providing a supplemental venue for access to the

[8] Association of Research Libraries, Monograph and Serial Costs in ARL Libraries, 1986-2010 (2010). `http://www.arl.org/bm~doc/t2_monser10.xls`

articles. Four years ago, in February of 2008, the members of the Faculty of Arts and Sciences at Harvard became the first school to enact such a policy,[9] by unanimous vote as it turned out.

In order to guarantee the freedom of faculty authors to choose the rights situation for their articles, the license is waivable at the sole discretion of the author, so faculty retain control over whether the university is granted this license. But the policy has the effect that by default, the university holds a license to our articles, which can therefore be distributed from a repository that we have set up for that purpose. Since the FAS vote, six other schools at Harvard — Harvard Law School, Harvard Kennedy School of Government, Harvard Graduate School of Education, Harvard Business School, Harvard Divinity School, and Harvard Graduate School of Design — have passed this same kind of policy, and similar policies have been voted by faculty bodies at many other universities as well, including Massachusetts Institute of Technology, Stanford, Princeton, Columbia, and Duke. Notably, the policies have seen broad faculty support, with faculty imposing these policies on themselves typically by unanimous or near unanimous votes.

Because of these policies in the seven Harvard schools, Harvard's article repository, called DASH (for Digital Access to Scholarship at Harvard),[10] now provides access to over 7,000 articles representing 4,000 Harvard-affiliated authors. Articles in DASH have been downloaded almost three-quarters of a million times.[11] The number of waivers of the license has been very small; we estimate the waiver rate at about 5%. Because of the policy, as faculty authors we are retaining rights to openly distribute the vast majority of the articles that we write.

The process of consultation in preparation for the faculty vote was a long one. I started speaking with faculty committees, departments, and individuals about two years before the actual vote. During that time and since, I have not met a single faculty member or researcher who objected to the principle underlying the open-access policies at Harvard, to obtain the widest possible dissemination for our scholarly results, and have been struck by the broad support for the kind of open dissemination of articles that the policy and the repository allow.

This approach to the access limitation problem, the provision of supplemental access venues, is also seen in the extraordinarily successful public access policy of the National Institutes of Health (NIH), which Congress mandated effective April, 2008. By virtue of that policy, researchers funded by NIH provide copies of their articles for distribution from NIH's PubMed Central (PMC) repository. Today, PMC provides free online access to 2.4 million articles downloaded a million times per day by half a million users.[12] NIH's own analysis has shown that a quarter of the users

[9]Text of the FAS policy and the other Harvard open-access policies is available at `http://osc.hul.harvard.edu/policies`.

[10]`http://dash.harvard.edu/`

[11]`http://dash.harvard.edu/mydash`

[12]National Institutes of Health, *NIH Public Access Policy Implications* (2012). `http://publicaccess.nih.gov/public_access_policy_implications_2012.pdf`

are researchers. The hundreds of thousands of articles they are accessing per day demonstrates the large latent demand for articles not being satisfied by the journals' subscription base. Companies account for another 17%, showing that the policy benefits small businesses and corporations, who need access to scientific advances to spur innovation. Finally, the general public accounts for 40% of the users, some quarter of a million people per day, demonstrating that these articles are of tremendous interest to the taxpayers who fund the research in the first place and who deserve access to the results that they have underwritten.

THE STANDARD OBJECTION TO OPEN-ACCESS POLICIES

The standard objection to these open-access policies is that supplemental access to scholarly articles, such as that provided by institutional repositories like Harvard's DASH or subject-based repositories like NIH's PubMed Central, could supplant subscription access to such an extent that subscriptions would come under substantial price pressure. Sufficient price pressure, in this scenario, could harm the publishing industry, the viability of journals, and the peer review and journal production processes.

There is no question that the services provided by journals are valuable to the research enterprise, so such concerns must be taken seriously. By now, however, these arguments have been aired and addressed in great detail. I recommend the report "The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?" by my co-panelist Elliott Maxwell,[13] which provides detailed support for the report's conclusion that "There is no persuasive evidence that increased access threatens the sustainability of traditional subscription-supported journals, or their ability to fund rigorous peer review." The reasons are manifold, including the fact that supplemental access covers only a fraction of the articles in any given journal, is often delayed relative to publication, and typically provides a manuscript version of the article rather than the version of record. Consistent with this reasoning, the empirical evidence shows no such discernible effect. After four years of the NIH policy, for instance, subscription prices have continued to increase, as have publisher margins. The NIH states that "while the U.S. economy has suffered a downturn during the time period 2007 to 2011, scientific publishing has grown: The number of journals dedicated to publishing biological sciences/agriculture articles and medicine/health articles increased 15% and 19%, respectively. The average subscription prices of biology journals and health sciences journals increased 26% and 23%, respectively. Publishers forecast increases to the rate of growth of the medical journal market, from 4.5% in 2011 to 6.3% in 2014."[14]

[13]Committee for Economic Development. *The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?* (2012). `http://www.ced.org/component/blog/entry/1/765`

[14]National Institutes of Health, *NIH Public Access Policy Implications* (2012). `http://publicaccess.nih.gov/public_access_policy_implications_2012.pdf`

Nonetheless, it does not violate the laws of economics that increased supplemental access (even if delayed) to a sufficiently high proportion of articles (even if to a deprecated version) could put price pressure on subscription journals, perhaps even so much so that journals would not be able to recoup their costs. In this hypothetical case, would that be the end of journals? No, because even if publishers (again, merely by hypothesis and counterfactually) add no value for the readers (beyond what the readers are already getting in the [again hypothetical] universal open access), the author and the author's institution gain much value: vetting, copyediting, typesetting, and most importantly, imprimatur of the journal. This is value that authors and their institutions should be, would be, and are willing to pay for. The upshot is that journals will merely switch to a different business model, in which the journal charges a one-time *publication fee* to cover the costs of publishing the article.

I state this as though this publication-fee revenue model is itself hypothetical, but it is not. Open-access journals already exist in the thousands. They operate in exactly the same way as traditional subscription journals — providing management of peer review, production services, and distribution — with the sole exception that they do not charge for online access, so that access is free and open to anyone. The publication-fee revenue model for open-access journals is a proven mechanism. The prestigious non-profit open-access publisher Public Library of Science is generating surplus revenue and is on track to publish some 3% of the world biomedical literature through its journal *PLoS ONE* alone. The BioMed Central division of the commercial publisher Springer is generating profits for its parent company using the same revenue model. Indeed, the growth of open-access journals over the past few years has been meteoric. There are now over 7,000 open-access journals,[15] many using the publication-fee model, and many of the largest, most established commercial journal publishers — Elsevier, Springer, Wiley-Blackwell, SAGE — now operate open-access journals using the publication-fee revenue model. Were supplemental access to cause sufficient price pressure to put the subscription model in danger, the result would merely be further uptake of this already burgeoning alternative revenue model.

In this scenario, the cost of journal publishing would be borne not by the libraries on behalf of their readers, but by funding agencies and research institutions on behalf of their authors. Already, funding agencies such as Wellcome Trust and Howard Hughes Medical Institute underwrite open access author charges, and in fact mandate open access. Federal granting agencies such as NSF and NIH allow grant funds to be used for open-access publication fees as well (though grantees must

[15] According to the Directory of Open Access Journals, `http://www.doaj.org/`.

prebudget for these unpredictable charges). Not all fields have the sort of grant funding opportunities that could underwrite these fees. For those fields, the researcher's employing institution, as de facto funder of the research, should underwrite charges for publication in open-access journals. Here again, Harvard has taken an early stand as one of the initial signatories — along with Cornell, Dartmouth, MIT, and University of California, Berkeley — of the Compact for Open-Access Publishing Equity,[16] which commits these universities and the dozen or so additional signatories to establishing mechanisms for underwriting reasonable open-access publication fees. The Compact acknowledges the fact that the services that journal publishers provide are important, cost money, and deserve to be funded, and commits the universities to doing so, albeit with a revenue model that avoids the market dysfunction of the subscription journal system.

ADVANTAGES OF THE OPEN-ACCESS PUBLISHING SYSTEM

The primary advantage of the open-access journal publishing system is the open access that it provides. Since revenue does not depend on limiting access to those willing to pay, journals have no incentive to limit access, and in fact have incentive to provide as broad access as possible to increase the value of their brand. In fact, open-access journals can provide access not only in the traditional sense, allowing anyone to access the articles for the purpose of reading them, but can provide the articles unencumbered by any use restrictions, thereby allowing the articles to be used, re-used, analyzed, and data-mined in ways we are not even able to predict.

A perhaps less obvious advantage of the publication-fee revenue model for open-access journals is that the factors leading to the subscription market failure do not inhere in the publication-fee model. Bergstrom and Bergstrom[17] explain why:

> Journal articles differ [from conventional goods such as cars] in that they are not substitutes for each other in the same way as cars are. Rather, they are complements. Scientists are not satisfied with seeing only the top articles in their field. They want access to articles of the second and third rank as well. Thus for a library, a second copy of a top academic journal is not a good substitute for a journal of the second rank. Because of this lack of substitutability, commercial publishers of established second-rank journals have substantial monopoly power and are able to sell their product at prices that are much higher than their average costs and several times higher than the price of higher quality, non-profit journals.

[16] `http://www.oacompact.org/`. See also Stuart M. Shieber, Equity for open-access journal publishing, *PLoS Biology*, volume 7, number 8 (2012). `http://dx.doi.org/10.1371/journal.pbio.1000165`

[17] Theodore C. Bergstrom and Carl T. Bergstrom, Can 'author pays' journals compete with 'reader pays'?, *Nature Web Focus* (2004). `http://www.nature.com/nature/focus/accessdebate/22.html`

> By contrast, the market for authors' inputs appears to be much more competitive. If journals supported themselves by author fees, it is not likely that one Open Access journal could charge author fees several times higher than those charged by another of similar quality. An author, deciding where to publish, is likely to consider different journals of similar quality as close substitutes. Unlike a reader, who would much prefer access to two journals rather than to two copies of one, an author with two papers has no strong reason to prefer publishing once in each journal rather than twice in the cheaper one.
>
> If the entire market were to switch from Reader Pays to Author Pays, competing journals would be closer substitutes in the view of authors than they are in the view of subscribers. As publishers shift from selling complements to selling substitutes, the greater competition would be likely to force commercial publishers to reduce their profit margins dramatically.

Again, the empirical evidence supports this view. Even the most expensive open-access publication fees, such as those of the prestigious Public Library of Science journals, are less than $3,000 per article, with a more typical value in the $1,000–1,500 range. By contrast, the average revenue per article for subscription journal articles is about $5,000. Thus, the open-access model better leverages free market principles: Despite providing unencumbered access to the literature, it costs no more overall per article, and may end up costing much less, than the current system. The savings to universities and funding agencies could be substantial.

CONCLUSION

I began my comments by quoting the mission of academics such as myself to provide the widest possible dissemination — open access — to the ideas and knowledge resulting from our research. Government, too, has an underlying goal of promoting the dissemination of knowledge, expressed in Thomas Jefferson's view that "by far the most important bill in our whole code is that for the diffusion of knowledge among the people."[18] The federal agencies and science policies that this committee oversees have led to knowledge breakthroughs of the most fundamental sort — in our understanding of the physical universe, in our ability to comprehend fundamental biological processes, and, in my own field, in the revolutionary abilities to transform and transmit information.

Open access policies build on these information technology breakthroughs to maximize the return on the taxpayers' enormous investment in that research, and magnify the usefulness of that research. They bring economic benefits that far exceed the costs. The NIH has shown one

[18]Thomas Jefferson, Letter to George Wythe (13 August, 1786). `http://hdl.loc.gov/loc.mss/mtj.mtjbib002184`

successful model, which could be replicated at other funding agencies, as envisioned in the recently re-introduced bipartisan Federal Research Public Access Act (FRPAA).

Providing open access to the publicly-funded research literature — amplifying the "diffusion of knowledge" — will benefit researchers, taxpayers, and every person who gains from new medicines, new technologies, new jobs, and new solutions to longstanding problems of every kind.

Chairman BROUN. Thank you, Mr. Shieber.

Mr. Plutchak, you are recognized for five minutes.

STATEMENT OF MR. SCOTT PLUTCHAK, DIRECTOR, LISTER HILL LIBRARY, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Mr. PLUTCHAK. Mr. Chairman, members of the subcommittee, thank you very much for the opportunity to speak to you today. My name is Scott Plutchak, and I am the Director of the Lister Hill Library in the Health Sciences at the University of Alabama at Birmingham.

For over a decade, much of my professional activity has taken place at the intersection of librarianship and publishing. Despite how dependent these two communities are on each other, it is striking how little real conversation and collaboration has taken place. One of the unfortunately results of this is that the debates over access to the peer-reviewed literature have been unnecessarily contentious and have diverted energy and attention from what could have been and should have been a careful examination of facts and opportunities.

Open-access advocates are lobbying hard for passage of FRPAA, the Federal Research Public Access Act, but I am afraid they are being shortsighted. The true value of the peer-reviewed literature comes from context when it is connected to the work that comes before it and that provides a foundation for what can be built upon it. Because FRPAA focuses so much on access to individual copies of articles, it falls far short of what could be and what needs to be achieved. I agree that the peer-reviewed reports of federally funded research should be made freely available. In the digital world, we certainly ought to be able to make that happen, but it must be done within a context that maximizes the value of those articles. Mere access isn't enough.

The report from the Scholarly Publishing Roundtable lays out the issues that need to be balanced—first, a focus on providing access to the version of record. Like the NIH policy on which it is based, FRPAA settles for access to the author's final manuscript version. In most cases, to meet a current need, that may be sufficient, but a robust system of scientific communication that is expected to persist over time requires access to the version that some entity is going to keep track of, ensuring that corrections are appropriately made, that retractions are handled when necessary, and that context is preserved.

Second—interoperability standards. Data mining of research reports can combine results to provide accurate summaries across a broad range of experiments. When these meta-analyses are performed by humans, they can be extremely beneficial but they are very expensive and they take a lot of time. In order to facilitate this data mining, we need to emphasize a standardized structure for research reports.

Third—a focus on digital preservation. Although there are several promising approaches underway, these are all still in the experimental stages. A recent study reported that only 15 to 20 percent of the e-journal content in two major research libraries was currently being preserved. Any set of policies designed to provide

access to the results of funded research must take into account methods for maintaining access for decades and even centuries.

Finally, a recognition that different disciplines have different practices and needs. One of the most contentious areas in the public access debates has to do with the length of the embargo. It became clear to us on the Roundtable as we investigated the differences among disciplines that 6 months may be longer than necessary in some disciplines and too short to be practical in others. While policy development across funding agencies must be carefully coordinated, some flexibility must be allowed and even encouraged so that agencies can develop policies that are acutely attuned to the needs of their disciplinary communities.

Librarians, researchers, educators, and publishers are all committed to achieving the widest possible distribution of the results of scholarly research. The challenge is to do that in ways that balance these elements. My experience with the Roundtable, my other work with the librarian and publishing communities over the years leaves me convinced that we will be more effective in developing those solutions when researchers, educators, publishers, librarians, and the public work together than if we continue on a path that sets us at odds with each other.

Across the publishing industry in both commercial and not-for-profit sectors, experiments in open-access publishing are proliferating. It should be clear to any objective observer that most publishers are in no way opposed to open access. They are quite sensibly seeking to develop business models that keep their organizations healthy while maximizing access to what they publish.

As authorized by the *America COMPETES Act*, the White House Office of Science and Technology Policy is currently reviewing the results from its latest request for information. Their approach to these questions suggest that they are attempting to achieve the kind of balance recommended in the report. Anything that the Congress can do to encourage policy development along those lines will be welcome. Anything that impedes or interferes with that work such as the passage of FRPAA-type legislation will paradoxically make the goal of a truly robust open-access infrastructure for scientific communication even more difficult to achieve.

Developing federal policies that will maximize the public's investment in research and provide incentives for the development of a robust scholarly communication system is complicated and achieving the appropriate balance of interests may not be as emotionally satisfying as advocating the simplicity of something like FRPAA but the American public deserves to have us do this right.

Thank you very much.

[The prepared statement of Mr. Plutchak follows:]

Testimony of

T. Scott Plutchak
Director, Lister Hill Library of the Health Sciences
University of Alabama at Birmingham

before the

Subcommittee on Investigations and Oversight

of the

House Committee on Science, Space and Technology

concerning

"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

March 29, 2012

Chairman Broun, Ranking Member Tonko, Members of the Subcommittee, thank you very much for the opportunity to speak to you today on the issues surrounding public access to the peer-reviewed reports of federally funded research. My name is Scott Plutchak, and I'm currently the director of the Lister Hill Library of the Health Sciences at the University of Alabama at Birmingham. The views I'm expressing today, however, are my own opinions and do not in any way represent any official position of my university.

Historians of the beginning of the print age refer to the fifty year period from about 1450 – when Gutenberg's press was invented – to about 1500 as the "incunabula period," from the Latin word meaning "from the cradle." This was a period of rich experimentation when entrepreneurs, scientists and scholars tested the opportunities offered by this powerful new technology. What resulted was the beginnings of the great age of print that we all grew up in.

We are now in the incunabula period of the digital age, and the opportunities before us are tremendous. Digital technologies for communicating scientific knowledge hold the promise for accelerating discovery in ways unimaginable to previous generations. We see them affecting every aspect of our social and economic life. Certainly the impact that they've had on my world of libraries and librarians has been earth-shattering, and equally so for the world of research and education that academic librarians have dedicated themselves to. We have it within our grasp to establish a new information infrastructure for the communication and advancement of scientific knowledge that builds on the best of what we've done in the past, while taking full advantage of the power of these new digital information technologies.

Unfortunately, the debates over access to the peer-reviewed journal literature that have taken place over the last decade or so have been unnecessarily contentious and have diverted energy and attention from what could have been, and should have been, a careful examination of facts and opportunities. The recent flurry of activity surrounding the introduction of the Research Works Act (RWA) and the reintroduction of the Federal Research Public Access Act (FRPAA) have resulted in a great deal of the kind of sloganeering, wishful thinking, and shading of fact in order to score rhetorical points that has characterized most of the public debate around Open Access during the past decade.

The most vocal of the Open Access advocates are lobbying hard for passage of the Federal Research Public Access Act, but I'm afraid they are being terribly short-sighted. Passage of FRPAA would be the digital equivalent of the first half of the 17th century, and we have the opportunity to do so much more.

The great achievement in scientific communication in the 17th century was embodied in the first two scientific journals – the *Journal des Sçavans* , first published in Paris in January, 1665, and *The Philosophical Transactions of the Royal Society*, appearing in London two months later. For the first time, scientific reports would be bundled together and disseminated systematically across Europe. It was the result of the best minds of the time using the latest technology to advance scientific progress.

We have the potential now for a similar significant leap forward, but the FRPAA approach risks actually take us backward – to a world in which individual reports of scientific results are isolated from context, held in scattered repositories, with very little in the way of connective tissue.

The true value of the peer-reviewed literature comes from context, when it is connected to the work that comes before it, that is laterally related to it, and that provides a foundation for what can be built upon it.

I completely agree with the notion that the peer-reviewed reports of federally funded research should be made freely available. In a digital world where distribution is cheap and easy, we certainly ought to be able to figure out a way to make this happen. But I want to see this done within a context that maximizes the value of those reports and that takes full advantage of the expertise that publishing professionals can bring. Mere access simply isn't enough.

The report from the Scholarly Publishing Roundtable,[1] which I was fortunate enough to be a member of, lays out the issues that need to be balanced in order to achieve a truly robust scholarly communication infrastructure that will take full advantage of the opportunities before us.

The key elements are these:

- A focus on the version of record. FRPAA, like the NIH public access policy on which it is based, settles for access to the author's final manuscript version, rather than the final published version – the version of record. In most cases, to meet a current need, this is probably sufficient. But a robust system of scientific communication that is expected to persist over time requires access to the final version, the stewarded version; the version that some entity is going to keep track of, ensuring that corrections are appropriately made, that retractions are handled when necessary, and that permanent access is ensured. At present, we are facing a world in which multiple versions of articles are available through various repositories and websites and it is increasingly difficult for the reader to determine which version they are accessing at any point in time. We need to be developing policies that minimize the potential for this sort of confusion, rather than exacerbating it.

- Standards for interoperability. Text-mining and data-mining published research reports enables deep analysis of multiple experiments, where individual results can be combined to provide accurate summaries across a broad range of experiments. When these sorts of meta-analyses are performed by humans, they can be extremely beneficial but they are very expensive and require a great deal of time. We know that there is tremendous knowledge embodied in existing research reports if only we could effectively analyze it. In order to facilitate this we need to emphasize a standardized structure for these reports. It is not necessary for them all to be in a single repository if they are structured in standardized ways. The National Library

of Medicine's "Journal Archiving and Interchange Tag Suite" is well on the way to becoming a de facto standard.

- A focus on preservation. Most of the scientific literature in biomedicine (the area that I know the best) is now "born digital," and, increasingly, there is no print counterpart. While print continues to be an important medium in some areas of the social sciences and humanities, this is rapidly declining. Although there are several promising approaches at play in providing permanent preservation to born digital documents, these are all still in the experimental stages. Any set of policies designed to provide access to the results of funded research must take into account methods for maintaining access for decades and even centuries into the future. This is not a trivial problem, and librarians and archivists understand that it is a critical one. A recent study reported that only 15% to 20% of the e-journal content in two major research libraries was being preserved by current initiatives.[2]

- A recognition that different disciplines may require different policies in order to achieve maximum benefit. Just as there is a broad range of adoption of "born-digital" objects across scientific disciplines, so are there differences in funding streams and in use of the literature. One of the most contentious areas in the public access debates has to do with the length of the embargo. The NIH policy specifies a maximum of 12 months; FRPAA would shorten that to 6 months. And yet it became clear to us on the Roundtable, as we investigated the differences among disciplines, the 6 months would be longer than necessary in some disciplines and too short to be practical in others. This highlights the fact that disciplinary differences are so wide that any single simple access policy that is intended to apply to all federally funded research will have negative unintended consequences in some areas. While policy development across funding agencies must be carefully coordinated in order to maximize interoperability, some flexibility must be allowed and even encouraged so that agencies can develop policies that are acutely tuned to the needs of their disciplinary communities.

