12 January 2012

To: Office of Science and Technology Policy

From: Eugene Arthurs
CEO, SPIE
Bellingham, WA 98225
Eugene@spie.org

Re: SPIE response to "Request for Information" (RFI) by the Office of Science and Technology Policy, Executive Office of the President, regarding "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research"

SPIE Background

SPIE (Society of Photo-Optical Instrumentation Engineers) is a not-for-profit society serving the international optics and photonics community. Our mission is to foster knowledge transfer, education, and networking among researchers, technology leaders, educators, and students engaged in the broad interdisciplinary sectors we serve. SPIE's goals to support technical innovation and the development of a well-educated and highly trained work force are concordant with those of the America COMPETES Reauthorization Act of 2010.

SPIE publishes seven peer-reviewed journals covering the fields of optical engineering, electronic imaging, biomedical optics, microelectronics, nanophotonics, remote sensing, and energy. We organize over 300 technical conferences annually in a diverse range of fields and publish proceedings for each of those. Combined, SPIE publishes 18,000 journal and proceedings research papers annually. We also publish 20 or so books each year and one print and two online magazines. We have a balanced distribution of members (currently 16,200) and authors from the academic, industrial, and government sectors as well as significant geographic diversity. About 40,000 individuals participate with SPIE each year. Our current estimate for the lower bound of the global photonics market is just under $500 billion. The enabled market and services is much greater-the entire internet is photonics enabled.

SPIE has already made a reasonable commitment to open access to support the education and research needs of our community. We publish two fully open access virtual journals containing letters, review articles, and tutorials (*SPIE Letters* and *SPIE Reviews*). SPIE fully subsidizes these two journals so there is no required fee for either authors or readers. Our subscription-based journals provide authors with an immediate open-access option with payment of a modest publication fee. In addition, for the *Journal of Biomedical Optics* we deposit NIH-funded articles with PubMed Central on the authors' behalf and make articles for which other authors pay voluntary page charges open access 12 months after publication. Finally, all articles published in our magazines are open access at no cost to readers or authors.

While we have some measure of commitment to open dissemination of our publications, we also face the reality that there is a substantial cost to maintain a diverse, innovative, and healthy program. SPIE has followed the open access dialogue closely for the past several years and recognizes the polarity of the issue and the significant implications that government policies could have on the scholarly publishing community. We appreciate this opportunity to provide input.

General Comments and Viewpoints

Society publishers are indeed fortunate for the volunteer support we receive from the expert editors and reviewers who vet and improve the content we publish. Many of these volunteers are our members and Society leaders and help define SPIE's organizational priorities, strategic direction, and altruistic activities. Their contributions do not appear on financial statements. The majority of our costs, therefore, are incurred in the administrative, editorial, and production functions that transform manuscripts into cogent journal articles, in the costs of disseminating this information to a global readership, in archiving published materials, and in making them findable and available in current media formats.

SPIE believes that an unfunded, broad-based free public access policy would result in unnecessary challenges and dilemmas for publishers, scholars, institutions, and potentially the funding agencies themselves. We offer the following perspectives for your consideration and discussion:

- Except for some recent examples of successful author-funded open access journals, subscriptions are the most common and time-proven way to provide access to technical information and to compensate publishers. Papers of particular interest to researchers and the public are accessible via subscriptions, pay per view, document delivery, public access libraries such as state universities, interlibrary loan, author websites, or directly from authors, most of whom welcome interest in their work. They can be obtained almost immediately for a reasonable fee or sometimes at no cost at all. Access to a journal article via the parent publication ensures that a cited article can be easily found via DOI link, OpenURL link, search engine, or some other referral source. Once journal articles are dispersed to other locations such as government or university repositories, they become more difficult to find and their long-term integrity and preservation become indeterminate.

- Those publishers that already have open access journals charge authors or their institutions a publication fee. This model biases science, discriminating in favor of researchers who can afford to have their work published in these journals, and those that can often allocate funds from their research grants to cover this publishing fee, thus cutting into available funding for the research itself.

- Public access to research funded by the United States Government will intrinsically be accessible to the entire world, not just the citizens and organizations whose taxes have funded the work. Therefore, many beneficiaries of this paradigm will not have contributed in any way to the research, nor will they have contributed to the cost required to publish and make it available. In effect, U.S. taxpayers, authors, and publishers will be subsidizing the research itself and providing the results to the rest of the world with no reciprocal contribution to the cost. There are good arguments for open science for the life sciences and medical research. Less so for the science and technology we serve, much of which turns into products and markets. We believe there already is an unbalanced net flow of publications of funded scientific research from the U.S. Open access will add to this, undermining of U.S. competitiveness.

