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

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
www.alpsp.org

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.