- And yes, free access. Librarians, researchers, educators and publishers are all committed to achieving the widest possible distribution of the results of scholarly research. The opportunity that we have in the digital world is to achieve much wider distribution than could be imagined in the print world. The challenge is to do that is ways that balance the elements that I mention above. My experience with the Roundtable, and my other work with the librarian and publishing communities over the years, leaves me convinced that we will be more effective in developing those solutions when researchers, educators, publishers, librarians and the public work together than if we continue on a path that sets us at odds with each other. Across the publishing industry, in both commercial and not-for-profit sectors, experiments in open access publishing are proliferating. It should be clear to any objective observer that publishers are in no way opposed to open access. They are, quite sensibly, seeking to develop business models that keep their organizations healthy, while maximizing access to what they publish.

Although many in the OA community disparage the contributions of traditional publishers, it is noteworthy that both the NIH policy and FRPAA are explicitly based on the assumption that there is *something* that publishers are doing that is absolutely vital. Investigators are already required by their granting agencies to report the results of their research back to the agency and yet compliance with these requirements is shockingly low. A recent study estimates that only 22% of registered clinical trials whose results were mandated to be reported within twelve months of the conclusion of the study had done so.[3] One could argue that better enforcement of the reporting requirements already in place at the funding agencies, along with mechanisms to make those reports easily searchable and available would do much more toward "making the results of federally funded research available to the public" than would result from the passage of something like FRPAA.

But it is clear that what OA advocates want is not just agency mandated reports – it's the peer-reviewed papers that we crave. And despite the endless claims that since peer reviewers do not charge for their work, there can't really be very much expense involved in doing peer-review, neither NIH nor any other agency has shown any inclination to set up their own peer review system. One might ask the question, if publishers add so little value, and peer review is accomplished at practically no cost, why not empower the agencies to set up their own peer review panels and cut the publishers out altogether? But however much derision and disdain the publishing industry has to put up with, it is clear that they continue to provide *something* that even the most fevered OA advocates believe is essential. The question we should be asking is how do we maximize that value. And yet, a study by Ross and others shows that the results of fewer than 50% of completed clinical trials even show up in the peer-reviewed literature within 30 months of the completion of the trial.[4] Without tackling these issues, far more research results will still remain unavailable to the public than will become accessible.

The Research Works Act was a terribly mistimed rearguard action that was designed to forestall federal legislation of any kind. It is noteworthy that although the Association of American Publishers lauded the act, many individual publishers were quick to distance themselves from it. But again, it should be noted that the goal of the Research Works Act was not to combat Open Access – it was to forestall federal regulation. Publishers don't object to open access, they object to federal regulation. This is hardly surprising. Businesses in general prefer less regulation to more regulation. There is nothing peculiar to the publishing industry in this.

Nonetheless, it seems to me that some level of government regulation is appropriate for at least that portion of the publishing industry that is focused on communicating the results of research. The investments made by taxpayers and the potential benefits to be gained from a well-organized and robust system of scientific communication are such that the inconveniences of some level of regulation are warranted.

But framing the question as, "How do we get better access to the peer-reviewed reports of research" is merely a pale version of the question that we *ought* to be asking, which is, "How we do we take advantage of digital technologies to develop a robust and innovative scientific communication infrastructure that fully takes advantage of the potential of 21st century digital technologies."

We have reached a point, after over a decade of squabbling about Open Access, when the terms of the debate seem to have come entirely down to whether or not someone supports something like FRPAA or something like the RWA. And yet, my experiences with the Roundtable, and the other work that I've done in this area over the past dozen years have convinced me that we can gain so much more than FRPAA offers. But to achieve this we will need to come together as a community that includes scholars, educators, librarians, publishers, and the public at large. We're going to have to listen to each other and be bold and creative. We have to recognize that mere "access" without paying attention to the other aspects of a scholarly communication system doesn't get us very much.

If something like FRPAA is passed into law, it will represent a huge missed opportunity. We have already wasted years on a battle that needn't have been fought. As authorized by the America COMPETES Act in December of 2010, the White House Office of Science & Technology Policy (OSTP) is currently reviewing the results from its latest Request for Information. Their approach to these questions suggests that they are attempting to achieve the kind of balance recommended in the Roundtable report. Anything that the Congress can do to encourage policy development along those lines will be welcome. Anything that impedes or interferes with that work, such as the passage of FRPAA-type legislation will, paradoxically, make the goal of a truly robust open access infrastructure for scientific communication even more difficult to achieve.

Developing federal policies that will maximize the public's investment in research and provide incentives for the development of a robust scholarly communication system is complicated, and achieving the appropriate balance of interests may not be as emotionally satisfying as advocating the simplicity of something like FRPAA. But the American public deserves to have us do this right.

---

[1] Report and Recommendations From the Scholarly Publishing Roundtable. January 2010. http://www.aau.edu/policy/scholarly_publishing_roundtable.aspx (Accessed 3/26/12).

[2] Preservation Status of e-Resources: A Potential Crisis in Electronic Journal Preservation. Fall 2011. http://www.cni.org/topics/digital-preservation/preservation-status-of-eresources/ (Accessed 3/26/12)

[3] Prayle AP, et. al. Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: cross sectional study. BMJ 2012;344:d7373

[4] Ross JS, et. al. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. BMJ 2012;344:d7292

Chairman BROUN. Thank you, Mr. Plutchak. The Chairman will now take a unanimous consent request.

Ms. LOFGREN. Thank you, Mr. Chairman. I would ask unanimous consent to insert in the record a letter I have received on this subject from 52 Nobel Prize winners in favor of open access.

Chairman BROUN. Hearing no objection, so ordered.

[The information appears in Appendix II.]

Chairman BROUN. The—we will recess this hearing. We have to go vote. We will be back—bring this back into order 10 minutes after the beginning of the last vote.

[Recess.]

Chairman BROUN. I call the Committee back to order.

I thank the panel for your testimony, appreciate you all getting through it and I apologize for rushing all of you. But we are going to have five series of vote today so it is going to be—hopefully, we can get through at least one round of questions before the next vote series but we will see.

So if you would please, Members as well as respondents to the questions, please try to make them as quick as possible so we can get through as many questions as we possibly can. And I will remind you that each member has only five minutes.

And at this point, I will open the round of questions by recognizing myself for five minutes.

A question to all of you: is there one-size-fits-all policy that can cover the entire Federal Government or do specific agencies and disciplines require different approaches? I know Mr. Shieber has already—I think it was Mr. Shieber or Mr. Plutchak already dealt with that so let me hear from others and we will hear from all you all real quick. One size fits all, is that appropriate? Dr. Dylla?

Dr. DYLLA. Chairman Broun, in an answer—simple answer, no. If you look at all the scholarship, it is a very diverse array of activities. And if you look at one of the things we have focused on in our testimony and your comments, it is this open-access model where the author pays the full costs of a publisher providing its added value. That works very well for a well funded field that has lots of federal funding. I think it appears to work well for medicine. It probably works well for some disciplines like physics and chemistry, parts of it. But for mathematics and social sciences there is very little grant structure out there. It would be very difficult for that particular one size to fit all.

Chairman BROUN. Mr. Maxwell?

Mr. MAXWELL. There is one aspect of one-size-fits-all that I think in fact is true across the fields and that is that increased public access has substantial public benefit. And that is why I would like to stress the notion of starting with that and if necessary tweaking that. As to Fred's comment, the—what you see in the publishing world are lots of different models and the models have been changing. We are seeing lots of proprietary journals having—becoming hybrids, having some part of open access. You see a remarkable growth of open-access journals, 7,500 over the last decade.

If the—if we start with the notion of increasing public access and that that is a benefit and then if you want to ratchet back and say this might not work in this particular environment, one way of doing it—and we didn't study this in the report but one way of

thinking about it is to say we will have a 6-month embargo across all federal funding but we will wait for a year and each of the departments can have a proceeding in that year that would say with respect to our granting, here is what would be the optimum solution. Here is what fits us. But without a kind of pressure from legislation, this can go on and on and on when the benefits that we lose every day from reduced public access are meaningful. They are meaningful for the progress of science, they are meaningful for innovation, they are meaningful for the benefits for the society.

Chairman BROUN. Dr. Taylor?

Dr. TAYLOR. Thank you, Mr. Chairman. With all due respect, I might question the premise of the question. I don't think mandates are required at all. So whether a one-size-fits-all mandate is okay or not is perhaps not the right question. So, for example, as to these journals made freely available by our choice 12 months after publication, we have decided in our little niche in our business structure that this is an okay way to provide public access to the literature that we publish. We have done that on our own without anybody telling us to do that. And that is true for many other publishers in a number of disciplines.

Chairman BROUN. Anybody else? Mr. Shieber?

Mr. SHIEBER. I was just going to mention in terms—one of the main areas in which people object to a one-size-fits-all solution is in the area of the length of embargos. And it may well be that different fields would like different embargo lengths, but the issue is not that. The issue is what is the embargo length that allows the maximum access while maintaining viability of the publishing system overall? As far as we know, there is no evidence that—certainly NIH embargo length of 12 months is problematic and of all biomedical funders, NIH seems to be an outlier in their embargo length, most requiring a 6-month embargo. So I think the burden of proof would be on finding evidence that a longer embargo would be needed rather than starting with a long embargo and shortening.

Chairman BROUN. Thank you, Mr. Shieber.

Mr. Plutchak, you want to——

Mr. PLUTCHAK. I think I have answered that in my testimony.

Chairman BROUN. Okay. I thought you did.

And my time is expired. I now recognize Mr. Tonko for five minutes.

Mr. TONKO. Thank you, Mr. Chair.

Mr. Plutchak rightly said that the question we ought to be asking is how do we take advantage of digital technologies to develop a robust and innovative scientific communication infrastructure that would then fully take advantage of the potential of our 21st century digital technologies. Much of the innovation in publishing and access is happening organically as we move further into the digital age. Some of this was pushed a little faster by the 2008 NIH policy and then the 2010 COMPETES language, but all of these efforts and the tone of the debate have come a long way since FRPAA was first introduced in 2006.

So let me ask each of our panelists the following: are there downsides to giving markets and technology time to work their magic, especially in the light of the immense amount of activity al-

ready going on and the COMPETES-driven agency efforts? Mr. Plutchak, maybe we will start with you, please?

Mr. PLUTCHAK. Yeah, there is a downside in the sense that I think Mr. Maxwell has addressed, the benefits of making things open quickly, but I always look at these things as a need to balance upsides and downsides. Everything has downsides. And my concern is that by focusing so much on the upside of making openness achievable quickly, we don't pay enough attention to the factors of interoperability, of preservation, the other things that are managed that have to be balanced. And I think overall, those—the downside of potential delay is worth taking that time to make sure that we are developing policies that are really effective for the long-term.

Mr. TONKO. Mr. Shieber?

Mr. SHIEBER. Yes, I am definitely a proponent of letting markets work their magic and for that reason I think it behooves us to make sure that the scholarly communication mechanism that we have is based on a well functioning market. Unfortunately, the reader-side payment, the subscription revenue model has economic properties that lead to market dysfunction and they are very clearly displayed. One nice thing about the writer-side or author-side payment model, publication fee revenue model, is that it doesn't have those same market dysfunctions. So ideally, we would be moving in the direction of open-access journals based on author-side fees.

And to say a word about why that is, from the point of view of a reader, two journals are complements of each other. You would like both. You want access to both. In fact, access to one provides motivation to get access to the other because they cite each other. From the point of view of a writer, two journals are substitutes and therefore compete with each other. So you have a competitive market on the writer side, a monopolistic market on the reader side. So we should be moving towards the writer side in general, that is in the direction of open-access journals.

Mr. TONKO. Okay. Thank you. Dr. Taylor?

Dr. TAYLOR. Thank you, Ranking Member Tonko. I think that, as Scott said, there are downsides to everything and I would agree with his assessment that in this instance the downside to letting the organic growth with some impaling inputs proceed is the right way forward. I think there is also an important distinction between access for the public and access for scholarships. So we can measure access. And as Mr. Maxwell said, the NIH policy is allowing a lot more people to access the content that is available through PubMed Central. But the evidence that I am aware of that scholars are using that content anymore than they were already using journal content available through other means is just not there in terms of citations to the articles that are made available. So I don't think that is not much in the way of downside and I think we should wait.

Mr. TONKO. Okay, thank you. And Mr. Maxwell?

Mr. MAXWELL. Yes, thanks for the opportunity. Just three points—one is there are very good economic indicators and research showing that access does in fact lead to more people reading them, more diverse forms of experimentation, more innovation. And from the MIT economists who do this study, the answer is clear not only

in terms of other people's citations but faster commercialization of the research.

The second point is that I do believe in markets, and the change in the markets as a result of the NIH policy have meant a growth of companies that are taking the PubMed Central information and turning it into useful products that were not available before for scholars, not for the general public. So it goes beyond not only benefits to the general public but the benefits of the scholars.

The third part is if you start from the premise as I do that the taxpayer funded this research and we want to get as much return from this research as we can, I would say you make this available by legislation and you let markets work as they are working now and as people are changing their models to adjust for that. But they—it is—the debate is not about this—the health or wellness of the proprietary publishing industry; it is about how to get the dissemination of government-funded research out to as many people so we can get as much innovation and as much commercialization as we can do it.

Mr. TONKO. Thank you. And Dr. Dylla I will get to you in the next round.

Chairman BROUN. The gentleman's time has expired. I am sorry about that, Mr. Tonko.

Ms. Lofgren, you are recognized for five minutes.

Ms. LOFGREN. Well, thank you very much, Mr. Chairman. And I think this is an important hearing.

I was mentioning that I was—I am a member of the Judiciary Committee and Intellectual Property Subcommittee and we have had a review of this subject there from the copyright point of view, and essentially really this is a copyright exception if you want to get down to what we are talking about. And copyrights are a one-size-fits-all. I mean right now you get the life of the author plus 70 years and you lock it down and there it is. And the question is do we want to make an exception for federally funded research in the interest of science.

And full disclosure, I am a cosponsor of the Federal Research Public Access Act because I think if you take—if you really dig down into it, you have to reach the conclusion that that is going to serve us better.

You know, when we have disruptive technologies, incumbents who have done—I am not critical; I mean they have played an important role but the world changes and it is difficult to change. And I think about there is for-profits that are making, you know—one company that does the outsourcing of journal production for smaller societies earned $3.3 billion last year. I mean that is something they don't want to give up. I understand that. The small societies, this is an important source of revenue for them, I mean, because they are paid fees and it is important I think to America that those societies persist. So the question is how can that happen if this model changes, which inevitably it will?

And finally, I guess it is to the peer-review process and how can we make sure that there is value added in the publication? And I guess that is the question I have for the witnesses. It is worth pointing out that the bill that Mr. Doyle has introduced permits peer review journals to have some exceptions and the like while al-

lowing the author to not be constrained as is often the case when an author gives up their rights to the publisher.

And the question is as the NIH has moved forward in this area is I haven't seen any evidence that there has been an adverse impact. And I am wondering if anybody can cite data to me that delineates the adverse impact from the policy.

Dr. TAYLOR. Thank you. I will start there in the middle of your questioning, Representative Lofgren, when you mentioned that small professional societies are vital to the fabric of the country and scholarship here. As a Director of one of those societies, I couldn't agree more.

As to how to help ASPB and other smaller societies——

Ms. LOFGREN. Actually, the question I had was about data damaged from the NIH policy. Do you have any data on that?

Dr. TAYLOR. The only data I have—and it would not necessarily imply damage—but is—has to do with the usage of our journal articles over time, so the fact that most of the usage is happening after the first six months.

Ms. LOFGREN. I don't have a lot of—Mr. Maxwell, maybe you have data.

Mr. MAXWELL. Well, the only data that we found in the public record were assertions by the publishers that there was damage and a citation of a study that took place before the policy went into effect that said librarians might cancel. We looked very hard for the evidence and we asked people about it. What I might suggest is that the people—that you look at the financial analysts who looked at this science and technology publishing sector and those folks say, one, that the publishers come to them and say this is not a real big deal for us. And they say that their rate of profit growth after the recession is likely to come back up to six and seven percent as opposed to the four percent profit growth during the recession. We don't see economic damage and we don't see the publishers telling the analysts who advise investors that this is a problem.

Ms. LOFGREN. Well, Elsevier had a 37 percent profit margin.

Mr. MAXWELL. They are higher in general than most businesses because they have per article a monopoly. If you want this article, you have to get it from Elsevier unless you are willing to wait for 6 months or 12 months or whatever. And people who need it, in particular the private sector people who would build on this don't want to wait.

Ms. LOFGREN. Well, I just—my time is just about up but I would note I haven't heard anybody cite data, and since this is the Science Committee, we are data-driven. But the loss cannot be calculated because it is what didn't get invented; it is what didn't get discovered because of the lack of information. That can never be calculated.

And with that, Mr. Chairman, I yield back.

Chairman BROUN. Thank you, Ms. Lofgren.

We will go to our second round of questions now, and Chairman recognizes himself for five minutes.

Mr. Dylla's testimony notes the projects at DOE and NSF related to public access. Would everyone please comment on projects such as these, whether they are a model for other agencies to follow and

whether they provide any guidance on how Congress should address public access? The whole panel, we will start out with Mr.—Dr. Dylla. Go ahead, sir.

Dr. DYLLA. The previous question addressed the proper role of government in the marketplace and what has been put in place with the collaborations dictated by the—Section 103 of the COMPETES Act I feel is working out very nicely with two agencies that I have spent 35 years working with—the Department of Energy and the National Science Foundation.

I give you one cogent example of about a half-dozen pilot projects that we have been working on since the COMPETES Act was passed. A very simple question that one of you might ask a program manager from the National Science Foundation or the Department of Energy is, please, Mr. Program Manager, can you tell me how many scholarly publications resulted from your funding last year? And that is a very difficult thing to answer because it turns out an author puts that required information that the underlying research was funded by one of those agencies in an arbitrary place in the journal. It might be in the acknowledgement; it might be in a footnote. It is buried. So for the agency to dig out that information by itself requires quite a bit of data mining.

A consortium of publishers called CrossRef, which includes many libraries and research institutions, has a funded project now that includes those two funding agencies that allow any of the 11 funding agencies to identify the articles that result from their funding with the Social Security number that my colleague, Crispin Taylor, mentioned, the Digital Object Identifier number. So a random member of the public could come to the NSF website and find the author, the institution, and the title of the articles by year, and when they see that DOI number, they would be linked to the version of record on the publisher's site. And if they don't happen to have a subscription, there are 40 publishers within our consortium that are—have been piloting for the last year a modest cost article rental model. We have been doing it for about a year and a half at the American Institute of Physics. Eight million people a year come to our website. They are mainly researchers. Three thousand people came which we assume were random members of the public to take advantage of this service.

But that is just one example of the partnerships and pilot projects that have been put in place as a result of the COMPETES Act.

Chairman BROUN. Mr. Shieber?

Mr. SHIEBER. Yes. I wanted to use Dr. Dylla's example, talk a little bit about that. He talked about a project to allow funders to figure out what—how many papers were being acknowledged as a project I think with NSF now. And first of all, of course, this is not a project to increase access to the articles but it does provide useful information. But it also gives us the opportunity to look at a phenomenon I think is important, which is how openness often trumps other kinds of functionalities.

For the open-access literature, of course, you can imagine automatic computer analysis of the open articles to data mine for exactly this information. In fact, a computer science colleague of mine at Penn State has actually implemented such a system only avail-

able for the open-access literature that automatically data mines the acknowledgements of articles and anybody can search—you can search for an NSF grant number and see all the articles that at least that are openly available for which the data mining is possible that acknowledge the funder. So these kind of projects are wonderful and necessary to the extent that they allow this kind of functionality for the closed literature, but of course, they would be redundant if the literature were just open in the first place.

Chairman BROUN. Thank you, Mr. Shieber. My time is about expired. And by the way, if any of you all—we are going to give you some written questions and any of the questions that you are asked if anybody wants to weigh in and give us an answer, we would appreciate that.

I now recognize Mr. Tonko for five minutes.

Mr. TONKO. Thank you, Mr. Chair.

Dr. Dylla, why don't we go back to the earlier question that unfortunately you didn't have a chance to answer? And that was about the downsides to giving markets and technology time to work their magic. Do you cite any downsides?

Dr. DYLLA. Post the Scholarly Publishing Roundtable and many of its recommendations showing up in the Section 103 of COMPETES, I think most of the downsides have been minimized because you have all cohorts working together, government as the funder of scientific research, the scientists and librarians and publishers working together on these issues as the COMPETES Act proposed. So as you mentioned in your opening statement, Representative Tonko, this has only been in place for about a year and we should let this play out and it is playing out. It won't solve all the problems as quickly as everyone wants those problems solved, but I think it is playing out very nicely.

Mr. TONKO. Thank you.

And a question to perhaps you and Dr. Taylor, grant awards are peer-reviewed before the research is done, and when the research is finished the only peer review the results receive is at the stage of submission for publication. And with our two Ph.D. scientists on the panel, I wonder if you can explain why peer review is important to the progress of science, and how emerging publishing trends may be changing that peer-review process? And by the way, any of you can respond to this but I thought perhaps our two scientists might want to. Dr. Taylor?

Dr. TAYLOR. Thank you very much, Ranking Member Tonko.

So I think peer review is important for a whole bunch of different reasons at the level of journal publication. Most importantly, the function of peer review is to assess the validity of the conclusions that the author of the article is drawing from his or her research work. So I think that is the fundamental level. And that toggle is really important because without that stamp of validity, it is impossible to accept the statements as valid and true.