- Patents and other forms of intellectual property rights are permitted for federally funded research, providing future income streams for funding recipients, universities, government contractors, small businesses, etc. This is a positive thing for U.S. competitiveness. We do not believe that the income stream generated from the publication of research should be differentiated from other forms of IP commercialization by beneficiaries of research funded in whole or in part by the U.S. Government. In fact, much of the published work funded by federal agencies is subject to copyright and all the protections afforded by U.S. Copyright Law, and a public access mandate may actually countermand

those protections. This potential dichotomy between public interest and private intellectual property rights should be considered carefully in formulating any public access policy.

- Recipients of U.S. Government research grants submit progress and final reports to funding agencies. These reports represent the most complete and comprehensive record of research outcomes available. Several agencies such as DOE and NASA already make some or all of these reports publicly available through databases such as Science.gov, Information Bridge, DTIC, etc. A unified effort by the agencies to integrate this abundance of research output in a U.S. Government database or portal would provide a far more complete, comprehensive, and timely public record of funded research than any form of required open-access publication in widely dispersed journals or repositories. While PubMed Central might be a model for such a repository, the NIH requirements on authors and publishers are increasingly burdensome and the cost and effort to establish and maintain similar sites would likewise be burdensome to the agencies, the researchers whose work they fund, and to the publishers that have to meet deposit requirements. A sweeping open access policy would be administratively challenging, and costly, to execute and to monitor. Why not entrust publishers and the information economy to continue to assume that responsibility and let the agencies focus on what they do best?

- A policy that focuses on papers published in peer-reviewed journals can only partially fulfill the goal of providing public access to all unclassified research funded by the U.S. Government. To the best of our knowledge there is no requirement that funded authors publish their research in a peer-reviewed journal. We are not aware of statistics related to how much of this work is actually published. Even if a research paper is submitted to a peer-reviewed journal, there is no assurance that it will be accepted. Undoubtedly, some fraction of the research is simply inappropriate or premature for journal publication. For example, the failure of a new drug is not normally published in the medical literature, yet negative studies can be very valuable. What is the best way to expose this information?

- Despite recent improvements in timeliness of peer-reviewed journals, there is still a time delay between completion of research and journal publication. Add to that a potential embargo period, and the results of publicly funded research are less than timely enough to satisfy the ongoing research needs of other practitioners, including those working on projects funded by the U.S. Government. A more immediate and direct dissemination of project reports and data would do a lot more to advance the innovation and competiveness goals of America COMPETES.

- In addition to support provided by government grants, researcher teams are supported by their universities, companies, and labs in the form of facilities and equipment, support personnel, and information resources, each of which is essential to sustaining a successful R&D program. Information is neither more nor less essential than the other needs. Information acquisition is part of an organization's business cost, and journals are no less a commodity than many of these other research needs and have similar economic realities. It is difficult to imagine a model where the cost to the information consumer is subsidized by the information creator, but isn't that effectively the result of open-access publication?

- Publishers add substantial value to the papers published in peer-reviewed journals, from the peer-review process itself to copy editing, typesetting, and production to global dissemination. This process occurs subsequent to the research itself and is funded primarily through subscriptions. Journal papers are derivative products of research, not the grant deliverable. There is no

justification for the government to claim IP rights to the journal article, with the value-added investment made by publishers. There is no dispute that the government owns the research reports and the data and it is this information that the government should seek to publish, again with some preferential access for the U.S. taxpayer.

- In addition to peer review, copy editing, and composition services, publishers ensure that journals are properly archived and preserved via services such as Portico and CLOCKSS. We ensure that content is indexed in A&I databases such as Web of Science, Inspec, Medline, Scopus, etc., and covered by search engines. Publishers make substantial investments to ensure the long-term integrity and access to content. This infrastructure ensures the long-term wide accessibility and preservation of the scientific research literature and is the result of significant investment and commitment within the private sector. We believe the needs of U.S. science and scientists are well served by these programs and that it would not be practical or productive for government to attempt to replicate them solely for taxpayer-funded research.

- Free public access to federally funded research will disrupt the economic model of many society publishers. Our subscription fees relative to most for-profit publishers are moderate and reasonable. While some economic reform within the information industry overall might be warranted, this is not the way to get there. A healthy economic system should be self-regulating and balanced by the usual supplier/consumer dynamics.

The Internet is a wonderful thing, putting information literally at the fingertips of those who need it, sometimes at a cost and sometimes for free. Where there is a cost, the consumer must decide whether the cost is justified, whether it is likely to provide value. Where information is free, there is usually another income option available to the provider, or another motivation.

Scholarly publishers like SPIE provide a level of rigor and objectivity that aims to ensure the provenance, quality and integrity of information published under our auspices. This sort of reliability is not always assured with information that is freely available on the Internet. The web contains a plethora of “open-access” information of questionable authenticity and value alongside an alarmingly large body of pirated and plagiarized content. The reputation of Society as well as the respectable commercial publishers depends on the public’s trust in the quality and provenance of information we deliver.