The more subtle parts of peer review have to do with things like novelty and impact, which are—novelty is not—is easy to quantify but impact is perhaps a little bit qualitative. And this has to do with author choice and their assessment of the novelty and impact of their journal articles in terms of where in the hierarchy of journals they choose to submit their work in the first place. So in addi-

tion to providing a basic level of this is good work, this is sound work, peer review is also sort of a ranking system in terms of the community—the scientific community's assessment of the impact of our work.

Mr. TONKO. And with any of the publishing trends, is there anything that may change that peer review?

Dr. TAYLOR. That is a good question. It is already changing. A very prominent publication called PLoS ONE that is relatively new has a form of peer review that focuses on the first part of what I said—this is sound work—but is less interested in the second part—this is particularly impactful or novel work. And it is a model that is working very well for PLoS ONE and that a number of other publishers are attempting to replicate.

Mr. TONKO. Thank you. Dr. Dylla?

Dr. DYLLA. I think everyone on the witness stand here would agree that the most fundamental value that the publishing enterprise brings to science is this peer-review process. And it is often said, well, publishers don't pay the peer reviewers. As an academic, we do this—we feel this is the most important part of our service to the Academy. We write articles, we expect them to be reviewed by our peers, we become peer reviewers for our colleagues.

But I will give you an example from the American Institute of Physics. We publish 15,000 physics articles every year in five journals. We have to manage the peer-review process for about 50,000 reviewers. That involves paying about 130 Ph.D. scientists all over the world. We pay down their academic time so they can help us manage this. So this is not a cost-free process. Often, when the argument goes to let's just put it up the web and have it crowd-sourced, I remind you the difference between a restaurant you—a restaurant review you might get from one of these crowd-sourced reviews and one that you get from Michelin when an expert spends five different visits. It is not quite the perfect analogy but it works somewhat.

Mr. TONKO. Thank you very much. I yield back.

Chairman BROUN. The gentleman's time is expired. Thank you, Mr. Tonko.

Now, I recognize Ms. Lofgren for five minutes.

Ms. LOFGREN. Well, thank you very much. You know, as I am listening here, you know, I am thinking that really the journals have—they served a function as gatekeepers, and I don't mean that necessarily in a negative sense. I mean there is a pro and con. You know, some say there is good research that never got out because of it and others, you know, you kept out flimflam and there is probably truth in both assertions. But it is a model that I think ultimately is unsustainable in the current environment. I mean you have got—the scientists are not paid anything. The peer reviewers are not paid anything. And I grant you that I am sure it does cost to manage a large number of the peer reviewers, so I am not discounting that and I see that as an identified need to address, because whether Mr. Doyle's bill passes or not, the current model is going to fall apart. Because there is no reason why a for-profit company should make $3 billion when the scientists aren't paid and the peer reviewers aren't paid and there is a possibility to structure this in a completely different way.

And so the question I have is whether or not Congress does something or whether we just let this morph. It is going to change. And how do we identify and protect the things that are valuable to protect, which is the need for peer review—that is not going to be a barrier to publication but certainly of value to the scientific community—and a further need which is to support the scientific societies, especially the smaller ones that get some revenue from the current system and will not be able to do that. When the Internet and connectivity blows them away, how do we preserve those good societies that give value?

And instead of—you know, I have seen this over in the Judiciary Committee where content holders cling to old business models and it is ultimately a failing effort, and it would be smart to think about how do we preserve the good things? So if people have comments on that.

Mr. PLUTCHAK. It is a great question and I think it really is the critical question. A couple of points to make. In regard to the cost issue and publishers, it is important to recognize that Elsevier, although they get all of our attention—and certainly as a librarian I share the frustration of my colleagues—they only control about 20 percent of the journal publishing. Most publishers don't come anywhere near these 37 percent profit margins. And as much as I would like to see a system in which we do not have a single company having that kind of control at those sorts of margins, I think we have to recognize that most publishers don't come close to that.

Look at the Public Library of Science, which is considered to be—their flagship journals are absolutely as high-quality as any. They charge around a little less than $3,000 a piece to publish an article. They recognize that there is that cost. So we have to be a little bit careful about that argument about how much—what is done for free and what is not. But then I think moving forward towards how we pull this together, I don't think it is a question of do we legislate something now or do we just let it go? I think we have to work through the process that brings these stakeholders together.

Ms. LOFGREN. Or not. I mean the Federal Government has paid for this research and we have an interest and the taxpayers have an interest in getting that information out as swiftly as possible if there is no damage done.

Mr. Maxwell, you were looking eager to add in.

Mr. MAXWELL. Well, the only thing I would probably differ with my colleague on the panel about the profitability of the sector. It is not only Elsevier that has had higher-than-normal returns of capital but a number of publishers. In fact, the largest publishers have had a higher return of capital and a long-term ability to raise their subscription rates over the last 30 years.

I do think that Dr. Taylor's point which he made earlier about the success of the PLoS model, which is one in which the author pays and which there is already a kind of working peer-review system for—that produces the quality that he had mentioned means that you have the kind of market innovation that you are seeking. There are ways of doing this, of ensuring quality consistent with open access and we should be encouraging that.

I don't think that we should believe that change will occur as quickly as I think it should absent the presence of legislation and——

Ms. LOFGREN. If I may, my time is running out. I agree which is why I am a cosponsor of the bill, but I think even if the bill does not pass, this model is dead. It is just a question of how long the patient is going to be on life support.

I yield back.

Chairman BROUN. Thank you, Ms. Lofgren.

We will go to our third and final round of questions so I will yield myself five minutes.

To the panel, PubMed Central blocks mass downloading for a variety of reasons, one of which is general piracy from countries such as China. How serious is this issue and are there examples of mass downloading attempts at other online journals and archives?

Second question, should limitations be placed on access to prevent misuse? How should that be decided and who should decide it? And thirdly, should research funded by U.S. taxpayers be made publicly available to non-U.S. taxpayers? Could this even be prevented? The panel? Dr. Taylor?

Dr. TAYLOR. I think I can say a few words about the first and the third question. It is very hard for us to assess whether and if it is happening the extent of damage from mass downloading from sites like PubMed Central because we get very little usage data from PubMed Central. We don't know who specifically is using our content on the PubMed Central site, so we can't speak to downloads—mass downloads; we can't even assess whether it is library users who would be potential subscribers for the Society that are using the content. So I can't answer that question.

In terms of making publicly—making content produced following research funded by the U.S. taxpayer available universally, you know, for a society like ASPB, that is okay. And one of the problems that I have with bills like FRPAA is they are focused only on the content that is produced from research funded by the U.S. taxpayer. As a professional society with members around the world publishing papers from authors around the world, if we want to disseminate that content, we want to disseminate it to everyone, not only with a narrow focus on the United States or on research funded by the U.S. taxpayer.

Chairman BROUN. Dr. Dylla?

Dr. DYLLA. Let me address the download problem. Let me start by saying the scholarly publishing industry is an industry that embraced the Internet in its early days. We are not a Johnny-come-lately to this Internet storm. Our journals at the American Institute of Physics and our sister societies went on—started going online in 1994, about two years after the web was invented by a group of physicists I remind you. And what happened——

Chairman BROUN. Not by Al Gore?

Dr. DYLLA. No, it was not. When we made the print-to-online transition, if you take a typical university in the area here like the University of Maryland, research university—our offices are on their campus—we, in the print days, probably had 10 subscriptions coming into the campus, the main library and sister libraries, we went online. There were suddenly all 40,000 people who had a

Maryland.edu address could get that content. And to address inappropriate use we would have circuit breakers on how many downloads at a particular email address. And if you suddenly saw 10,000 downloads, that was somebody that was doing no good.

And so the industry as itself—and I am a member of the Executive Committee of two of our trade associations for scholarly publishers that include a broad representation of nonprofits and profits—we carefully monitor this but it is—you can only pick and choose an occasional flagrant. And they—it is like piracy in other types of intellectual property.

Chairman BROUN. Okay. Mr. Shieber?

Mr. SHIEBER. Yeah, I did want to question perhaps a presumption in part of your question which is that mass downloading is somehow inherently bad. There are, of course, good reasons to allow mass downloading where rights to do so are available. And I think the prohibitions against PubMed Central mass downloading have to do with the fact that they don't have the rights to allow mass downloading, arbitrary use of the data. They just have rights to allow researchers to download individual articles. For the subset that they do have rights, the so-called open access subset of PubMed Central, they do allow mass downloading. And in fact, Harvard participates in that and downloads those articles en masse. So—and the various uses that can be made including text mining of that data. So to the extent that there were broad rights to allow that more generally, mass downloading would be a useful thing to be able to do.

Chairman BROUN. Well, we are just trying to prevent folks who are up to no good from mass downloading and make it harder to be able to do so that is questioned.

My time has expired. I now recognize Mr. Tonko for five minutes.

Mr. TONKO. Thank you, Mr. Chair.

This public policy area is a bit new to me so I am trying to get a better understanding of the interpretation of the NIH policy. Perhaps across the panel if you could address the adopting of NIH as a policy by agencies, all agencies. Is that a good thing? Is it a functional policy or is there modification that would have to be made? I just would like to hear your thoughts about imposing that or adopting it at all agencies. Sure, Dr. Dylla.

Dr. DYLLA. Going back to the discussion on the floor already about one size doesn't fit all, there are fields where the 6-month proposed embargo I feel would be very inappropriate, even for the fields that the American Institute of Physics represents. The journals that we publish for the physics community and one of our sister sites, the American Physical Society, has equal number of journals. If you look at how those are distributed just here in the United States, 40 percent of our subscription income comes from small colleges and universities that are not research universities.

So if we were to go to a different business model where that content would have to be given away in 6 months, I remind you what the unit of time that a librarian works with. It is one subscription year. At these small colleges and universities, it is the librarian who has intense budget pressures might say, well, we don't want to have those subscriptions anymore. And I think we would be looking at a sizeable disruption of our source of income, which is

used to promote physics for the general welfare. We do a lot of outreach in media and student education activities with those funds. We would find it very disruptive to put our entire business model into that one type.

Mr. TONKO. Mr. Maxwell, we will go across the table here.

Mr. MAXWELL. As I said earlier, I think that the thing that is common across every field is that there would be benefits to the public if materials were made available. And we should start there because that is the undisputable fact across the panel. More access means more research means more commercialization and quicker commercialization and more economic growth. If there is to be a differentiation—because the NIH policy is I think by most lights a very great success—if we take that as a model and say perhaps that the agency could have a policy process that says within a year you decide whether for your community this works, that at least would be able to move this process and get the benefits more quickly available. And then let people try to see whether there is something unique about that. But it is this lost innovation that bothers me when—which would be if we kind of dribble out the policy while people protect their margins.

Mr. TONKO. All right, thank you. Dr. Taylor?

Dr. TAYLOR. Thank you, Ranking Member Tonko.

I would like to just remind everybody that the NIH policy came into effect as a few sentences in an appropriations bill, arguably without due consideration in a committee like this one. That being said, I think there are two parts to the policy. One is the imposition of an embargo period. The second is a requirement that articles be deposited in a centrally operated, government-run archive. I think, you know, we have talked about embargos but with regard to the second aspect of this, in this day and age such an archive is simply not necessary. The information is there available on the Internet and discoverable and usable without having to have it in one place. So in that sense, the NIH is in my opinion wasting money on PMC.

Increasingly, PMC is duplicating. It is looking an awful lot like our platforms, our publishing platforms and it is becoming increasingly competitive with the publishers directly, which I think is a problem, too.

Mr. TONKO. Thank you. We have about a minute left so if I could have Mr. Shieber and Mr. Plutchak split that time, I would love to hear your comment.

Mr. SHIEBER. I won't repeat other comments that have been made but I will make one mention about the issue of whether there should be a central repository or some other mechanism. To the extent that the mechanism in whatever bill that one imagines doesn't mandate a central repository, that is fine so long as the articles can then be mined and mirrored in government-run repositories and more broadly than that. But if—but I would avoid a prohibition against doing the kind of—making the kind of uses that do allow for preservation and government repositories and the like.

Mr. TONKO. Thank you. Mr. Plutchak?

Mr. PLUTCHAK. Just very quickly, the two basic limitations that I see with extending the NIH policy we have alluded to this central

repository which I think is unnecessary and a lack of focus on the version of record. I think that needs to be stressed in any policy.

Mr. TONKO. Thank you.

Chairman BROUN. The gentleman's time is expired.

I now recognize Ms. Lofgren for five minutes.

Ms. LOFGREN. Yes, I am wondering if you could tell me on behalf of the American society for the plants and the physics institute, what percentage of the revenue for your institute, for your society comes from publishing and the publishing—this whole business?

Dr. TAYLOR. Thank you very much for the question, Representative Lofgren. For ASPB it is roughly 80 percent of our revenues come from publishing.

Ms. LOFGREN. Okay. How about physics?

Dr. DYLLA. It is more than 90 percent. We don't have members; we just have Member Societies so a typical scientific society will have income from member dues and from running meetings. All of our income virtually comes from publishing.

Ms. LOFGREN. Okay. So that is, I mean, a major challenge in terms of if we want to preserve a multiplicity of societies as this model changes, whether quickly through legislation or naturally because of the disruptive technologies, we have got to come up with some strategies to support these scientific societies that are workable. And I would hope—this is not the subject of this hearing, but I think it would be a good subject for a hearing. How are we going to be able to do that?

I would like to—let me just ask a question—and I see Mr. Maxwell and Dr. Taylor want to address the first subject, but Dr. Taylor said something I thought was very interesting, which is the exemption being proposed is limited to taxpayer funded and why that would be. And one of the options that we have is to remove copyright protection from publishers who do not compensate—and have the copyright ownership remain only with the author him or herself. And that would be true worldwide because anything published in the United States, no matter who the ethnicity or origin of the author, is subject to copyright law within the United States. I am wondering—that would be even more disruptive than Mr. Doyle's bill but would that actually serve—Mr. Maxwell, you are nodding. I don't know if that is what Dr. Taylor had in mind but that is the logical extension of his comments. How would that work at a Harvard? How would that work, Mr. Maxwell?

Mr. MAXWELL. Well, I think that would throw all the cards up in the air. I am not advocating that and the report didn't advocate that. What I did want to—the point that I did want to make about the societies is that the revenue streams that are generated by—for the societies through the publishing isn't necessarily going to go away. The PLoS model is that the author pays. And so there is a revenue stream that keeps coming. It is not saying you are deprived ultimately of the revenue; you are deprived of a particular business model, which is a subscription-based business model. And more and more of the subscription-based models are becoming hybrids in which they are partially author-paid.

Ms. LOFGREN. If I may, I question whether that can survive as a model either. I mean we are moving from—away from a scarcity-based system, and why would an author pay a gatekeeper when

you can develop through universities or others completely other systems that meet the scientific requirements just as well?

Dr. TAYLOR. Because the scholarly publishers are providing the services of peer review, which we talked about earlier and ranking, which are—the peer review is fundamentally important for the reasons we have discussed. The ranking allows other scholars to assess that particular piece of research in the context of the larger body.

Ms. LOFGREN. Mr. Shieber?

Mr. SHIEBER. Yeah, Representative Lofgren, I think you are exactly right on focusing on where the scarcity is and you are right that the technology changes that balance and the scarcity of access is now completely artificial. But there is still—as Dr. Taylor says, there is still an item of scarcity and that is the effort that goes into—that is involved in the peer-review process that ends up being uncompensated, but the management of the peer-review process is compensated to the publishers, the copyediting, typesetting, and imprimatur of the journal. All of those kinds of things are.

Now, those are all costs—real costs and important—provide important services. They scale with the number of submissions and articles, not with the amount of access. And that is why a revenue model that also is based on that side of the equation makes a lot of sense, in addition to the fact that it is not based on a monopolistic product——

Ms. LOFGREN. Um-hum.

Mr. SHIEBER. —and the fact that it doesn't embed a moral hazard and all kinds of other economic advantages. So I don't—I am not sure I agree with your claim about the unsustainability of the whole system. I think——

Ms. LOFGREN. Well, I always like to throw disruptive statements out there in Christensen form. But I think your comment is a very thoughtful one and actually delineates the very useful role played by the societies and the journals, and the potential way for that to be sustained economically.

Mr. SHIEBER. Can I mention one other thing? It highlights potentially extraordinarily important role for government, which is to—just as government funding agencies and universities have been subsidizing on the reader side by buying subscriptions, government funding agencies and universities should be willing to subsidize on the writer side for these kind of author-side payment models to put them at a level playing field. In fact, Harvard has done just that in cooperation with a small set of other universities by committing to paying reasonable open-access publication fees.

Ms. LOFGREN. I see my time has expired. I would just like to say, Mr. Chairman, I appreciate that you have held this hearing because although the Judiciary Committee with jurisdiction over copyright has looked at this, there is no way that we would ever have gotten into this discussion which is how to organize the funding of science so that this is preserved. So that is very helpful. Thank you very much.

Chairman BROUN. Thank you, Ms. Lofgren. In fact, I have a very strong opinion about property rights, including intellectual property. And I thank you and all the members for their questions.

And I thank all of you all, the Members of the panel, for you all's valuable testimony. If you don't know southern, y'all means all of you all. But Members may have additional questions, and actually, I would like to submit some of our oral questions to all of you all so that if you want to add something if somebody did not have a chance to speak, I would appreciate you all's expeditious answer in writing to those questions. Members can submit questions for written answer. The record will remain open for two weeks for additional comments from Members. The witnesses are excused and the hearing is now adjourned. Thank you all.

[Whereupon, at 11:26 a.m., the Subcommittee was adjourned.]

Answers to Post-Hearing Questions

ANSWERS TO POST-HEARING QUESTIONS

Responses by Dr. H. Frederick Dylla

U.S. House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Investigations & Oversight

Questions for the Record

"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

Thursday, March 29, 2012
9:30 a.m. – 11:30 a.m.
2318 Rayburn House Office Building

Dr. Frederick Dylla,
Executive Director and CEO, American Institute of Physics

Questions submitted by Dr. Paul Broun, Chairman, Subcommittee on Investigations & Oversight

Pricing Issues

1. **Concerns have been raised that the prices of many scholarly journals continue to increase sharply. Do you agree with these concerns? If so, what are the reasons for these price increases and how does this impact the exchange of new scientific knowledge?**

In compliance with anti-trust considerations, publishers do not engage in discussion of journal prices nor do we undertake analyses of journal pricing patterns in the industry. However, there are research patterns to be taken into account and library-produced analyses of serials' pricing.

Research output
Starting in the 1950s, the increase in scientific output as the number of practicing scientists and number of research articles produced have grown at a fairly consistent rate of 3.5% annually. Research funding has also continued to grow, albeit at different rates in different countries at different times. Thus, the scholarly journal literature has expanded considerably.

Since the mid-1990s, scholarly publishers have invested in new technologies that have dramatically increased the utility of published resources, increasing productivity, expanding availability, accelerating the speed of delivery and enhancing discoverability, thus increasing the impact of scientific progress.

Library Statistics

The Association of Research Libraries publishes an annual report of library statistics. The *ARL Library Statistics for 2009 – 2010* (see http://www.arl.org/stats/annualsurveys/arlstats/index.shtml), has a chart tracking changes in serials unit costs, expenditures and number of purchases at ARL universities for the period 1990 – 2010:

- During that 20-year period, according to statistics published annually by *Library Journal*, the average serial's list price has increased by 613%.
- During that same period, the overall increase in libraries' total spending on serials has increased by far less: 327%.
- This is despite a large increase in the number of serials purchased, which grew by 315%.
- The price paid per serial has actually declined in the past decade and is now close to the level it was 20 years ago.
- The reason for this dynamic is attributable to the migration from print to digital content where publishers have been able to offer various bundles or clusters of journals at little or no extra cost through licenses with consortia or state-wide university systems.

The migration from print to digital content has resulted in a many order of magnitude increase in access to scholarly journals by the research community, and a concomitant steep drop in price for an individual to access an article. According to the forthcoming "AAP Industry Statistics Report for Professional and Scholarly Journals for 2010", the research community downloaded over 1 billion articles from a collection of approximately 5800 journals published by 26 scholarly publishers. The collection of journals examined in this study represented more than 60% of the citations reported in the scholarly journal literature for 2010. The summation of reported revenues linked to electronic institutional circulation from these publishers for 2010 was approximately $2.7 billion; therefore, the average cost per download by a researcher was on the order of $2.70 per article download. This cost is in the same range that a group of 40 scholarly publishers is offering to the public for access to journal articles through a new article rental service, which allows individuals to "rent" single articles for periods of days for a nominal fee, similar to Apple's iTunes™ service.

2. **If open access journals and repositories continue to grow and provide more competition to existing journals, does this eliminate any pricing concerns since subscribers will have more choices on where to spend their resources?**

The fundamental challenge for AIP and many other publishers is how to generate sufficient revenue to cover the costs of filtering worthy articles, adding value to an author's manuscript

once accepted for publication and producing a journal. These costs include: editorial office software, peer review management, plagiarism detection, editorial board management, online hosting platform, copy editing, creating files for online distribution and delivery to mobile devices, marketing, referencing, and long-term archiving. This must be accomplished while generating marginal net revenue allowing us to re-invest in future publishing technologies and innovations to best serve our stakeholders (subscribers, authors, scholars, editors, peer reviewers, funding agencies).

There are currently two main business models supporting publication of peer-reviewed journals and the global dissemination of scholarly articles. The predominant business model is based on subscription charges or access fees generally paid for by academic library customers. This is also known as "reader pays". The other model is termed open access or "author pays," in which the author pays an article processing charge, allowing the article to be made available at no charge to readers.

For AIP, the predominant model is a "reader pays" with subscription prices and license fees set in a way that provides customers around the world with a superior value at cost effective prices. Of note, AIP scores extremely well on the "price per article" metric on journalprices.com. AIP also offers tiered pricing for its journal subscriptions where customers are categorized by their level of research activity and amount of usage. We also track other value for other metrics such as cost per download. The trend in increased usage far exceeds price increases. AIP has also enhanced access through consortia licenses, resulting in almost 5,000 institutions around the world having access to AIP journals. In the print world less than 10 years ago, this figure was approximately half of this number.

AIP has responded positively and creatively to open access initiatives in a number of ways. Firstly, AIP's copyright policy allows authors to deposit a version of their paper on the author's and employer's web pages and on e-print repositories, such as institutional and subject repositories. Secondly, AIP's "Author Select" program allows any author to make their article open access by paying an article processing charge. In this model, the author retains copyright under a Creative Commons license. Thirdly, AIP is developing a program of open access journals. With this range of products and services to the author, reader and library communities, AIP will continue to provide a sustainable future to strengthen the scholarly research community.

Regarding pricing, whatever business models are adopted, the challenge for AIP will be to continue to offer a competitive service that is highly valued – whether it is reader pays or author pays. The free market demands that we take these steps. As a medium-sized, non-profit

learned society—AIP publishes 15 leading journals, three of which are published in partnership with other organizations; the institute also publishes on behalf of five smaller Member Societies—AIP is well aware of mounting competition from large commercial publishers throughout the world that are publishing hundreds and in some cases thousands of journals.

Peer review

3. **It appears that the majority of peer reviewers donate their time. If the peer review process is valuable, would AIP support a mandatory payment for peer reviewers paid for by the researcher or the funding entity?**

I am not aware of any peer reviewers for scholarly journal articles being reimbursed for their effort expended in reviewing articles. Academics consider this effort to be a standard practice in the profession. Scholars donate their time to review scholarly articles in their fields of study with the assumption that their peers will reciprocate. This system of volunteer, anonymous reviewing has been in place since external peer review became the norm for many journals (expanding beyond editorial review) since early in the last century. I believe establishing a payment system for peer reviewers would distort and overly complicate the present system which is generally working very well.

Even though publishers do not pay for a peer reviewer's time, a significant cost expended by publishers in producing journal articles is managing the peer-review system. AIP is medium-sized publisher, publishing 15 journals with approximately 15,000 articles published each year in our suite of journals. Approximately 20% of our expenses are for the administrative tasks of managing 30- 40,000 reviewers for this number of articles; this includes stipends for 130 academics that serve as full time or part time editors.

Centralized database vs. a distributed model

4. **With open access archives spreading in number, please share your thoughts for a distributed model in which there are numerous databases versus a centralized model like PUBMED? Which access model would AIP prefer?**

Although a centralized data platform would appear to have some obvious advantages regarding simplicity of operation, the use of a centralized, government-controlled platform for a large body of scholarly content has significant downsides. *This would almost certainly increase federal expenditures.*

A centralized, governmental approach would almost certainly increase federal expenditures and likely draw funding away from research budgets. It also discourages innovation by driving traffic away from private market innovators, including publishers, thus minimizing scientific and commercial opportunities.

An important role for government in this arena would be to drive and fund the development of interoperability standards (the ability for separate systems to work together) and promote the widespread use of such standards. AIP supports the recommendation of the 2010 Scholarly Publishing 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."

It is AIP's position that stewardship of publications should be the collaborative responsibility of the publishing, library, and research communities. Federal 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 and as one of many stewards of specific data platforms that need to be linked across public and private boundaries. Please see page 8 of AIP's prior written testimony for the March 29, 2012 hearing for additional detail.

Notably, the Department of Energy (DOE) Office of Scientific and Technical Information (OSTI) has recently reported recommendations supporting a distributed archive. OSTI is currently engaged in pilot projects with several scholarly publishers to make the citations of DOE-funded journal articles available in the search and retrieval applications operated by OSTI. OSTI is improving DOE's ability to demonstrate its research results by collaborating with journal publishers. This involves engaging with publishers to identify and broaden access to the journal articles reporting on research funded by DOE. AIP applauds this effort as an exemplary way for government to engage in a win-win collaboration. Through these efforts, OSTI reports they have gained insight on how projects like these might be structured to achieve success. In part, OSTI recommends that citations should include abstracts and hyperlinks to a landing page for the publisher's version of the article. In addition, DOE notes that an algorithm should be developed in concert between the journal publisher and OSTI to identify DOE-funded articles in the publishers' data stores and an automated process should be supported to allow for processing new articles submitted to the publisher on a weekly or monthly basis.

Piracy concerns

5. **PUBMED Central blocks mass downloading for a variety of reasons, one of which is journal piracy from countries like China. How serious is this issue for AIP and are there examples of mass downloading attempts at other online journals and archives?**

Piracy is a major concern for all creators of intellectual property, including the scholarly publishers. AIP relies on national and international enforcement of copyright and anti-piracy laws. PUBMED Central's broad, world-wide access enables piracy in countries where intellectual property protection is poor, thus damaging an important US export market. Indeed, according to the Association of American Publishers submission to the OSTP Request for Information on public access, one negative consequence of the NIH policy on public access is a global increase in the piracy of US scientific and scholarly articles.

Estimates are that annual losses from this theft cost US publishers and scientific societies up to $100 million annually. (See page 10 and references therein: http://www.pspcentral.org/documents/OSTPRFI-ScholarlyPublications-final1-10-2012.pdf). This sort of theft and piracy also works against AIP's and other publishers' efforts to develop viable business relationships with Chinese libraries, relationships that adhere to international principles for protecting intellectual property. Thus, not only does piracy sap revenue from US publishers, it also jeopardizes future revenue growth as the US tries to expand its markets into China.

Embargo period

6. **How important is the embargo period to the economic health of journals and to the public interests as a whole? Is the importance of an embargo period likely to decline over time? What is the ideal embargo period, and is it the same duration for all disciplines?**

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. Economic impacts will vary depending upon whether the journals are published weekly, monthly or quarterly.

Publishing is AIP's primary source of revenue, supporting our publishing and outreach activities that serve the broader physics community and the general public. Over 40% of our subscriptions are from non-research institutions. We believe that these institutions would strongly consider cancelling their subscriptions if access to articles was mandated within a short time period of time [such as the 6 month period as currently stated in the FRPAA bills (S. 2096; H.R. 4044) before both Houses of Congress]. Government mandates in this area could impose a significant threat to our business model and sustainability.

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. Based on the evidence related to AIP journals and to journals covering physics and related sciences, significant economic threats to our enterprise would arise with the assignment of minimum embargo periods. In looking at a sample of several physics, 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. Furthermore, AIP journals have a longer cited half-life than their respective physics category averages. This analysis is presented on page 15 of the first attachment to AIP's written testimony for the March 29, 2012 hearing.

Instead of 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. This system allows government to enhance public access in a way that is not only effective, efficient, and sustainable, but also maintains the economic viability of the US scientific enterprise into the future.

To summarize the key components, AIP and a number of our colleagues from the scientific publishing community propose the following plan to improve public access to the results of publically funded research:

1. Scholarly publishers as a group propose modifications to their author submission software so that all journal articles written would include funding agency information along with the standard metadata that is already being deposited in CrossRef (crossref.org) 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, and many have already done so. Such a license provides access to the article metadata, including the critical article identifier (the DOI – digital object identifier). This pilot

project, known as FundRef, was introduced on May 2, 2012 with first results expected by the end of this year. More information available here in the press release: http://www.crossref.org/01company/pr/news050212.html.

2. With the successful implementation of FundRef, agencies would have access to the standard metadata from published articles. By displaying this information on agency websites, visitors—ranging from the research community to the general public—could follow the link [enabled through the DOI] to the publisher's platform where article abstracts are freely available and the Versions of Record (VoR) (maintained by the publishers) are 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.

OSTP Report

7. **The OSTP report required by the COMPETES Act was recently released. Do you agree with the contents of the report and do you feel it addresses your concerns? Were any issues of interest to you not addressed? What specific next steps should OSTP or the Committee take in regards to the issue?**

The OSTP report on public access that was released at the end of March was responding to the one of the COMPETES Bill Section 103 statutes that required OSTP to report on progress on their oversight role for interagency planning one year after the passage of the COMPETES bill. OSTP did a significant amount of work on this complex topic by commissioning and managing two interagency panels as described in the report (one on access to data and one on publications). The primary content in the report refers to developments on public access from three agencies: NIH, which already implemented a public access policy for publications in 2008 by Congressional mandate, and from DOE and NSF, which began development of public access policy as a result of the COMPETES statutes.

The report notes that OSTP will continue to monitor progress at the agency level and leaves open the possibility for additional reports downstream as agencies make further progress on policy development and implementation. From my standpoint, I am pleased that the focus has

shifted to DOE and NSF, which can be models for how the other agencies respond to the related issues of access, interoperability standards and archiving.

As a scientific publisher I am pleased to note the following two quotations from the OSTP report:

"... agencies and public commenters are cognizant of the essential role that publishers and the peer review system play in advancing the scientific enterprise. The PASP therefore set out to explore what steps could be taken to expand public access while preserving the value that publishers provide to the scientific enterprise, creating new business opportunities, and maximizing the economic and societal benefits of the Federal investment in research and the resulting publications."

"It should also be stressed that it is the intention of the Administration to continue a robust dialog with the private sector and the public to ensure that policies developed will benefit the public interest and to maintain a level playing field for all interested parties."

Impact of Access Policies

8. **Do existing public access policies have an adverse impact upon federal grant applications? Does the quality of proposals suffer because researchers do not want to make their work or data public? Understanding that researchers typically want to communicate their findings to the greatest possible audience, is this even a problem?**

AIP cannot comment on the quality of proposals since agencies receive federal grant proposals in confidence and protect the often proprietary nature of their contents. We note that any reported impacts on the type and level of grant applications under the existing NIH public access policy should not be extrapolated to other fields or agencies. Medical research as funded by NIH is markedly different in type, scope and duration from physical science research funded by DOE or NSF, for example.

We can say with confidence that researchers typically can and do communicate their research results as widely as possible. Publishing scientific journals is the primary means by which scientists communicate advances in research to the scientific community. However, journal articles are only one way in which research results are communicated and the scientific community is only one audience. In addition to scholarly journals, scientists also use meetings, lectures, reports, media, and public outreach materials and events to communicate research results. Audiences range from the general public to commercial interests. Also, a variety of organizations, including universities, government agencies, the media, patient advocate groups

and scientific societies such as AIP try to communicate results of scientific research to these varied audiences.

Data Access

9. **Should all information and data associated with federal research published by journals be made public? Is this reasonable or even possible?**

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. As a scholarly publisher, AIP believes that better access and reuse of original research data are to be encouraged at all levels and among all stakeholders. AIP believes that data resulting directly from federally funded scientific research should be made freely available in a sustainable manner, and that this is best accomplished through appropriate policies that leverage public-private collaboration. One way to achieve this balance is for the government to adopt a sensible, flexible, and cautious approach to drafting public access policies. This approach should engage 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.

For example, NSF is considering a proposal from AAS and AIP to fund a pilot project to link data-behind-the-figures and tables with scholarly publications. The pilot involves a reasonably small community: astronomy/astrophysics and plasma physics. Significant issues to be examined include author participation, peer review, and selection of appropriate datasets.

10. **Aside from the publishing industry and scientific society concerns regarding greater public access, does providing greater access impact the quality of scientific research in any way?**

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. As referred to in my oral testimony, all parties in this debate agree that enhanced public access is a good thing. It's good for the economy and it is good for science.

Responses by Mr. Elliot Maxwell

U.S. House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Investigations & Oversight

Questions for the Record

"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

Thursday, March 29, 2012
9:30 a.m. – 11:30 a.m.
2318 Rayburn House Office Building

Mr. Elliot Maxwell,
Project Director, Digital Connection Council, Committee on Economic Development

Questions submitted by Dr. Paul Broun, Chairman, Subcommittee on Investigations & Oversight

Pricing Issues

1. Concerns have been raised that the prices of many scholarly journals continue to increase sharply. Do you agree with these concerns? Since Congress usually avoids becoming involved in pricing debates, why should this be an issue for Congress?
   ANSWER: I am concerned about prices because of their effect on access, particularly to reports on research funded by taxpayers. To the extent that prices rise, fewer institutions and individuals will be able to afford subscriptions, and therefore fewer individuals will benefit from the research funded by taxpayers thus, reducing the return on taxpayer investment in the research. Policies such as the NIH public access policy address this problem of diminishing access without having Congress take any particular position on journal pricing. Continued price increases suggest that the proprietary journals retain pricing power and are not suffering from the existing public access policies in the United States or elsewhere. If they were losing substantial subscription revenue at least some of them might be experimenting with lower prices to encourage increased subscriptions. As the NIH Director has pointed out, subscription prices overall have continued to go up..

2. If open access journals and repositories continue to grow and provide more competition to existing journals, does this eliminate any pricing concerns since subscribers will have more choices on where to spend their resources?
   ANSWER: I believe that the scientific publishing market will be heterogeneous with open access and proprietary journals competing for articles and readers and with increased numbers of repositories providing access to some but not all articles. Increased competition among models is a helpful development, but no substitute for a policy that ensures that government -funded research is available to everyone in order to maximize

the return on taxpayer investment in research, promote innovation, increase progress in science, and provide the resulting economic benefits.

Peer review

3. Although federal taxpayers may have funded the research behind a submitted manuscript, they have not automatically paid the costs of editing and peer review unless a journal charges for that service. How would the debate over public access change if the federal government either:
   - Required that the submitted manuscript, not the peer reviewed version, be made available to the public immediately, or
   - Required that journals be compensated for all reasonably related peer reviewing costs?

   ANSWER: A balanced policy should use the peer review process to identify manuscripts worthy of publication, thereby providing some quality control on what is held out as good science. One should be careful about assuming that there are large costs associated with coordinating such peer review. The results of research funded by taxpayers is provided by researchers to the publishers without compensation because researchers want their work to be broadly available and are competing to have their work made available by the most prestigious outlet. Academics provide their time and effort to review journal submissions, provide editing, and serve on editorial advisory boards, usually without any compensation, because of their belief in the importance of service to their disciplines. There are indeed coordinating costs but these have declined as the Internet is used for the coordination.

   Proprietary publishers should have a reasonable opportunity to recover these coordination costs during embargo periods, but embargo periods should be kept as short as possible so as not to unduly delay the achievement of the powerful benefits that increased access provide for science, innovation, and economic growth. Alternatively these coordinating costs can be paid for by the authors as is the case with the growing number of open access journals or by other means. Competition between the various publishing models should provide some check on the reasonableness of the costs of peer review.

Centralized database vs a distributed model

4. With open access archives spreading in number, please share your thoughts for a distributed model in which there are numerous databases, versus a centralized model like PUBMED?

   ANSWER: I do not have a strong opinion about whether to prefer a centralized model like PUBMED versus a decentralized model. This is a question that OSTP can help

answer. I do believe that the PUBMED model is working and provides an existing vehicle that could be reasonably and quickly modified, if necessary, to accommodate other taxpayer funded research. Just as I favor open access to get the greatest return on NIH funded research, I would favor getting the greatest return on the investment in PUBMED absent a showing that other models are superior.

If a case is made for a distributed model, I would very strongly support requirements for interoperability. I would suggest requiring that all databases be easily searchable from a centralized access point. This would reduce the burden on users who would not have to visit multiple sites and make it easier to publicize the means by which potential readers could obtain access to taxpayer funded research. I would also be very cautious about attempts by any individual database operator to discriminate or act anti -competitively with respect to any other database, That behavior could present a substantial problem that is already apparent in efforts on the Internet by site managers to make their sites "sticky" and to discourage users from leaving the site to visit others. Allowing anti competitive discrimination or unnecessary restrictions on access would defeat the purpose of increasing access to taxpayer funded research.

Negative effects of open access

5. Your testimony indicates that those opposed to greater public access should bear the burden of showing the appropriate delay for works being made available. Why shouldn't federal agencies and Congress wait to change the current model until there is strong evidence that there won't be harm to publishers, many of whom have been working to improve science in their filed for decades? Your study alone does not mean there is strong evidence.
6. ANSWER. One of the aims of my study was to look carefully at each and every claim made by publishers of harm caused by the NIH public access policy. I believe that the report demonstrated, based on ***all*** of the public testimony and regulatory submissions by the proprietary publishers, that they have not provided any convincing evidence of harm. While forecasting great harm in their public policy files, publishers' comments to financial analysts downplay any negative effects and suggest that the publishers will grow their profits in the future. The report also looked at a wide range of studies which demonstrated the value of greater access for science, innovation and economic growth.

   I believe that publishers bear the burden of demonstrating harm given that they are the ones in possession of the data about the impact of increased public access policies on their businesses. That proprietary data is not available to me or to other disinterested scholars. Given the strong case for greater access, if publishers, ask policymakers to delay greater access and to withhold the benefits to science and to the economy, then

publishers should be asked to provide the data that would demonstrate that the harm resulting from increased access outweighs the benefits. This is something that they have not yet done in any public setting.. If they can provide substantial evidence that the harm to science, innovation, and the economy from increased public access outweigh the benefits, I would change my view on what the public policy should be. But it is harm to science and the society, not harm to any particular publisher or group of publishers that is the test. No particular publisher or group of publishers should be protected from the disruptive effects of the Internet if the value for society of increased public access outweighs the as of yet undemonstrated harm to society and the scientific enterprise.

OSTP Report

7. The OSTP report required by the COMPETES Act was recently released. Do you agree with the contents of the report and do you feel it addresses your concerns? Were any issues of interest to you not addressed? What specific next steps should OSTP or the Committee take in regards to the issue?
ANSWER. I see the OSTP report as more of a status update than a deep analysis of the questions regarding public access. I would encourage the committee to support the extension of the NIH public access policy to other federally funded research and to ask OSTP to move expeditiously to address the issues required to obtain the benefits of increased public access as quickly as possible.

Impact of Access Policies

8. Do existing public access policies have an adverse impact upon federal grant applications? Does the quality of proposals suffer because researchers do not want to make their work or data public? Understanding that researchers typically want to communicate their findings to the greatest possible audience, is this even a problem?
ANSWER: I have not seen any evidence that the quality of proposals has suffered because of concerns about public access; the competition for funding is so intense that researchers know they must put forward their best efforts in order to obtain financial support.
But the existing restrictions on access to data, protocols, computer programs etc. and the existing restrictions on reuse need to be the subject of intense review by policymakers as soon as possible. These restrictions are preventing scientists from making the fullest use of research funded by taxpayers. It is these restrictions which are the most important for the working scientists; they have been overlooked in the struggles over increased public access to published works.

Data Access

9. Should all information and data associated with federal research published by journals be made public? Is this reasonable or even possible?

ANSWER: I think there are a number of issues that need to be addressed regarding access to underlying data, protocols, computer programs etc. but we should begin to address these as soon as possible with the goal of increasing access. I think it unlikely that ALL information and data will be made public. Privacy protections, for example, would be likely to require that the identities of participants in studies be stripped out. But much of what is withheld today can and should be made available after a careful review of the issues and the development of appropriate guidelines.

10. Aside from the publishing industry and scientific society concerns regarding open access, what are the science research implications of providing full open access? I believe that in the years to come we will look back at today's arguments over published articles as relatively simple. The much more difficult questions about increasing access to underlying data, protocols, and computer programs will be seen as the next frontier for policymakers--and the area where policymakers can benefit scientific researchers and the scientific enterprise the most. Full open access may not be the most desirable result given issues of security, data integrity, and privacy among others, but the direction in which science is moving is toward greater openness. Policymakers should work with the various stakeholders to determine the limits on openness, if any. Science thrives on openness and policy should limit that openness only when there are substantial societal costs that outweigh the benefits.

U.S. House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Investigations & Oversight

Responses to Questions for the Record

"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

Thursday, March 29, 2012
9:30 a.m. – 11:30 a.m.
2318 Rayburn House Office Building

Submitted by Dr. Crispin Taylor,
Executive Director, American Society of Plant Biologists
June 6, 2012

Responses to questions submitted by Dr. Paul Broun, Chairman, Subcommittee on Investigations & Oversight

The responses below reflect my own opinions, as well as the perspectives of the American Society of Plant Biologists (ASPB), where I am CEO, and the Association of Learned and Professional Society Publishers (ALPSP), a trade association of non-profit publishers to which ASPB belongs. I respectfully refer committee members to my written testimony, submitted on March 27, 2012, for additional information regarding ASPB and ALPSP.

Pricing Issues

1. *Concerns have been raised that the prices of many scholarly journals continue to increase sharply. Do you agree with these concerns? If so, what are the reasons for these price increases and how does this impact the exchange of new scientific knowledge?*

No, I do not agree with concerns that prices for many scholarly journals continue to increase sharply. Indeed, recent data published by the Association of Research Libraries [1] indicate that although list prices for journals are rising, per journal costs to libraries have fallen back to where they were 20 years ago. This is despite the fact that publishers are now producing multiple versions of journals – whether in print, online, or optimized for e-readers – and investing substantially in ensuring that the electronic journal versions are ever more useful and interconnected than they were previously. At the same time, more people in more places are accessing more information via scholarly journals than has ever been the case before, so the substantial investments that publishers have made over the past 15 to 20 years to bring

[1] http://publications.arl.org/ARL-Statistics-2009-2010/

their journals online have enormously improved the capacity for the exchange of new scientific knowledge – and the economic benefits that derive from that knowledge. Indeed, ASPB currently serves over 2,000 research institutions, and every person affiliated with these institutions has instant access to ASPB journal content online.

Furthermore, although costs are obviously important to the purchasers (which, in the case of subscription model journals, are most often libraries), it is important to also consider the value of a product to its end users. It is an axiomatic business rule that higher value products can attract higher prices. And it is also clear that, in any given scholarly discipline, there is a hierarchy of journals, with those valued most highly by a scholarly community publishing the most important and innovative research findings.

As governments around the world continue to invest substantially in fundamental and applied research – and as that research is published in scholarly journals – the total number of journals has risen steadily over the past several decades. With the rising number of journals – and the fact that by far the most prevalent business model in scholarly publishing remains the subscription model – it is undoubtedly true that libraries are finding it increasingly difficult to pay for all the journals their patrons wish to read. But in my opinion that concern has more to do with constrained or shrinking library budgets than it does with the prices of individual journals.

2. *If open access journals and repositories continue to grow and provide more competition to existing journals, does this eliminate any pricing concerns since subscribers will have more choices on where to spend their resources?*

The complexity currently inherent in the scholarly publishing ecosystem makes this a difficult question to answer succinctly. Many journals that, like ASPB's, operate primarily under a subscription access business model, nevertheless offer opportunities for authors to make their articles freely available immediately upon publication. (And many make all of their content freely available after an embargo period that meets their business needs and the needs of their reader communities.) Meanwhile, some journals (a minority) are operating under an "author pays" business model, under which all research articles are made Open Access upon publication. And other models abound. In all instances, publishers host the content on their own journal platforms (or those leased from third parties) – these are the formal, definitive, Version of Record (VoR) articles.

Repositories – whether organized around an institution, a field of study, or, in the case of PubMed Central (PMC), a government agency – operate in parallel to the formal scholarly publishing marketplace. As such, they either duplicate work done by publishers – in my opinion an unnecessary and wasteful practice in a time of tight budgets everywhere – and/or

provide access to article versions that are not definitive. This is important, because studies have demonstrated that researchers prefer to access the publisher-created and hosted VoR from a peer-reviewed journal as the authoritative, definitive version, over versions in subject or institutional repositories[2],[3].

So, the growth of open access journals and repositories does not eliminate pricing concerns (which, as I pointed out earlier, are in my opinion largely misplaced). However, unfunded mandates requiring publishers to release the value-added works they produce at an arbitrarily defined time point following publication interfere with publishers' abilities to support themselves, compete and experiment with business models, or create and improve products and services that could benefit to the publisher and the consumer. Three recent reports provide evidence that these concerns are valid. A study of journal publishing in the humanities and social sciences, fields in which scholarly articles are known to have a very long shelf live, concluded that the imposition of embargo periods like those in biomedicine community could threaten the sustainability of these journals[4]. A 2006 Publishing Research Consortium (PRC) report concluded librarians would likely cancel scientific journal subscriptions if content is available for free, even with embargo periods[5], a finding that has been substantially reinforced by a more recent survey conducted by PRC on behalf of ALPSP and the UK Publishers Association[6].

The bottom line is that a fair market has to find its own business models; they cannot be imposed. Although governments, as I stated in my testimony, have a legitimate interest in the scholarly publishing domain, that interest is most appropriately promulgated via policies and approaches that foster non-disruptive change. Disruptive change (perhaps inadvertently triggered by a government tipping the scales too suddenly toward one business model or another) is likely to be far more harmful to scholarly communication in the short to medium term than a steady evolution in the direction of further increased public access.

Neither ALPSP nor ASPB is in favor of mandated deposit to centralized open repositories. In addition to significant concerns about long-term sustainability and piracy, open repositories have deleterious effects on the publishing model; for example, NIH does not currently provide publishers with full, detailed usage statistics from PMC, which means publishers are unable to supply libraries with the complete picture with regard to their institution's use of a wide range of journals. Such usage data are crucial in determining renewals and while this

[2] http://www.peerproject.eu/reports/ D4.2 PEER Behavioural Research – Final Report
[3] http://www.publishingresearch.net/projects.htm Research Publication Characteristics and Their Relative Values
[4] The Future of Scholarly Journals Publishing Among Social Science and Humanities Associations," Report on a study funded by a Planning Grant from the Andrew W. Mellon Foundation (February 2009), available at: http://www.nhalliance.org/bm~doc/hssreport.pdf
[5] Publishing Research Consortium Report "Self-Archiving and Journal Subscriptions: Co-existence or Competition" (July 2006). Accessible at http://www.publishingresearch.org.uk/documents/Self-archiving_report.pdf
[6] http://www.alpsp.org/Ebusiness/AboutALPSP/ALPSPStatements/Statementdetails.aspx?ID=407

situation persists, librarians are making cancellation decisions on the basis of incomplete usage data.

Peer review

3. *It appears that the majority of peer reviewers donate their time. If the peer review process is valuable, shouldn't their time be paid for, either by the researcher or funding entity?*

Peer review by scholarly journals is a multifaceted process. Ultimately, the decision on whether or not to publish a manuscript is made by a member of the journal's editorial board. These individuals are frequently provided modest honoraria by publishers, in recognition of their expertise and the time and energy that they put into their decisions. Chief editors are the journals' ultimate arbiters; they can also be compensated or reimbursed by publishers. (ASPB provides its journal editors-in-chief with both an honorarium and an institutional allowance that they use to offset the time that leading our journals takes away from their own research programs.) Publishers also incur additional expenses for managing the peer review process (which typically involves both staff and an online infrastructure) and for holding periodic meetings of editorial boards at which journal policies, practices, and scholarly directions are discussed and determined.

All that said, it is indeed the case that each manuscript deemed worthy of review by an editor is evaluated by peer scholars (typically two or three per manuscript) who usually are not paid for their contributions. However, providing service as a peer reviewer – whether to journals, to agencies that are evaluating the merits of submitted grant applications, or to a professional organization that is judging students' poster presentations at a scientific meeting – has long been a central tenet of academic/scholarly culture. The peer reviewer's reward comes in the form of formal acknowledgments of their service from journals and other entities – and the opportunity to list that peer review service on a CV as both a contribution and an implicit recognition of their expertise and the respect held for their opinions.

It is my assessment that the current system is operating very well. Furthermore efforts to compensate peer reviewers would be complicated by the fact that the journal peer review process is set up to evaluate manuscripts no matter where they come from; so it is frequently the case that an American peer reviewer is asked to assess a manuscript submitted by a Japanese scholar – and vice versa. It would be impractical – and, in my opinion, inimical to scholarly discourse – to endeavor to restrict peer review on the basis of which country a manuscript came from or a reviewer was based in.

4. *Although federal taxpayers may have funded the research behind a submitted manuscript, they have not automatically paid the costs of editing and peer review unless a journal charges for that service. Would the debate over open access be less divisive if the federal government either:*
    - *Required that the submitted manuscript, not the peer reviewed version, be made available to the public immediately, or*
    - *Required that journals be compensated for all reasonably related peer reviewing costs?*

Regarding a requirement that the submitted manuscript, not the peer-reviewed version, be made publicly available, such an approach might suffice for public access, although with the important caveat that it might actually put a brake on innovations derived from research if it impedes scholars from submitting manuscripts that include patentable information until after those patents had been applied for or awarded. However, requiring that the submitted manuscript be made available is not necessary to support scholarship and research, where, as noted above, data show that more scholars already have greater access to more scholarly information than ever before. Moreover, as also noted above, there are data indicated that scholars strongly prefer to have access to the definitive VoR, which typically resides on the publisher's site.

A far more appropriate – and immediately available – option would be to make grantees' reports available to the public via the granting agencies' websites. Some already make such research reports available (e.g. the DoE Information Bridge[7]), but others do not. Making all such reports freely available would solve the "public access" issue. Moreover, by cooperating with publishers – as they already are through the experimental FundRef program[8] – agencies would be able to link grantee reports directly to any peer reviewed journal articles that are derived from the funding, regardless of the access mechanism via which those articles are available.

Regarding a requirement that journals be compensated for peer review costs, such an approach would not account for the many other investments that publishers make in post acceptance processing, production, hosting, distribution, long term preservation and archiving, and so on. It also, as previously mentioned, would not take account of the fact that most scholarly journals review and publish articles from all over the world, not just the US. So although the US Government's interests re public access may pertain specifically to those articles reporting on research funded by US government agencies, the larger interest – re scholarship and using research findings as a basis for economic development and growth – would be much more effectively served by developing

[7] http://www.osti.gov/bridge/
[8] http://www.crossref.org/fundref/

policies and approaches that move toward making the entire corpus of the scholarly literature available to all.

Centralized database vs. a distributed model

5. *With open access archives spreading in number, please share your thoughts for a distributed model in which there are numerous databases versus a centralized model like PUBMED? The Internet itself is a distributed network by design. Shouldn't the same be true for how federally funded research is made available to the public, using search engines to find what is of interest?*
Yes, the distributed nature of the Internet is a fundamental feature of its design and operation, and there is no sound reason to require that federally funded research be made available via centralized repositories, like PMC. In an interconnected age, with current and ever-improving technology, centralization is not required and moreover, requires unnecessary duplication of effort at considerable expense. Indeed the report from the Scholarly Publishing Roundtable (on which I served) in January 2010[9] recommended decentralization to achieve the interoperability needed to "enhance the impact of the scholarly literature and ignite the generation of new knowledge".

So, it is easier, quicker, and cheaper to continue to support the distributed model than to invest in centralized databases that not only duplicate effort but can also undermine publishers' capacity to innovate and secure returns on their investments. The definitive VoR exists on publisher websites, and that is where we should be aiming to point people interested in finding research articles.

The capacity for search engines to discover distributed information is enormously enhanced by the development and utilization of standard approaches and metadata. Publishers have gone to considerable lengths in developing tools to ensure interoperability between different access systems. For example the Digital Object Identifier (DOI[10]) system, to provide persistent identification of digital objects, the CrossRef [11] organization and its various ongoing projects aimed at connecting users with primary research content and the Open Research and Contributor ID (ORCID[12]) initiative, to solve author name ambiguity in scholarly communications and latterly resolving institutional naming ambiguity.

[9] http://www.aau.edu/WorkArea/DownloadAsset.aspx?id=10044
[10] http://www.doi.org
[11] http://www.crossref.org
[12] http://orcid.org

Publishers have also collaborated with librarians and database providers to establish COUNTER (Counting Online Usage of NeTworked Electronic Resources[13]), which has produced an international set of standards and protocols governing the recording and exchange of online usage data. These standards enable publishers to better understand the usage patterns of their digital content and for librarians to track the usage of their digital collections. A variety of Internet search engines, abstracting services, and other tools do an excellent job of ensuring the discoverability of research, and innovations and advancements of these and other tools continue to be developed.

An important role for government in this arena would be to drive and fund the development of interoperability standards that would facilitate and enable ever richer connections among journal articles and other types of scholarly information available online and promote the widespread adoption and use of such standards. Indeed, ASPB supports the recommendation of the Scholarly Roundtable Report that states that government policies should be guided by the need to foster interoperability and encourage "additional multiagency programs supporting research and development to expand interoperability capacity and to develop and promote additional interoperability practices and standards." The Scholarly Roundtable Report further notes that the NSF, DOE, and other agencies provide important funding for the development of interoperability capacities through their cyberinfrastructure programs.

Piracy concerns

6. *PUBMED Central blocks mass downloading for a variety of reasons, one of which is journal piracy from countries like China. How serious is this issue and are there examples of mass downloading attempts at other online journals and archives?*
Journal content and manuscripts hosted by PMC are universally available for free, including in countries, like China, in which intellectual property protection is relatively poor. So, ASPB and ALPSP are concerned about this matter.

That concern is heightened by PMC data showing that more than half of all its users are from outside the US. This repository is therefore likely to have an adverse impact on the export market for the US publishing industry which, in total, employs around 50,000 people and contributes roughly $3.5 billion to the US balance of trade.

But digital piracy is a more general concern. A recent article in the ALPSP Journal *Learned Publishing*[14] provides some examples of the challenges in controlling online piracy and protecting the copyrights in the works for the authors and publishers. The

[13] http://www.projectcounter.org/
[14] Ed McCoyd. "Online piracy of publishers' content: a primer," Learned Publishing, 2012 (25); 21-28.

Association of American Publishers reported data to the US Department of Commerce in December 2010 on piracy monitoring of 10 publishers[15]. The publishers detected more than 299,000 online files infringed copyrighted products within a 24 month period. The seriousness of piracy needs to be further explored so that the content that publishers make available is properly represented and not resold without permission. This is a concern regardless of whether the content is freely available or not.

That said, the need for archiving digital information has been recognized by publishers, librarians, funders and researchers alike. Collaborative projects already exist to ensure the long term preservation of scholarly information through initiatives such as Portico[16], LOCKSS[17], and CLOCKSS[18].

The embargo period

7. *Mr. Dylla's testimony stated that there should be different embargo models for different fields. Do you agree with his testimony? How important is the embargo period to the economic health of journals and to the public interests as a whole? Is the importance of an embargo period likely to decline over time? What is the ideal embargo period, and is it the same duration for all disciplines?*

Regarding different embargo periods for different fields, yes, I completely agree with Dr. Dylla's testimony that these are needed.

The importance of embargo periods can be assessed by examining trends in article usage and impact over time; the journal half-life forms a useful measure of these trends in any given discipline. For example, the American Physiological Society reports journal half-life from 4.3 to over 10 years[19]. The quarterly journals of the American Anthropological Association also have a cited half-life of over 10 years, and 90% of downloads occur 12 months after the date of publication. In mathematics papers published in 2009, 50% of citations were found to be to papers originally published before 1999, with 20% of citations to papers published before 1985[20].

[15] Comments of the Association of American Publishers in response to the Department of Commerce 'Inquiry on Copyright Policy, Creativity and Innovation in the Internet Economy'. Association of American Publishers, 2010, p. 11. Available at http://publishers.org/_attachments/docs//issues//piracy-aap%20response%20to%20noi%2012-10-2010.pdf

[16] http://www.portico.org/digital-preservation/

[17] http://www.lockss.org/lockss/Home

[18] http://www.clockss.org/clockss/Home

[19] http://www.the-aps.org/publications/journals/info/impact_factors.htm

[20] http://www.msri.org/attachments/workshops/587/MSRIfinalreport.pdf Donald E McClure (2011) Dynamics of Mathematics Journals, 2000 to 2009

So, embargo periods are crucial for subscription-based publishers, because they provide the temporal window during which those publishers monetize their products. As they exist now, with the exception of the NIH requirement regarding PMC, they are not arbitrary but have been calibrated by publishers seeking to balance both their business needs and obligations and the needs of the communities they serve. On top of this, many publishers have free access policies that allow authors to choose whether to make their articles freely available immediately upon publication. Publishers are also innovating and adopting other models, such as inexpensive pay-per-view and article rental models that are intended to provide access outside of subscriptions at low or very low cost.

It seems unlikely to me that the importance of the embargo period will decline with time, unless it happens in the context of a collaborative, sustainable, and orderly evolution of journal business models toward further increased public access.

8. *How important is the embargo period to the economic health of ASPB, other journals, and to the public interests as a whole? Is the importance of an embargo period likely to decline over time?*

The embargo period is vitally important to ASPB; our usage data indicate that over half of our journal *Plant Physiology*'s online usage – and, thus, half of the journal's potential value, comes after 12 months. Nevertheless, because we have found that we are able to sustain subscriptions with a 12-month embargo window, we have opted for many years to make the entire content of both ASPB journals free 12 months after publication.

ASPB is exploring novel ways in which to raise revenues to that will sustain the society's efforts to provide ongoing support for plant biology and plant biologists. To the extent that those efforts are successful and allow us to adjust our journals' business model, it is conceivable that the importance of a 12-month embargo period may decline over time. But this is only feasible if government policies regarding public access continue to allow ASPB the space and time it needs to innovate. Mandates that shorten embargo times beyond that which ASPB has determined is sustainable will close the window on ASPB's capacity to innovate. Put simply, shortened embargo periods would undermine the society's capacity to invest for the future.

<u>OSTP Report</u>

9. *The OSTP report required by the COMPETES Act was recently released. Do you agree with the contents of the report and do you feel it addresses your concerns? Were any*

issues of interest to you not addressed? What specific next steps should OSTP or the Committee take in regards to the issue?

In general I found the OSTP report to helpful in its descriptions of OSTP and NTSC activities and in its call for continued analysis and engagement. I particularly welcome the following passage:

> [Responses to the RFI] *have indicated a strong support for broad public access to scholarly publications resulting from federally-supported research. Similarly, agencies and public commenters are cognizant of the essential role that publishers and the peer review system play in advancing the scientific enterprise. The PASP therefore set out to explore what steps could be taken to expand public access while preserving the value that publishers provide to the scientific enterprise, creating new business opportunities, and maximizing the economic and societal benefits of the Federal investment in research and the resulting publications.*[21]

The report also emphasizes the need for US government agencies to engage with similar organizations in other countries – a point that resonates strongly with me and echoes another of the recommendations of the Scholarly Publishing Roundtable – and it points toward a continued collaborative process involving all stakeholders in developing and coordinating any agency policies, as envisioned in the America COMPETES Reauthorization Act. In particular, the report highlights productive public-private partnerships underway between the DOE and the NSF and publishers.

On the other hand, the report also makes some assertions regarding the current policy regarding mandated deposit of articles in PMC that I consider to be either questionable or inaccurate. For example, I would question the assertion that "*NIH awards are used to produce peer-reviewed papers*"[22], because scholarly papers, per se, are not produced with award funding – and award funding does not cover the added value that publishers provide. Furthermore, although the statement that "*This policy, and its subsequent fine tuning, has led to a dramatic increase in the number of NIH papers posted to PMC*"[23] may indeed be true, the fact of the matter is that this "dramatic increase" has come largely as a result of publishers voluntarily submitting author manuscripts on the authors' behalf. Without such cooperation from publishers, the number of deposits would be much lower.

[21] Page 5, from the section describing the "Task Force on Public Access to Scholarly Publications."
[22] Page 11
[23] Page 12

Impact of Access Policies

10. Do existing public access policies have an adverse impact upon federal grant applications? Does the quality of proposals suffer because researchers do not want to make their work or data public? Understanding that researchers typically want to communicate their findings to the greatest possible audience, is this even a problem?

No; in my opinion, public access policies have no adverse impact on federal grant applications. There is no problem regarding access to the scholarly literature in academia; indeed, as I stated previously, publisher and librarian policies over the past 15-20 years have rendered enormously more material available to many more scholars than was ever available before journals started publishing on the Internet. A recent survey from the Publishing Research Consortium found that 97% of researchers in North America have very or fairly easy access to research journals[24]. This study also demonstrated that North America enjoys one of the best 'access to information' versus 'importance of that information', profiles of any of the regions investigated.

Furthermore, grant applications and the material in them are confidential unless/until the grants are funded, and researchers routinely include a mix of unpublished and published data and information in their applications. The published data, of course, have been vetted via peer review and publication, and the journal in which they have appeared is one important measure of their quality and impact. Unpublished data included in grant applications are assessed by peer reviewers identified by the agencies to review those applications in the context of the published data. To put it another way, unlike journal articles, grant applications are not a medium for communicating research findings to a broad community; they are a (confidential) vehicle for proposing novel research work that builds on prior work.

Data Access

11. Should all information and data associated with federal research published by journals be made public? Is this reasonable or even possible?

This is already happening; it is exactly what publishers (and those that run databases) do, and the entire content of all the journals that I know of is publicly available. But making something public is not the same as making it freely available, whether immediately or after some period of time (i.e., a publisher-determined embargo period). So, in many instances, there is a fee associated with accessing publicly available journal articles, especially shortly after they are first published (or "made public").

[24] http://www.publishingresearch.net/projects.htm Access vs. Importance

It is indeed feasible to work toward a goal of immediate free public access to the scholarly literature in some fields. But such an objective cannot be sustainably or reasonably accomplished via mandates – particularly if they are unfunded or unilateral; it will require international collaboration among publishers, agencies, and other stakeholders, as well as the development of robust standards for interoperability and exploration of new business models in controlled environments.

12. *Aside from the publishing industry and scientific society concerns regarding open access, what are the science research implications of providing full open access?*
It is my opinion that the business and practice of scholarly publishing have changed enormously over the past two decades. Publishers are innovating furiously, developing novel products and services that take advantage of the Internet, and exploring new business models. As a result, the scholarly publishing milieu is already richly supporting – and greatly benefiting – research and discovery.

Presuming that scholarly publishing is allowed sufficient space to continue to innovate in a sustainable manner, I am confident that the industry will further support research and discovery. But I continue to believe that that it would be in the best interest of the US Government and all other stakeholders to strike a balance between public access and the needs and interests of the scholarly publishing industry because of the positive impact and value the latter brings to the progress of science and its contributions to American society and the national economy. Such a balance can be achieved based on shared principles, including the importance of peer review, the recognition of economic realities, the exploration and adoption of adaptable and viable publishing business models, the need to ensure secure long-term archiving and preservation of scholarly information, the increasing need to establish connections among disparate information sources and repositories online, and the desirability of broad access.

One way to achieve this balance is for government to adopt a sensible, flexible, and cautious approach to drafting and revising public access policies—an approach that engages all concerned parties, including federal agencies, scientists, university administrators, librarians, publishers, and the public. Indeed, it is ASPB's position that government agencies should develop flexible public access policies through voluntary collaborations with nongovernmental stakeholders, including researchers and publishers. Policies should be guided by the urgent need to foster global interoperability of information across multiple databases and platforms. Agencies' efforts and resources could then be directed toward facilitating cyberinfrastructure and collaborative programs with and among agencies and other stakeholders to develop robust standards for the structure of full text and metadata, navigation tools, and other applications to achieve interoperability across the scholarly literature and other information sources.

Responses by Mr. Stuart Shieber
"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

Thursday, March 29, 2012
9:30 a.m. – 11:30 a.m.
2318 Rayburn House Office Building

Mr. Stuart Shieber,
Director, Office for Scholarly Communications, Harvard University

Questions submitted by Dr. Paul Broun, Chairman, Subcommittee on Investigations & Oversight

Pricing Issues

1. *If open access journals and repositories continue to grow and provide more competition to existing journals, does this eliminate any pricing concerns since subscribers will have more choices on where to spend their resources?*

No, waiting on open-access journals and repositories to grow does not eliminate pricing concerns. The issue of choice alluded to is a chimera. Any given article is published only in a single journal. Therefore, from the point of view of readers, journals are not economic substitutes that compete for readers' resources, since a reader cannot turn to a "competing" journal to read the same article. Rather, they are economic complements. So providing more journals, even open-access ones, does not add competition to a market that is structured as the journals market is.

Compounding this problem of the natural monopoly that inheres in the subscription journals market, the market is also subject to a moral hazard, since the readers of the content (the researchers) are not the purchasers of the content (the libraries). This leads to a market dysfunction that publishers can leverage to continue the hyperinflation of journal prices. Again, the moral hazard insulates subscription journals from any real competition from open-access journals, because from the point of view of individual readers, access to both open-access journal articles and articles from subscribed journals appears to be free.

For these reasons, the actual competition between subscription and open-access journals is not on the *reader* side, but on the *author* side. Both kinds of journals do compete for

submissions from scholar-authors. From the point of view of authors, different journals are economic substitutes. But here again, subscription journals have an inherent advantage. To an author, subscription journals are free to publish in, because the government through its funding agencies, together with the universities, support the subscription journals by underwriting the subscription fees that the publishers charge. But many open-access journals, since they eschew subscription fees, recoup revenue by author-side publication fees. An author in choosing a journal must pay this fee to publish in the open-access journal, but typically does not pay a fee to publish in the subscription journal. The open-access publication fee thus serves as a direct disincentive for authors to publish in open-access journals.

The solution to this anti-competitive situation is for the government to explicitly support open-access journals in the same way that it supports subscription journals: Government funding agencies should provide incremental funds to grantees to pay reasonable open-access publication fees for articles that they fund. Doing so would eliminate the disincentive for authors to publish in open-access journals, placing open-access and subscription journals on a level playing field, and opening up the market to true competition on the author side.

Because the underlying economics of subscription publishing means that there can be no competitive market on the reader side, the role of the federal government there is to at least make sure that the reading public can have some level of access to the research literature. Policies like the NIH Public Access Policy achieve this. The pending FRPAA bill would go a long way toward maximizing such public access.

Peer review

2. *It appears that the majority of peer reviewers donate their time. If the peer review process is valuable, shouldn't their time be paid for?*

Not only do the *majority* of peer reviewers donate their time, essentially *all* peer reviewers donate their time. As academics, we view our participation in the peer review process as editors and reviewers as part of our responsibility to the social structure of scientific research.

No one is calling for peer reviewers to be paid for their time. The substantial additional

funds required to pay peer reviewers for their time at market rates would merely exacerbate the current unaffordability of the journal ecology. Academics have established a social norm that participation in the peer review system is expected as part of our contribution to the advancement of knowledge. Eliminating this social-norm-based system in favor of a market norm is likely to be counterproductive. Research in behavioral economics has convincingly shown that when systems move from social norms to market norms, expectations of the participants can change abruptly and counterproductively. In the case of reviewing, it may be impossible to provide sufficient payment to academics to convince them to participate merely as a market function.

Fortunately, so far at least, academics have been willing to provide these reviewing and editorial services pro bono. However, reviewers are not likely to have infinite patience with a social norm that, because of dysfunctions in the market, is providing excessive profits to publishers, while not providing them with access to the literature that they voluntarily write, review, and edit. The solution is not to pay the reviewers. The solution is to adjust the scholarly publishing market so that it does not exhibit the market dysfunctions. Moving to an open-access journal publishing system would achieve this adjustment.

The government could take two actions that would have the effect of supporting the open-access journal publishing system. First, government funding agencies could require that any articles deriving from government-funded research be made publicly available. The FRPAA legislation would achieve this to a large extent, though the embargo period is still problematic. Second, government funding agencies could provide incremental funds to grantees to underwrite reasonable publication fees for articles in open-access journals (just as funders and universities now provide the subscription fees on behalf of readers of the subscription journals).

3. *Although federal taxpayers may have funded the research behind a submitted manuscript, they have not automatically paid the costs of editing and peer review unless a journal charges for that service. Would the debate over open access be less divisive if the federal government either:*

- *Required that the submitted manuscript, not the peer reviewed version, be made available to the public immediately, or*
- *Required that journals be compensated for all reasonably related peer*

reviewing costs?

The premise underlying the question is that federal taxpayers via government funding agencies fund the research that underlies an article, but do not fund its editing, peer review, or other publisher services.

But subscription journal publishers *are* compensated for these services. They are compensated in the form of subscription revenues, revenues which are made possible because of the monopoly rights of copyright voluntarily given to the publisher by the author. The publisher, having been given the copyright, then monetizes that copyright, using its ability to limit access to only those willing and able to pay. The average value of that copyright monetization, according to the Scholarly Publishing Roundtable, is about $5,000 per article. Given that subscription publishers are already compensated by the funded authors through monetization of copyright, asking for further compensation constitutes double-dipping.

The requirement to provide public access to an author's final manuscript to the public does not affect this monetization of copyright. As demonstrated in detail in the 2012 report "The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?" by Elliott Maxwell for the Committee for Economic Development, "There is no persuasive evidence that increased access threatens the sustainability of traditional subscription-supported journals, or their ability to fund rigorous peer review." Thus, public access does not change the fact that subscription journal publishers are already compensated for their services.

The subscription fees that compensate the publishers for their services are paid overwhelmingly by academic research libraries, under subvention by overhead charges on federal grants. It is the universities and funding agencies who thus pay for the publisher services.

For publishers that do not use the subscription revenue model, so-called open-access publishers, this argument does not hold. Open-access publishers do not receive revenue by monetizing copyright. For those publishers, funding agencies should, and are willing to, underwrite open-access publication fees. Since these fees are the only revenue for the journal, there is no double-dipping going on.

However, the mechanisms by which most funding agencies underwrite open-access publication fees are not ideal. (There are notable exceptions, in particular, the Wellcome Trust and Howard Hughes Medical Institute.)

1. Funding agencies require that authors pre-budget for these fees, despite the fact that it is impossible to predict the required amount of open-access publication fees.
2. Funding agencies require that grantees use direct grant funds for publication fees, forcing grantees to trade off using grant funds for publication fees against using grant funds for other research uses, and so disincentivizing publication in open-access journals.
3. Funding agencies limit underwriting of the fees to the duration of the grant, even though publication of the articles based on grant research often comes after the grant has ended.

All of these problems can be easily corrected, however, by funding agencies making available a small amount of *incremental* funds to grantees for the purpose of underwriting reasonable open-access publication fees. These funds would be available during the grant period and for a limited period thereafter. I've described this proposal in detail at <http://hvrd.me/sLhk3H>.

The embargo period

4. *Mr. Dylla's testimony states that there should be different embargo models for different fields. How important is the embargo period to the economic health of journals and to the public interests as a whole? Is the importance of an embargo period likely to decline over time? What is the ideal embargo period, and is it the same duration for all disciplines?*

The ideal embargo period is zero. Open-access journals provide this embargo period by definition, as well as providing a host of other advantages over subscription journals, and can do so in economically sustainable ways. Major open-access journal publishers — Public Library of Science, BioMed Central, Hindawi Publishing — are already operating in the black, demonstrating that embargoless publishing is sustainable.

Furthermore, as has been repeatedly noted, there is no evidence that shorter embargo

periods have any adverse effect on subscription journal revenues. The issue is covered exhaustively in the 2012 report "The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?" by Elliott Maxwell for the Committee for Economic Development, which concludes that "There is no persuasive evidence that increased access threatens the sustainability of traditional subscription-supported journals, or their ability to fund rigorous peer review."

The Harvard response to the White House RFI on "Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research" states that "if publishers believe that short embargo periods would harm them, they should release data showing it. Researchers, research institutions, and taxpayers cannot be expected to prove the negative, or to prove the harmlessness of short embargoes. Until there is data to show harm, we must act in the public interest and provide early or immediate public access to publicly funded research. If publishers provide data showing substantive harm, then it may become appropriate to consider what kind of compromise with the public interest might be justified."

5. *How important is the embargo period to the economic health of journals and to the public interest as a whole? Is the importance of an embargo period likely to decline over time?*

It is important to answer this question based on evidence, and not brute assertions. Crucially, there is no evidence that short embargo periods cause any damage to the economic health of journals let alone to the public interest as a whole. (To the contrary, the public interest is of course advanced by eliminating embargo periods.) The burden of proof is on publishers to provide concrete, verifiable evidence of damage before compromising the public's interest in having access to the research results that they have funded. All existing evidence is to the contrary. Many subscription journals with zero-length embargoes are operating at a profit. Moreover, open-access journals, which have no embargo by their very definition, are being operated profitably at the highest levels of quality.

OSTP Report

6. *The OSTP report required by the COMPETES Act was recently released. Do you*

agree with the contents of the report and do you feel it addresses your concerns? Were any issues of interest to you not addressed? What specific next steps should OSTP or the Committee take in regards to the issue?

To my knowledge, the only report issued by OSTP pursuant to the COMPETES Act was the March 2012 report "Interagency Public Access Coordination: A Report to Congress on the Coordination of Policies Related to the Dissemination and Long-term Stewardship of the Results of Federally Funded Scientific Research". This report provides information about the process that the Interagency Public Access Committee (IPAC, established pursuant to Section 103(a) of the COMPETES Act) has been undertaking, but does not provide concrete conclusions, because "The NSTC groups are continuing to consider the public comments received from the RFIs and how they should be incorporated into the objectives required by ACRA. Once they have finalized their decisions, the objectives of all three groups will be combined and presented to the CoS." Until the IPAC submits its conclusions, any comment on their sufficiency would be premature.

However, given (as IPAC states) that "The Administration has long recognized the importance of improving the management of and access to the results of federally funded scientific research including digital data and peer-reviewed publications" and that the responses to the OSTP's RFIs "showed broad support for increasing public access to scientific publications" (page 1), the OSTP should not delay in stating its support for current legislation such as FRPAA, which is intended to achieve exactly this aim.

Impact of Access Policies

7. *Do existing public access policies have an adverse impact upon federal grant applications? Does the quality of proposals suffer because researchers do not want to make their work or data public? Understanding that researchers typically want to communicate their findings to the greatest possible audience, is this even a problem?*

I know of no empirical evidence showing that existing public access policies have had an adverse impact upon federal grant applications, and cannot imagine how such a result would arise. Public access policies provide *more* access to articles, access that can be used to improve not only the quality of scientific research, but also, the quality of

research proposals and the thoroughness and accuracy of proposal reviews.

Indeed, the opposite may well be true: Lack of open access could be harming federal grant applications. Proposal quality may suffer because many articles are effectively inaccessible to proposal developers, many of whom do not fall within the scope of subscription access to all of the cited and pertinent literature for a given proposal. Thus, a public access policy could have the effect of improving the quality of grant applications, since researchers developing proposals would have access to a broader range of the pertinent literature. Similarly, first principles suggest that the availability of research results should make review of proposals more systematic and complete, and therefore more accurate and effective.

The utility of open access to improve the writing and vetting of grant proposals is but one example of the advantages of open access in general in advancing not only science research but also public knowledge and the federal economy as well.

Data Access

8. *Should all information and data associated with federal research published by journals be made public? Is this reasonable or even possible?*

Ideally, all information and data associated with federal research published by journals would be made public. However, the incentives for open access to data are different from the incentives for open access to articles. Researchers have no incentive to keep articles secret, hence open access to articles is easier to achieve. Researchers may wish to exploit data that has been painstakingly collected for a period of time to draw further research conclusions. They may therefore perceive an incentive to keep such data secret. For this reason, the two efforts — to make articles and data accessible — should be dealt with separately. Crucially, there is no reason to delay policies to provide open access to articles while developing more speculative policies on open access to data.

With this caveat in mind, work on appropriate methods for opening up access to data to enable the replication of research conclusions and further analysis is an important goal to augment the progress of science, and should be rigorously pursued. Such methods will have to trade off the advantages of openness with the incentives for private exploitation that inhere in the issue of data accessibility.

9. *Aside from the publishing industry and scientific society concerns regarding open access, what are the science research implications of providing full open access?*

The only implications of providing full open access — absent the financial interests of a small set of scholarly publishers — are positive, not only for science research but for public knowledge and for the federal economy as well.

Open access benefits science research by making the fruits of research available to all who wish to build upon it, not just the few who happen to fall within the scope of a subscription to the particular journal in which an article is published. No researcher on earth has access to all of the pertinent scholarly journals — not even at my own institution, Harvard, which holds the largest academic library in the world. All researchers thus benefit from open access to scholarly articles.

Open access benefits public knowledge by placing scholarly articles in the hands of the taxpayers who underwrote that research, and in the hands of the press and other nongovernmental organizations who can interpret that research on behalf of the public.

Open access benefits the economy by supporting the access needs of businesses, small and large alike, who rely on the latest scientific results to generate new products, methods, and treatments. John Houghton's modeling of the effect of broader public access to federally funded research shows that the benefits to the US economy come to the billions of dollars and are eight times the costs (John Houghton, *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*, July 2010).

Even if it were true (contrary to the empirical evidence) that broad public access to the science literature could harm the financial interests of some publishers, the goal of the federal investment in scientific research is not to maintain the financial interests of that particular multinational industry, but to maximize the financial interests of the country as a whole. Thus, the federal government should promulgate policies that provide full open access.

Responses by Mr. Scott Plutchak

U.S. House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Investigations & Oversight

Questions for the Record

"Federally Funded Research: Examining Public Access and Scholarly Publication Interests"

Thursday, March 29, 2012
9:30 a.m. – 11:30 a.m.
2318 Rayburn House Office Building

Mr. Scott Plutchak,
Director, Lister Hill Library, University of Alabama at Birmingham

Questions submitted by Dr. Paul Broun, Chairman, Subcommittee on Investigations & Oversight

Pricing Issues

1. Concerns have been raised that the prices of many scholarly journals continue to increase sharply. Do you share these concerns? Since Congress usually avoids becoming involved in pricing debates, why should this be an issue for Congress?

I do share these concerns. Since the costs of materials in the health sciences continue to increase at annual rates of roughly 7% to 8% and state budgets, in recent years, have tended to fluctuate between 6% cuts and 2% increases, it becomes impossible for us to maintain existing subscriptions and to add access to new journals. Thus, the ready access to scientific research that my students and faculty require continues to diminish. While I understand the reluctance of congress to get involved in pricing issues, the economics of scholarly publishing are unlike most other sectors of the economy. Because each journal title is unique, it is not possible to substitute one neurology journal for another neurology journal. Librarians have had very little success in negotiating with publishers to lower prices in response to increasingly constrained acquisitions budgets. Because of the importance to the public welfare of a robust and innovative infrastructure which is very dependent on ready access to the peer-reviewed literature, I believe that this is an area where the Congress could be justified in considering legislation that addressed pricing issues.

2. What harms for your library and your students arise if you are not able to subscribe to all of the journals you would prefer to subscribe to? Is this just a matter of setting priorities of university budgets?

At UAB (University of Alabama at Birmingham), we have done extensive surveying, focus groups and interviews to examine the impact on our education and research missions of reduced

access to publications. The impact on undergraduate students is minor, since instructors can generally find substitute articles available if their first choice is unobtainable. It becomes more of a concern for graduate students and post-docs, particularly in the sciences. For example, a young post-doc in one of our focus groups said that often what she really just needs is to read the methods sections of five or six articles in the area of her research in order to help her refine here technique. If she has to make do with one or two, she risks using a method that is not as effective and that may result in a greater risk of her grant applications not being funded. Faculty find themselves in the position of trying to find an article on the same theme that may be "good enough," making do with the abstract of an article rather than reading the whole thing, or taking the time to request an article through interlibrary loan (virtually everyone that we talk to says that the 24 to 48 hours it takes to obtain an article through interlibrary loan is a significant impediment to their being able to work efficiently). Repeatedly we hear that when people can't get access to an article they'll contact the author or a colleague at another institution to get a copy. It is likely that in these cases, copyright or licensing terms are being violated.

The issue is certainly one of setting priorities for university budgets, although I wouldn't say it is "just" a matter, since there are always many competing priorities. For most health sciences libraries, materials costs increase 7% to 8% per year, and account for around 45% of the library's budget. In addition, new specialty journals arise on a regular basis (it was recently reported that global submission rates for academic journals are at their highest level in six years), which requires additional funding. With university budget increases at substantially lower levels than that, an institution that tried to prioritize library funding above all else would be on an endless slope, draining resources each year from those other priorities. Until pricing increases overall can be brought more in line with average increases in university budgets overall, it will not be possible for higher education institutions to prioritize library collections above all of the other essentials of a modern university.

3. If open access journals and repositories continue to grow and provide more competition to existing journals, does this eliminate any pricing concerns since subscribers will have more choices on where to spend their resources?

This is a complex question. Whether or not a proliferation of open access journals would mitigate pricing concerns depends on how they are funded. If funded through the "author-pays" mechanism, they would either reduce grant funding available or require institutional funding – in either case this could lead to pressures on library budgets as institutional budgets are reallocated. What would really make a difference is if the funding for OA journals, through whatever mechanism, turns out to increase annually at rates lower than the current increases in subscription journals. At this point, I don't think we have enough data to make accurate predictions. Most librarians that I know support OA because it will increase access to the literature, not because they think it will mitigate budget concerns.

Regarding repositories – the various institutional mandates (along with the NIH public access mandate) rely on a robust subscription journal market. If all of the relevant journals were OA to begin with, there'd be no need to deposit copies in repositories, but if the subscription journals go out of business, they won't any longer be facilitating peer review. So to the extent that repositories are useful, they won't mitigate pricing concerns. If they put journals out of business that will mitigate pricing concerns in the sense of there being fewer journals to buy, but then the repositories themselves will become useless unless other entities pick up the peer review facilitation responsibilities.

Peer review

4. Although federal taxpayers may have funded the research behind a submitted manuscript, they have not automatically paid the costs of editing and peer review unless a journal charges for that service. Would the debate over open access be less divisive if the federal government either:
    - Required that the submitted manuscript, not the peer reviewed version, be made available to the public immediately, or
    - Required that journals be compensated for all reasonably related peer reviewing costs?

There have been some suggestions that if government agencies required public deposit of the various progress and final reports that are required of grantees, this would address the taxpayer interest in having access to the results of federally funded research. Very few OA supporters have shown signs of favoring this approach. I suspect the reaction would be similar to a requirement to make the submitted manuscript available. Despite the strong divisions overall on the OA question, there does seem to be a broad consensus that peer review provides an essential service in vetting the reports of funded research.

On the other hand, a requirement that journals be compensated would certainly be favorably received by the publishers and I think that many OA advocates would accept this as well, particularly if it resulted in access to the final version of record immediately upon publication. My personal view is that this would be highly desirable. It is not without consequences, however, particularly if the funding were accomplished with existing agency grant budgets. For example, commonly accepted estimates for the cost of publishing all of the NIH funded research articles in OA are about 1% to 2% of the NIH extramural budget. While that seems small on a percentage basis, for an institution like UAB, 2% would be $4 to $5 million. This equates to something like 12 grants – a steady state loss of 3 or 4 new grants a year that wouldn't get funded. Given the importance of that first grant to a young researcher, it is likely that this would result in at least a couple of promising researchers a year leaving academia, or UAB needing to

come up with the funding to keep those individuals on board. So there would be some resistance to this from the research community.

Centralized database vs a distributed model

5. With open access archives spreading in number, please share your thoughts for a distributed model in which there are numerous databases versus a centralized model like PUBMED? The Internet itself is a distributed network by design. Shouldn't the same be true for how federally funded research is made available to the public, using search engines to find what is of interest? Would this impact your patrons in any way?

I favor a distributed model. Many people who have gotten involved in the public access debates don't realize that PubMed Central was created long before the NIH policy was developed as a way for NLM to meet its statutory responsibility to preserve the biomedical literature. Since the NIH policy relied on author's manuscripts, some sort of repository was needed, so PMC was convenient. Creating additional PMC-like repositories at other agencies would add substantially to the costs. On the other hand, providing linkages to existing articles on publisher websites adds very little infrastructure cost. In order for this to be as efficient as possible, however, publishers must be encouraged to adhere to developing standards for how documents are structured. The NLM document type, for example, is increasingly being used by other organizations and would serve as a useful basis for such standards.

The embargo period

6. Mr. Dylla's testimony stated that there should be different embargo models for different fields. How important is the embargo period to the economic health of journals and to the public interests as a whole? Is the importance of an embargo period likely to decline over time? What is the ideal embargo period, and is it the same duration for all disciplines?

The optimal embargo period is one that is set at such a point that most subscribers will be unwilling to wait for articles to become freely available and will thus maintain their subscriptions, but no longer than is necessary for that. That's the point in time where the public interest and the economic health of journals are in balance. For a cutting edge journal, published weekly, in a field where UAB has a strong research interest, the embargo period could be two weeks and I would probably still maintain a subscription. For an allied health journal in an area where we do not have a research interest and where the material is primarily of use to students, I'd probably cancel if there were a six-month embargo or perhaps even longer. For a quarterly journal in the social sciences where we don't have a strong research interest, I'd cancel if the embargo was a year.

The point is that there are quite a number of factors that play into whether or not a critical mass of institutions will cancel a particular journal and therefore there can be no such thing as an ideal embargo period. Supporters of FRPAA are fond of pointing to the apparent lack of evidence of mass cancellations caused by the NIH policy, but the journals that are highly affected by that are primarily biomedical research journals where the demand for the most current material is the highest. It is folly to conclude that extending the NIH policy to other agencies, with a shortened embargo, would have no negative impacts.

Divergent opinions

7. Your testimony noted your dislike with the two pieces of legislation out there. If Congress could start with a plain piece of paper, what federal policy or legislation would you support, or do you prefer no legislation at all?

If additional legislation is called for, in my view it needs to balance several elements, all of which were addressed in the Roundtable report. It should not stipulate a specific embargo period, relying on negotiations among the stakeholders in the different disciplines to address an appropriate time frame; it should encourage development of standards for text- & data-mining purposes, and recommend a framework for encouraging collaborative policy making among all stakeholders. It should require OSTP to play a coordinating role, but should also set reasonable deadlines for when effective public access policies would be developed. Such legislation would be less than optimal for the strong OA advocates, and more intrusive than desired by the publishers most concerned about preserving unregulated control of their business practices – so it would probably be striking the right balance.

OSTP Report

8. The OSTP report required by the COMPETES Act was recently released. Do you agree with the contents of the report and do you feel it addresses your concerns? Were any issues of interest to you not addressed? What specific next steps should OSTP or the Committee take in regards to the issue?

The OSTP report certainly represents a step in the right direction, but there is still more that needs to be done. In crafting the report, OSTP relied on the responses to the RFIs. That was fine as far as it goes, but the development of robust and effective policies at the agency levels that will balance all of the appropriate concerns will require substantial additional engagement among all of the stakeholders. I recognize that there are hurdles involved in setting up appropriate working groups that can effectively engage in this kind of policy development, but I believe it is necessary to get past some of the impasse that we seem to still be left with. OSTP should be encouraged to develop a framework for such public/private cooperation in policy development at the agency level, with specific timeframes involved for delivering policy recommendations.

Impact of Access Policies

9. Do existing public access policies have an adverse impact upon federal grant applications? Does the quality of proposals suffer because researchers do not want to make their work or data public? Understanding that researchers typically want to communicate their findings to the greatest possible audience, is this even a problem?

I do not believe that existing public access policies have an adverse impact on grant applications. Researchers *do* want to make their work public and in many cases they want to make their data public, once they've finished the work that the data is used for. Wider public access to peer reviewed journal articles (or versions thereof) should not present any problems for researchers appropriately protecting their data or their work before they are ready to make it public.

Data Access

10. Should all information and data associated with federal research published by journals be made public? Is this reasonable or even possible?

This is a much more complicated problem than it appears at first glance. "All information and data" is potentially a huge amount, much of which is useless without substantial context and much of which, depending on the discipline (in computer science, for example) would be more useful if it were re-generated when needed rather than stored and made public. A more useful way of framing the desired outcome might be that all information and data necessary for someone to replicate and validate the research published by journals should be made public. I do think that would be in the public interest. But defining exactly what that data is and how it should be stored and managed, and what are the complementary roles of government, research institutions and publishers is still not well understood, although there is much good work being done in this area. It is certainly one in which the federal government has a strong interest and an important role to play.

11. Aside from the publishing industry and scientific society concerns regarding open access, what are the science research implications of providing full open access?

Full open access, including the ability to text- and data-mine across standardized databases of journal articles has the potential to greatly increase the rate of discovery beyond what has been achieved with existing tools. In this sense, the passion of open access advocates is well-founded. The challenge is how to get there from here, without substantial negative consequences. We are in a time of great upheaval and it is clear that sustaining the current scholarly publishing system should not be a goal of the government or of that industry itself – and most publishers recognize that. People in publishing are working very hard to take advantage of the new digital technologies to make more information more widely and effectively available than ever before. Open access is a part of that. But there are a lot of elements to this change and government

policy makers need to find that balance between reasonable caution and strong incentives to change in order to help the entire scholarly communication system deliver to the public what it has the potential to do.

Appendix II

Additional Material for the Record

LETTER SUBMITTED BY REPRESENTATIVE ZOE LOFGREN

An Open Letter to the U.S. Congress Signed by 52 Nobel Prize Winners

March 28, 2012

The Honorable Zoe Lofgren
United States House of Representative
1401 Longworth House Office Building
Washington, DC 20515

Dear Representative Lofgren:

As scientists and Nobel Laureates, we write to express our strong support for H.R. 4004 and S. 2096, the Federal Research Public Access Act (FRPAA). This broadly supported bi-partisan legislation would enhance access to federally funded, published research articles for scientists, physicians, health care workers, libraries, students, researchers, academic institutions, companies, and patients and consumers.

Broad dissemination of research results is fundamental to the advancement of knowledge. For America to obtain an optimal return on our investment in science, publicly funded research must be shared as broadly as possible. Yet, too often, research results are not available to researchers, scientists, or members of the public. We believe Congress can and must act to ensure that all potential users have free and timely access on the Internet to peer-reviewed federal research findings. This ultimately magnifies the public benefits of research by promoting progress, enhancing economic growth, and improving the public welfare.

As the pursuit of science is increasingly conducted in a digital world, we need policies that ensure that the opportunities the Internet presents for new research tools and techniques to be employed can be fully exploited. The removal of access barriers and the enabling of expanded use of research findings has the potential to dramatically transform how we approach issues of vital importance to the public, such as biomedicine, climate change, and energy research. As scientists, and as taxpayers too, we support FRPAA and urge its passage.

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 more effectively compete 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.

FRPAA builds on established public access policies that have been adopted by government agencies in both the U.S. and abroad. The National Institutes of Health (NIH) have implemented a successful comprehensive public access policy, mandated through the Consolidated Appropriations Act of 2008. All seven of the Research Councils in the United Kingdom have public access policies as does the Canadian Institutes of Health Research. This bill is also consistent with the growing number of institutional open-access policies that have been adopted at universities such as Harvard, Massachusetts Institute of Technology, and the University of Kansas.

The federal government funds over $60 billion in research annually. Research supported by the NIH, which accounts for approximately one-third of federally funded research, produces a more than 90,000 peer-reviewed journal articles each year. The return on our investment in scientific research is best realized with policies that promote access to the published results of that research. Passage of FRPAA will make it easier for scientists worldwide to better and more swiftly address the complex scientific challenges that we face today and expand shared knowledge across disciplines to accelerate breakthrough and spur innovation. As the undersigned Nobel Laureates, representing both U.S. interests and those of the rest of the scientific world, we ask you to co-sponsor and support the Federal Research Public Access Act.

Signed by 52 Nobel Laureates

U.S. Laureates:

Name	Category	Prize Year
Alexi Abrikosov	Physics	2003
Peter Agre	Chemistry	2003
Paul Berg	Chemistry	1980
J. Michael Bishop	Medicine	1989
Linda Buck	Medicine	2004
Martin Chalfie	Chemistry	2008
Elias Corey	Chemistry	1990
Robert F. Curl Jr.	Chemistry	1996
Johann Deisenhofer	Chemistry	1988
Andrew Z. Fire	Medicine	2006

Edmond H. Fischer	Medicine	1992
Riccardo Giaccardo	Physics	2002
Sheldon Glashow	Physics	1979
Roy Glauber	Physics	2005
Paul Greengard	Medicine	2000
Roger Guillemin	Medicine	1977
David Gross	Physics	2004
John L. Hall	Physics	2005
Leland H. Hartwell	Medicine	2001
Dudley Herschbach	Chemistry	1986
Roald Hoffmann	Chemistry	1981
Louis Ignarro	Medicine	1998
Wolfgang Ketterle	Physics	2001
Roger D. Kornberg	Chemistry	2006
Mario Molina	Chemistry	1995
Kary B. Mullis	Chemistry	1993
Arno Penzias	Physics	1978
H. David Politzer	Physics	2004
Stanley Prusiner	Medicine	1997
Robert C. Richardson	Physics	1996
Richard J. Roberts	Medicine	2012
Daniel Shechtman	Chemistry	2011
George Smith	Physics	2009
Hamilton Smith	Medicine	1978
George Smoot	Physics	2006
Jack W. Szostak	Medicine	2009
Joe Taylor	Physics	1993
James Watson	Medicine	1962

Non-U.S. Laureates

Name	Category	Prize Year

March 28, 2012
4 of 4

Aaron Ciechanover	Chemistry	2004
Richard Ernst	Chemistry	1991
Avram Hershko	Chemistry	2004
Tim Hunt	Medicine	2001
Klaus von Klitzing	Physics	1985
Yuan T. Lee	Chemistry	1986
Jean-Marie Lehn	Chemistry	1987
Sir Peter Mansfield	Medicine	2003
Hartmut Michel	Chemistry	1988
Erwin Neher	Medicine	1991
John Polanyi	Chemistry	1986
Venkatraman Ramakrishnan	Chemistry	2009
Brian Schmidt	Physics	2011
Akira Suzuki	Chemistry	2010
Sir John Walker	Chemistry	1997

RESPONSE TO FEDERAL REGISTER NOTICE REQUEST FOR INFORMATION SUBMITTED BY DR. H. FREDERICK DYLLA

AIP American Institute of Physics

One Physics Ellipse, College Park, MD 20740-3843 H. Frederick Dylla, Executive Director and CEO

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 Commons[5] 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. 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 CrossMark[9] 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, Portico[11] 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	*Applied Physics Letters, Journal of Applied Physics, Journal of Low Temperature Physics, Review of Scientific Instruments*
Chemical Physics	7.1	>10	*The Journal of Chemical Physics*
Physics—Fluids & Plasmas	6.6	7.6	*Physics of Fluids, Physics of Plasmas*
Mathematical Physics	6.5	>10	*Journal of Mathematical Physics*
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.

AIP American Institute of Physics
One Physics Ellipse, College Park, MD 20740-3843 H. Frederick Dylla, Executive Director and CEO

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 Digital Data Resulting From Federally Funded Scientific Research; FR Doc. 2011–28621**

Submitted by: H. Frederick Dylla, Executive Director and CEO, American Institute of Physics
Tel. +1 301-209-3131; Dylla@aip.org

Electronically submitted to: digitaldata@ostp.gov

The American Institute of Physics (AIP) appreciates this opportunity to submit comments and would be delighted to continue working with OSTP and other federal partners through a process of active engagement.

About AIP
The American Institute of Physics (AIP) 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. As one of the largest publishers of scientific information in physics, AIP employs innovative publishing technologies and offers publishing services for its Member Societies. AIP's suite of publications includes 15 journals, three of which are published in partnership with other organizations; magazines, including its flagship publication *Physics Today*; and the AIP Conference Proceedings series. 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.

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.

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. As a scholarly publisher, AIP believes that better discoverability and reuse of original research data are to be encouraged at all levels and among all stakeholders. AIP also believes that data resulting directly from federally funded scientific

research should be made freely available in a sustainable manner and that this is best achieved through appropriate policies that leverage public-private collaboration.

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. Policies should recognize that hosting, maintaining and preserving raw data or data sets, and continuing to make such data available over the long term, has a cost, which, in certain circumstances, the host should be entitled to recover. 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 specifically recommends that federal grants set aside funds to support researcher data management and deposit efforts. Federal agencies could also play a role in supporting and encouraging the establishment of discipline-specific data archives where these are currently lacking. The amount and type of support should be determined in collaboration with key stakeholders involved in the deposit, storage, and preservation of data.

Federal policies should also focus on supporting and encouraging the development of community standards for the citation and reuse of data sets, thereby facilitating the creation of a system that gives researchers an incentive to share data resulting from federal grants.

AIP Responses to RFI Questions

Preservation, Discoverability, and Access

(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?

We would make the distinction that it is not "public access" in the broadest sense that is important but rather access by other scientists who can use the digital data for the further advancement of science.

As data are not copyrightable, policies about access become policies about deposit by the data owner or proxy into an accessible system. It should be noted, though, that any policies should recognize and take into account differences between 'databases' (information products created for the specific display and retrieval of data) and 'data sets' (sets or collections of raw relevant data captured in the course of research or other efforts). Policies could require that data generated from federally-funded research be deposited in a certified and openly accessible repository; furthermore, researchers could be encouraged to make these deposits upon submission of their first manuscript showing results that were based on the data set. Although some agencies already have a preservation/access role (for example, DOE Order 241.1B), AIP agrees with the Interagency Working Group on Digital Data that "data stewardship is best accomplished in a system that includes distributed collections and repositories maintained where the custodian has trusted community-proxy status with the relevant communities of practice." Agency policies should support and encourage such a distributed system for both access and preservation; that is, policies should recognize and build upon the broad set of capabilities that exist for both access and preservation within the library and publishing communities for both documents and data, such as Portico, LOCKSS.

The integrity of preserved data would also need to be taken into account and supported by any policy.

(2) What specific steps can be taken to protect the intellectual property interests of publishers, scientists, Federal agencies, and other stakeholders, with respect to any existing or proposed policies for encouraging public access to and preservation of digital data resulting from federally funded scientific research?

All policies should comply with current copyright and patent law. Data should be embargoed to the principal researcher until conclusions drawn from the data can be published in the research literature. An additional maximum embargo of one year would also provide for the filing of patents by the grantees (or their institution) as allowed by many, if not all, funding agencies (HR 1249 Sec 102(b)(1)(A)). See also the distinction between databases and datasets as addressed response to question 1.

(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?

Differences between scientific disciplines and types of digital data must be taken into account by domain experts at the time of proposal review (note the language used in the Data Management Plan FAQ's of NSF in a variety of instances: "to be determined by the community of interest through the process of peer review and program management.") Only such experts will be able to determine if the data to be generated by the proposed research will be of longer term value to the scientific community of interest and if its type conforms to acceptable community standards.

Metadata—data about the data—which would include information both about what the data is and how it was collected, is addressed further in this response.

(4) How could agency policies consider differences in the relative costs and benefits of long-term stewardship and dissemination of different types of data resulting from federally funded research?

Policies must first recognize that not all data is worth preserving. Each type of data should be assessed regarding long-term stewardship. Policies would have to take into account not just the size of the datasets but also long-term usability, which depends on the rate of technology change, and level of documentation required. Along with the data, enough information needs to be preserved to reproduce the dataset. As noted in the answer to question 3, agencies will need to call upon data experts as well as scientific experts.

(5) How can stakeholders (e.g., research communities, universities, research institutions, libraries, scientific publishers) best contribute to the implementation of data management plans?

There needs to be an interconnected system for access to and sharing and preservation of data based on community-developed standards and best practices. The system needs to encourage innovation and must support multiple solutions—data as an information resource is inherently more complicated than scholarly articles. Each stakeholder will then need to contribute based on their specific skills and expertise. Libraries, through Institutional Repositories, could take on a stronger preservation role. Publishers have been adding value to the research process and providing access to and preservation of the scholarly literature for hundreds of years and could extend this to data, well beyond current support for supplemental material. Universities and research institutions have both scientific domain knowledge and data and information experts. Any system will need to preserve incentives for innovation.

Consider, for example, work being done by the Data Preservation Alliance for Social Sciences through their partnership with the Library of Congress, LOCKSS, and Dataverse to prototype a policy-based replicated data archive.

Other examples include:

- linking between datasets and their resulting scholarly publications based on community-accepted standards, thus ensuring datasets become part of the scientific literature;
- Having clear standards and guidelines for the certification and auditing of data repositories; encouraging a system that incentivizes data repositories to maintain the accuracy or integrity of the data once it has been deposited;
- Incentivizing the deposit of datasets and ensuring that the administrative burden this imposes on researchers minimal.

(6) How could funding mechanisms be improved to better address the real costs of preserving and making digital data accessible?

Require data management plans and coordinate plan requirements across agencies and to community standards (see the Open Archive Information System Reference Model – ISO standard 14721:2003). What constitutes data that needs to be preserved should be clearly identified through the process of peer review and program management. Preserving and disseminating digital data should then be considered "part of the cost" of funding and doing research, not "an additional cost". Funding agencies could emphasize that proposals must take into account data fit for reuse and preservation. Again, this

should be the approach across agencies. Research labs/institutions/university overhead rates would need to include costs of data preservation.

As pointed out in the final report from the Blue Ribbon Task Force on Sustainable Digital Preservation and Access (*Sustainable Economics for a Digital Planet*): "Policy mechanisms can play an important role in strengthening weak motivations" as there is often "misalignment of incentives between communities that benefit from preservation (and therefore have an incentive to preserve), and those that are in a position to preserve (because they own or control it) but lack incentives to do so."

(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?

If data is created in the course of federally-funded research, then the funding agency could require that any such data deemed to be "preservation data" be deposited in a recognized archive. Through direct agency involvement in creating a "comprehensive framework for data access and preservation" based on community-accepted standards and best practices for data citation and reuse, agencies would maintain lists of certified repositories. Certified repositories could be similar to the data center members of the DataCite organization (of which DOE's Office of Scientific and Technical Information is a member) or participants in the SafeArchive program of Data-PASS. In addition, grantee data management plans could be required to identify all datasets expected to be produced from funded work.

Certification of compliance would then simply require grantee reporting to include in reports on their funded proposal the data citations and the repository where the data was deposited.

As work is already being carried out to develop standards in this area (i.e. *The ISO 16363 Standard for Trusted Digital Repositories),* it would be more expedient for federal agencies to work within and help support such standards.

(8) What additional steps could agencies take to stimulate innovative use of publicly accessible research data in new and existing markets and industries to create jobs and grow the economy?

AIP agrees with the statement from the Interagency Working Group on Digital Data (IWGDD) in its report, *Harnessing the Power of Digital Data for Science and Society,* that "the current landscape lacks a comprehensive framework for reliable digital [data] preservation, access, and interoperability". We feel that there is a very important role for the federal government and its science funding agencies to play to help create and promulgate such a comprehensive framework.

Federal investment in creating stable, standardized, and accessible data will be an essential base from which innovation can occur. The ease of reuse could then lead to developments akin to IBM Research's "Many Eyes" product for data visualization (www-958.ibm.com), or spur the private sector to offer data services for researchers.

(9) What mechanisms could be developed to assure that those who produced the data are given appropriate attribution and credit when secondary results are reported?

This ecosystem of attribution and credit already exists with respect to scholarly articles. A researcher's standing in their field is largely a result of their list of authored scholarly articles and the number of citations to those published articles. The credit comes in the form of respect from peers, funding for further work, and career advancement, and rests in large part on the underlying quality control provided by peer review. Not providing appropriate attribution is considered unethical scientific behavior and can lead to the retraction of published work.

The mechanisms to be developed would support an extension of this system to cover data. The elements to support are:

- data must be recognized as a primary research output,
- data must have unique and persistent identifiers and be fully citable, thereby allowing its use and reuse to be tracked and recorded in the same way as scholarly publications, and
- data citation information must be used for research evaluation and reward.

Persistent identifiers for data could be handled through use of digital object identifiers already used for scholarly articles or similar (see Datacite.org). There are also examples of recommended practice for citing data. [For example: creator (publication year): Title, Publisher, identifier; see http://datacite.org/whycitedata and DOE's Data ID Service.]

Publishers could support the development of such a system by requiring that all data needed to reproduce the results and conclusions of a published scholarly article must be cited according to community standards.

Funding agencies could support the development of such a system by recognizing data that has been archived and made available to the research community as "first class research objects" at the same level as articles. Agencies should also recognize any reuse of these data which could then be counted via citations.

See the Australian National Data Center's "Building a Culture of Data Citation" poster available at http://ands.org.au/cite-data/index.html.

For a hybrid example spanning the world of digital data and scholarly publication, see the *Journal of Physical and Chemical Reference Data*, a long and successful collaboration between AIP and the National Institute of Standards and Technology.

Standards for Interoperability, Reuse and Re-Purposing

(10) What digital data standards would enable interoperability, reuse, and repurposing of digital scientific data? For example, MIAME (minimum information about a microarray experiment; see Brazma et al., 2001, Nature Genetics 29, 371) is an example of a community-driven data standards effort.

First, it is important to separate metadata standards from data format standards. Metadata standards could be developed that are lightweight enough to be widely interoperable and extensible so as to accommodate discipline-specific needs (within the XML publishing standard). These standards would need to cover both bibliographic information (data creator, date of creation, what the data describes, where it can be accessed, etc.), and how it was collected (experimental apparatus, experimental conditions, location, etc.).

Data format standards that would enable reuse and repurposing would need to be developed at the discipline-specific level. There need not be one solution per discipline: it may be that the communities in question need a handful of solutions that correspond to the various types of data and/or modes of scientific research that produces the data. So while it is true that actual data solutions need to be discipline appropriate, there may be logical clusters of solutions for the connections between publishing and data depending on the nature of the data.

There is a role for federal agencies in coordinating across discipline boundaries (covering all funded areas) and internationally. In its October 2011 report, *Federal Engagement in Standards Activities to Address National Priorities: Background and Proposed Policy Recommendations,* the Subcommittee on Standards of the National Science and Technology Council noted that "There was agreement among respondents that the US government should continue to play the role of participant in private sector standards setting processes. There was also general agreement that the effectiveness of government participation depends on the level and consistency of involvement and commitment of resources, both staff and budgetary, to the process. Lack of coordination among agencies...was cited by many respondents as having a negative impact on government effectiveness."

(11) What are other examples of standards development processes that were successful in producing effective standards and what characteristics of the process made these efforts successful?

The Digital Object Identifier, or DOI, is an example of a successful standard. Its development and adoption involved a multi-stakeholder, community-driven approach that solved a practical problem and provided benefit to the end-user.

(12) How could Federal agencies promote effective coordination on digital data standards with other nations and international communities?

AIP supports the recommendation of the Interagency Working Group on Digital Data (IWGDD) that an NSTC Subcommittee for digital data preservation, access, and interoperability be created. This subcommittee would then be able to provide coordination among the US funding agencies and collaborate with its international counterparts. Coordination at the national level should extend beyond

science funding agencies as relevant work is being done elsewhere within the US government (for example, the work of the Library of Congress through its National Digital Information and Infrastructure Program [NDIIP], particularly its "partnership with the National Science Foundation in 2005 to undertake a program of pioneering research to support advanced research into the long-term management of digital information").

In addition, this subcommittee could ensure that each Federal agency is itself required to adopt and implement digital data standards developed within the global community.

Federal agencies can support conferences and other initiatives on a discipline level by funding standards and preservation work as well as pure research.

(13) What policies, practices, and standards are needed to support linking between publications and associated data?

See answer to question 9. The mechanism for linking between publications and associated data essentially exists with the digital object identifier, which is already used widely for linking between publications. The federal government could provide additional logistics and financial support for making this mechanism standard practice with respect to data and coordinating/aligning policies across federal agencies to encourage use of those standards by grantees.

Agency involvement and/or support of current initiatives such as the NISO/NFAIS Working Group on Supplementary Journal Information (www.niso.org), which is working on recommended practices for publishers who are increasingly attaching data sets as supplementary information appended to publications, would also help address some of the issues at a practical level.

RESPONSE TO FEDERAL REGISTER NOTICE REQUEST FOR INFORMATION SUBMITTED BY DR. CRISPIN TAYLOR

American Society of Plant Biologists

12 January 2012

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

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

Electronically submitted to: publicaccess@ostp.gov

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

About ASPB

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

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

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

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

Introduction

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

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

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

ASPB Responses to RFI Questions

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

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

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

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

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

As stated in the 2010 *Scholarly Publishing Roundtable* report[3], many publishers have made the decision to move toward increasingly open structures and archives[4] as enabled by Open Access business models and new solutions to associated permissions, such as Creative Commons[5] licenses. These licenses provide a means for exercising certain rights regarding the re-use of an item. For example, these licenses could provide reuse rights if the resulting new works are also made available to the public. The *Roundtable Report* also notes that the number of journals making a change in business model is appreciable but small within the universe of more than 25,000 scholarly peer-reviewed journals[6]. ASPB echoes the *Roundtable Report* assertion that no existing digital business model has demonstrated its viability to the satisfaction of all, and we caution against de facto government endorsement of any single approach.

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

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

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

To maximize the effectiveness of its efforts, government does have an important role to play in convening stakeholders to develop standards for data and metadata, thereby helping to make research more readily searchable and discoverable. Publishers are already working in partnership to develop standardized information and collections through initiatives such as CrossRef[7].

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 have already adopted this approach (e.g., Howard Hughes Medical Institute, Wellcome Trust, and Max-Planck Institutes).

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

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

Government mandates for public access come at a significant cost to the US economy and to the scientific enterprise. Data from the National Institutes of Health's (NIH's) PubMed Central (PMC) repository indicate that two-thirds of PMC's users are from overseas. This suggests that critical export opportunities for the industry may be compromised, potentially resulting in the loss of US jobs[9]. Significant economic value added by the publishing industry could be wasted if revenue

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

What constitutes a publication and the nature of publication is changing with technology. A publication is no longer just a chunk of text fixed in time forever but a fluid representation. Publications can include supplemental material, multimedia files, software, and links to resources on the web and can be revised and corrected over time by the authors and publishers, hence the emergence of new community initiatives such as CrossRef's CrossMark[10] service, which electronically watermarks an article's Version of Record (VoR), and DataCite[11], which extends the CrossRef-promoted Digital Object Identifier (DOI) to datasets. Any plan for the future should recognize that the static aggregation/library model is not likely to hold up well in the distributed and dynamic Internet milieu.

ASPB believes that it is unlikely that one optimal procedure for preservation and stewardship will emerge to become applicable across all of scholarly publishing. For now, ASPB strongly recommends that agency policies embrace diversity, decentralization, and interoperability. In the long term, systematic collaborations among stakeholders (government, publishers, universities and their libraries, and other not-for-profit participants in the scholarly publishing system) will be necessary to achieve maximum benefit. We note that libraries, in partnership with publishers, have established entities for preservation of digital documents that are already in wide use, for example, Portico[12] and CLOCKSS[13].

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

(4) Are there models or new ideas for public-private partnerships that take advantage of existing publisher archives and encourage innovation in accessibility and interoperability, while ensuring long-term stewardship of the results of federally funded research?
Yes, please see detailed response to Question 5 below.

(5) What steps can be taken by federal agencies, publishers, and/or scholarly and professional societies to encourage interoperable search, discovery, and analysis capacity across disciplines and archives? What are the minimum core metadata for scholarly publications that must be made available to the public to allow such capabilities? How should federal agencies make certain that

such minimum core metadata associated with peer-reviewed publications resulting from federally funded scientific research are publicly available to ensure that these publications can be easily found and linked to federal science funding?

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

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

Specifics on these items are discussed below.

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

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

Standards and Identifiers: Agency Funding Information

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

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

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

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

Standards and Identifiers: Promoting Interoperability

ASPB is seeking to collaborate with operators of a prominent knowledge base in plant biology that incorporates a rich array of genomic information from a wide variety of plant species to establish mechanisms for algorithmically connecting journal articles to database entries upon publication. Specifically, the collaborators propose to enable the retrieval of functional gene annotations and molecular annotations from ASPB journal articles using data-mining tools such as Textpresso[14] and BioCreative[15], both of which make use of Natural Language Processing and are organized around robust and highly structured ontologies. The collaborators plan to create a reference library that includes known and predicted gene names, symbols, functions, phenotypes, and pathway annotations in three target plant species. Together with the ontologies, which will play a key role in structuring data annotation, the library will also help establish data capture architectures that the ASPB journals would implement with their authors as manuscripts are being submitted, thereby directly, immediately, and algorithmically connecting published journal articles with the underlying datasets and knowledgebase. Both collaborators envision developing proof-of-concept data-mining methodologies that would be broadly applicable in other fields of research.

Standards and Identifiers: DOIs for Data Sets and Supplementary Material

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

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

Standards and Identifiers: Author Name Disambiguation

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

Discovery Tools: Content Mining

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

Such pilots should focus on goals such as the following:

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

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

Pilot Projects: Sponsored Access to Published Research

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

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

Pilot Projects: Linking to/from Research Reports

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

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

Possible outcomes of such a pilot might include:

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

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

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

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

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

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

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

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

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

References

[1] Cox, J. and L. Cox, *Scholarly Publishing Practice: Academic Journals Publisher's Policies and Practices in Online Publishing*, 3rd ed., ALPSP (2008), http://www.alpsp.org/Ebusiness/ProductCatalog/Product.aspx?ID=44.
[2] Outsell, "An Open Access Primer-Market Size and Trends" (2009), http://www.outsellinc.com/contact_us/open_access_primer_2009.
[3] Report and Recommendations of the Scholarly Publishing Roundtable, January 2010, available at www.aau.edu/WorkArea/showcontent.aspx?id=10044. Referred to throughout this document as the *Roundtable Report*.
[4] Morris, S., *Journal Authors' Rights: Perception and Reality* (London: Publishing Research Consortium, 2009), http://www.publishingresearch.net/documents/JournalAuthorsRights.pdf.
[5] Creative Commons (http://creativecommons.org/about) is a nonprofit corporation that provides free licenses and other legal tools to mark creative work with the freedom the creator wants it to carry, so others can share, remix, use commercially, or any combination thereof.
[6] Ware, Mark and Michael Mabe, *The STM Report: An Overview of Scientific and Scholarly Journals Publishing*. September 2009.
[7] CrossRef (www.crossref.org) is a not-for-profit group founded by publishers in 2002 that now maintains 50 million items. Almost 1,000 publishers participate, assigning Digital Object Identifiers (DOIs) to published content items. Development of the CrossRef service has resulted in seamless navigation of the research literature by users so that researchers using the bibliography in one article can link from a reference to the full text of the referenced article.
[8] This would ensure readability to the broadest audience. NSF is already pursuing such a policy, see http://www.nsf.gov/pubs/policydocs/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/.
[9] See, e.g., 2009 U.S.-China Joint Commission on Commerce and Trade (JCCT) Factsheet. Available at http://www.ustr.gov/about-us/press-office/fact-sheets/2009/october/us-china-joint-commission-commerce-and-trade.
[10] CrossMark (www.crossmark.com) is a current pilot project of CrossRef to that will allow readers to easily determine whether they are looking at the publisher-maintained, stewarded version of a journal article.
[11] DataCite (http://datacite.org) is a not-for-profit organization established to facilitate easier access to research data on the Internet, increase acceptance of research data as legitimate, citable contributions to the scholarly record, and support data archiving that will permit results to be verified and re-purposed for future study.
[12] Portico (http://www.portico.org/digital-preservation/) is a digital preservation service provided by a not-for-profit organization with a mission to help the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways. It is among the largest community-supported digital archives in the world, working with libraries, publishers, and funders to preserve e-journals, e-books, and other electronic scholarly content.
[13] CLOCKSS (*Controlled* LOCKSS) is a not-for-profit joint venture between the world's leading scholarly publishers and research libraries whose mission is to build a sustainable, geographically distributed dark archive with which to ensure the long-term survival of web-based scholarly publications for the benefit of the greater global research community (http://www.clockss.org/clockss/Home).
[14] http://www.textpresso.org/
[15] http://biocreative.sourceforge.net/

The Association of Learned and Professional Society Publishers
Shaping the Future of Learned and Professional Publishing

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

To:
Office of Science and Technology Policy (OSTP)
publicaccess@ostp.gov

From:
Dr. Audrey McCulloch
Acting Chief Executive
The Association of Learned and Professional Society Publishers (ALPSP)
15 Windmill Way
Tring
Herts HP23 4HQ
United Kingdom
audrey.mcculloch@alpsp.org

January 2012

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

1. The Association of Learned and Professional Society Publishers (ALPSP) is the international trade association representing scholarly and professional publishers across all academic disciplines. ALPSP has a broad and diverse membership of over 300 organizations in 37 countries who publish over half the world's total active journals, as well as books, databases and other products.

2. ALPSP's mission is to connect, train and inform the scholarly and professional publishing community and to play an active part in shaping the future of academic and scholarly communication.

3. In the US, ALPSP represents 60 organizations in 14 states employing an estimated 3,000 employees.

4. ALPSP welcomes the opportunity to respond to the Office of Science and Technology Policy (OSTP) Request for Information on Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research. Our response addresses issues relevant to the ALPSP membership.

5. Scholarly publishing is an international enterprise, with around 1.5 million articles published annually[1]. US researchers dominate this output with a 29% share of the total. The majority of publishers (95%) are small, publishing one or two journals. At the other end of the scale, the 100 largest publishers account for 67% of the total number of journals.

6. Publishers are dedicated to providing the widest dissemination of the peer-reviewed results of research and to supporting the scientific enterprise. In addition to investing heavily in staff and technology, not-for-profit learned and professional society publishers redirect their 'surplus' back into the community through organization of conferences, scholarly awards, teaching fellowships, skills transfer through workshops and seminars, enhancing professional standards and benchmarking, travel and other grants, provision of patient information and public understanding of science initiatives. Commercial publishers also invest directly in the scientific community, through grants, awards and other sponsorship schemes.

7. Publishers support any *sustainable* models of access, the most common being the subscription-based model. Gold Open Access, where the author (via the institution or funder) provides payment to fund publication, is gaining popularity, though it should be noted that this is not a fully tested model with regard to long-term sustainability. Publishers are working with funding organizations to investigate the issues surrounding this new access model to ensure it can provide sustainable business models for publishers to continue to disseminate value-added peer-reviewed literature.

8. Policies which require open access publication but do not provide funding for that publication, such as Green Open Access (author self-archiving in openly accessible

[1] http://www.stm-assoc.org/industry-statistics/the-stm-report/

repositories) threatens to undermine the publication system on which it depends, as evidenced in a recent report from the Research Information Network[2].

9. The PEER project[3] in Europe has been investigating the effects of large-scale, systematic deposit of the Accepted Manuscript (see NISO/ALPSP definitions for Journal Article Versions[4]) in repositories. This project is a rational approach towards defining the problems and thereby identifying potential solutions. It is a broad ranging project encompassing economic, behavioral and usage aspects. The behavioral study has reported and noted that authors value highly peer-reviewed journals and whilst there is still some confusion regarding open access publishing, there were reservations about peer-reviewed papers being held in open-access repositories. It also found that readers were unlikely to go to a repository to search for journal articles.

(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?

10. Current markets for peer-reviewed publications exist globally and publishers have invested heavily to ensure that there are many channels of access to publications. The markets are already well-served and a recent survey from the Publishing Research Consortium found that 97% of researchers in North America have very or fairly easy access to research journals[5]. This study also demonstrated that North America enjoys one of the best 'access to information' versus 'importance of that information' profiles of any of the regions investigated.

11. Publishers have recognized the needs of the myriad communities they serve and have responded appropriately, leading the way with technical tools and services to enhance the access, usability and analysis of published research, collaborating widely with various stakeholders in the process.

12. In this regard, a number of publisher-led initiatives have increased access to many different user groups. For example, DeepDyve[6], an article rental system, enables anyone to access thousands of scholarly and academic journals. Users may browse an article online and subsequently purchase the article for download if desired. patientINFORM[7] brings up-to-date, authoritative information from the world's leading medical journals to patients and caregivers. Information is provided in a summarized form, with links to free or reduced-price access to the full article on the publisher website. The Emergency Access Initiative[8] is a partnership between the Association of American Publishers (plus other publishers), the National Library of Medicine and the

---

[2] http://www.rin.ac.uk/news/press/heading-open-road-costs-and-benefits-transitions-scholarly-communications
[3] http://www.peerproject.eu/
[4] http://www.niso.org/publications/rp/ NISO RP-8-2008 Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group
[5] http://www.publishingresearch.net/projects.htm Access vs. Importance
[6] http://www.deepdyve.com/
[7] http://www.patientinform.org/
[8] http://eai.nlm.nih.gov/docs/captcha/test.pl?url=

National Network of Libraries of Medicine with the aim of providing temporary and free access to those affected by disasters and those providing assistance to them. It includes public access.

13. In addition to the collaborations in paragraph 12, publishers also provide free or very low cost access to universities and colleges, research institutes, schools, hospitals, governmental offices and national libraries in the lowest gross national income per capita countries throughout the world through initiatives such as Research4Life[9], eIFL[10] and PERii[11].

14. It is clear that publishers are keen to ensure that the needs of different markets in accessing scholarly information are met appropriately and are keen to do so in collaboration with other stakeholders. Publishers are keen to engage with the US Government to address the further gaps it has identified in public access. It would be useful for agencies to detail the particular needs of such user groups and to collaborate with publishers to establish the most efficient and appropriate ways in which to address those needs.

15. The need for archiving digital information has been recognized by publishers, librarians, funders and researchers. Collaborative projects already exist to ensure the long term preservation of scholarly information through initiatives such as Portico[12], LOCKSS[13], CLOCKSS[14] and the National Library of the Netherlands (Koninklijke Bibliotheek) eDepot[15].

16. Very careful consideration needs to be given to archiving and public access policies, if these are to be tied to growth in the US economy and improving output of the US scientific enterprise. Public access cannot be restricted to one local region. Ensuring public access to publications resulting from federally-funded research will result in global access, therefore benefiting researchers and other users all over the world (and potentially also their economies), not just the US. This removes any competitive advantage for the US economy and research output.

17. Data from the National Institute of Health reports that more than half of all PubMed Central users are from outside the US. This repository is therefore reducing the export market for the US publishing industry which, in total, employs around 50,000 people and contributes c. US$3.5 billion to the US balance of trade.

(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?

---

[9] http://www.research4life.org/
[10] http://www.eifl.net/
[11] http://www.inasp.info/
[12] http://www.portico.org/digital-preservation/
[13] http://www.lockss.org/lockss/Home
[14] http://www.clockss.org/clockss/Home
[15] http://www.kb.nl/index-en.html

18. The US government is clearly aware that allowing global public access to the peer-reviewed published output from federally-funded research has the potential to open such content to piracy and other unauthorized dissemination.

19. Such piracy undermines the income that scholarly publishers require to continue their investment in the aforementioned projects, tools and collaborations for the benefit of the scholarly community.

20. The most efficient way to ensure appropriate protection of intellectual property interests of all stakeholders would be to make the final Research Report, provided by the researcher to the funder, freely available. This would allow a rapid and very broad dissemination of the research results obtained directly from federal funding. This would also facilitate such reporting to be tied back to the original grant made by the federal agency. Final project reports could also be linked to the peer-reviewed published research, available online whether free, via rental or for full purchase as the publisher business model dictates.

21. ALPSP is not in favor of mandated deposit to centralized open repositories. In addition to significant concerns about long-term sustainability and piracy, open repositories have deleterious effects on the publishing model; for example, NIH does not currently provide publishers with full, detailed usage statistics from PubMed Central, which means publishers are unable to supply libraries with the complete picture with regard to their institution's use of a wide range of journals. Such usage data is crucial in determining renewals and whilst this situation persists, subscriptions are being cancelled based on incomplete usage data.

(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?

22. Studies have demonstrated that researchers prefer to access the publisher-created Version of Record (VoR) from a peer-reviewed journal as the authoritative, definitive version, over versions in subject or institutional repositories[16],[17].

23. In an interconnected age, with current and ever-improving technology, centralization is not required and moreover, requires unnecessary duplication of effort at considerable expense. Indeed the report from the Scholarly Publishing Roundtable in January 2010[18] recommended decentralization to achieve the interoperability needed to "enhance the impact of the scholarly literature and ignite the generation of new knowledge".

24. Publishers have gone to considerable lengths in developing tools to ensure interoperability between different access systems. For example the Digital Object Identifier (DOI[19]) system, to provide persistent identification of digital objects, the

[16] http://www.peerproject.eu/reports/ D4.2 PEER Behavioural Research – Final Report
[17] http://www.publishingresearch.net/projects.htm Research Publication Characteristics and Their Relative Values
[18] http://www.aau.edu/WorkArea/DownloadAsset.aspx?id=10044
[19] http://www.doi.org

CrossRef [20] organization and its various ongoing projects aimed at connecting users with primary research content, and the Open Research and Contributor ID (ORCID[21]) initiative, to solve author name ambiguity in scholarly communications and latterly resolving institutional naming ambiguity.

25. Publishers are also continuing to invest in the development of discipline-specific tools to enable users to interact with and analyze specialized content. Such tools would be lost with centralization.

26. Publishers are continuing to invest in metadata standards, which improve the ease with which relevant articles can be discovered. With such excellent standards, search tools are all that is required to connect users with the most appropriate content for their needs, and importantly to the VoR. Such metadata standards include those developed by EDItEUR[22], IDEAlliance (PRISM)[23] and NISO[24] (see also paragraphs 33 and 34 below).

(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?

27. In addition to the many public-private partnerships already mentioned, publishers are keen to engage further with Government and its agencies. Proposals have already been put to NSF for collaborative projects to enhance the public access, utility and preservation of publications resulting from federally-funded research.

28. Such proposals include standardizing the collection, display and use of metadata to indicate the federal grant supporting the research from which a scholarly publication derived and potential linking back to the Federal Agency website. A further example is the proposal for a project to understand the requirements for and benefits derived from content mining and to establish a methodology for overcoming current barriers, such that publishers can facilitate such content mining with sustainable business models.

29. These are just two of the proposals under discussion with the NSF.

(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?

30. As already mentioned above (paragraph 28), publishers are already undertaking a project with CrossRef and the Department of Energy (DoE) to standardize the way

[20] http://www.crossref.org
[21] http://orcid.org
[22] http://www.editeur.org/
[23] http://www.idealliance.org/specifications/prism/
[24] http://www.niso.org/standards/

funding information is collected publishers and included in article metadata. This would enable Federal agencies to easily obtain information about publications resulting from federally-funded research.

31. Such collaborative projects enable cost-effective standardization across all Federal agencies and publishers.

32. Metadata allows users to discover information and find related information without the requirement of accessing the full text. Two initiatives are important in this regard.

33. The Dublin Core Metadata Initiative[25] provides key specifications and best practice regarding the use of metadata for the description of various digital resources (including books and journal articles). It enables interoperability of different applications and vocabularies and optimizes the metadata for searching.

34. CrossRef[20] provides a cross-publisher linking network. This allows readers to easily link to other resources of interest on other publisher platforms. This works seamlessly through DOIs and metadata which are embedded in articles and other content as part of the value-added publication process.

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

35. Federal agencies funding scientific research should maximize the products that they invest in, that is the Research Reports required by Federal agencies from the research scientist. Some already make such research reports available (e.g. the DoE Information Bridge[26]), but others do not. Making all such reports freely available would solve the "public access" issue.

36. Federal agencies do not invest in peer-reviewed journals. Publishers add significant value to peer-reviewed publications and this is reflected in researcher preference for the VoR[16,17]. Publishers should then be at liberty to employ appropriate business models by which they may recover their investment and to reinvest.

(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?

37. No. Publishers invest considerably in all types of content they produce to add value to the scholarly and academic community that utilize them. Such publications should not be appropriated without rightsholder permission and compensation. To behave otherwise would compromise the sustainability of high quality publication, dissemination and preservation of the research results.

[25] http://dublincore.org/
[26] http://www.osti.gov/bridge/

(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?

38. There is no single "appropriate" embargo period. Federal agencies should not impose inappropriate embargo periods on non-federally funded businesses. Individual publisher business models are not arbitrary, but are carefully calibrated to meet the needs of the market and the investment made.

39. The most common current embargoes range from zero, for gold Open Access material, to 12 months, as a result of the NIH-mandate. Publishers, however, should be able to set their own appropriate embargo, depending on the material they publish and the market for which they publish, and this may be more or less than 12 months.

40. An indication of the length of usage an article in a given discipline received, the journal half-life forms a useful measure. For example, the American Physiological Society reports journal half-life from 4.3 to over 10 years[27]. The quarterly journals of the American Anthropological Association also have a cited half-life of over 10 years and 90% of downloads occur 12 months after the date of publication. In mathematics papers published in 2009, 50% of citations were found to be to papers originally published before 1999, with 20% of citations to papers published before 1985[28].

41. Imposing mandates on the potential to recover investment from such usage further undermines publishers' ability to continue to innovate and add value for the benefit of the scholarly and academic community.

42. In the current economic climate, recovering investment is all too important. Journal budgets are being squeezed and foreshortening the length of time a publisher is able to recoup their investment has the potential to seriously damage publishers and therefore the overall economy.

43. As already referred to, the lack of transparency demonstrated by NIH has the potential to undermining the entire system. Librarians utilize usage statistics as part of their considerations for journal renewals. Whilst publishers have worked with NIH to assist authors in fulfilling their mandated deposit, NIH has been unwilling to provide publishers with detailed usage statistics, which would allow publishers to provide a more accurate picture to librarians of the usage of journals by their faculty.

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.

---

[27] http://www.the-aps.org/publications/journals/info/impact_factors.htm

[28] http://www.msri.org/attachments/workshops/587/MSRIfinalreport.pdf Donald E McClure (2011) Dynamics of Mathematics Journals, 2000 to 2009

44. Scientific research and scholarly communication is an international enterprise. Any efforts to improve "public" access through collaborations, standards or other projects, should necessarily be considered on an international, rather than national scale, if the real benefits of improving access to data are to be efficiently and cost-effectively recognized.

45. Publishers are very willing to enter into collaborative projects to explore the nature of these issues with the aim of producing the most cost-effective and appropriate solutions for all stakeholders.

Dr. Audrey McCulloch
Acting Chief Executive

On behalf of the ALPSP membership.