A government policy that requires open access publication will inevitably disrupt a system that has served the needs of scholars and inventors, classrooms and laboratories, publishers and librarians, through a long and fruitful period of amazing scientific and technological advancement. We don’t presume to believe that free public access or some disruption of publishers’ business models will be the undoing of any of this. Some would argue that such a disruption is needed. We assert, however, that the contemplated action will have the effect of imposing a new and not necessarily fair or sustainable economic model on the enterprise of scholarly publishing and have additional implications as suggested above. We appreciate the opportunity to submit this input and urge the Task Force to look closely at alternatives to mandated open access publication of journal articles derived from federally funded research.

Comments on OSTP Questions

(1) Are there steps that agencies could take to grow existing and new markets related to the access and analysis of peer-reviewed publications that result from federally funded scientific research? How can policies for archiving publications and making them 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?

All U.S.-based science and technology publishers and societies are stakeholders in a productive scientific enterprise within the United States. SPIE is a society first and a publisher second and our priorities lie with supporting growth in optics and photonics science and industry and the constituents we serve. The U.S. is the world leader in aspects of our field, and research supported by the U.S. Government is essential to retaining technology leadership; therefore our goals are closely aligned with America COMPETES. Analogously, access to information is essential to enduring progress in science and technology. In our publisher role, our first goal is wide dissemination and availability at reasonable cost. The majority of our subscribers are U.S. institutions. We are continuously seeking ways to expand access to our publications and have reduced prices to meet that objective. Our pricing history is starkly different to that of commercial publishers, as are our goals.

The key is to get the right information into the right hands at the right time. With today's pace of development, that means immediately. Therefore, we believe a comprehensive public research report repository posted on acceptance by contract managers is the correct approach. If laypersons want access, for example, to medical information, that's an ancillary benefit. The key is to enable the people who need it most to be able to access information quickly, including all the methodology and data. We live in an information-rich world and it's challenging to navigate this abundance. Therefore, it would also be beneficial to consider how the information could best be filtered, or even to consider some standardized reporting using XML templates in order to facilitate dissemination, alerting, etc. If U.S. funded research is publicly available, inevitably it will be available to anyone anywhere. Perhaps some additional information filtering or user authentication layers should be developed to provide an advantage to U.S. constituents.

We do not see it being economically practical for the U.S. Government to replicate the NIH PubMed Central model for a host of disciplines and contend that formal publication and dissemination is best left to publishers. It seems to us that the most economical way to achieve the access goals of America COMPETES is to invest in a robust integrated database including all publicly funded research and to invest in the technology to facilitate access to this information by those who need it most—our researchers.

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

This is an important topic for the fields we cover as much of what makes up SPIE's area of interest has or might be commercialized or applicable to national security. We recognize that what is covered by other societies is valuable in expanding human knowledge, and different considerations apply for IP. Allowing grant recipients broad patent and other IP freedoms encourages innovation and commercialization. Competition and free enterprise are the mainstays of the U.S. economy, and any measures that

discourage innovation and commercialization will only impede progress. This applies to technological enterprise and to publishing as a competitive business. The pace of change in the information industry over the past two decades has been extraordinary. Just as labs and companies need resources and incentives to fuel investment and innovation, so too do publishers. Policies that destabilize the economic foundation of the publishing industry put that industry at risk of consolidation through mergers and acquisitions that can only give advantage to the wealthiest commercial publishers, a large number of which have their origins and base of operations outside the U.S. If an analogous situation were true for the science and technology industrial base, competition would be undermined and the pace and direction of innovation would inevitably slow, or move elsewhere. There is great advantage in providing incentives and opportunities to start-ups and to entrepreneurs because often those are the pathways to innovation. We do not believe anything that weakens a vigorous competitive environment in science and technology or in publishing is in anyone's best interest.

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

The agencies already have the documents needed to disseminate the results of federally funded research. What is lacking is an integrated database or portal to enable access to this corpus and the infrastructure to manage and maintain it. Even if the government were to centrally host agency-funded journal articles, because these comprise only a fraction of all published articles researchers would still need access to the body of information not directly funded by the U.S. Government. Therefore, subscriptions would still be necessary. Using Google Scholar or Web of Science or Inspec would still be necessary. A central repository of federally funded U.S. papers will not sufficiently meet the information needs of U.S. scientists, engineers, teachers, and students. There's no clear evidence that such a repository will catalyze an industry-wide paradigm shift toward open access journals and therefore it really can't truly meet or transform the information needs of the U.S. research community.

Search is best left to entities for which search is their core competency and business. Publishers are not in the search engine business—Google, Microsoft, Endeca, and others do it far better than we ever could. The reality is that most researchers use Google or Google Scholar, even when they have other resources available to them. We would not advise reinventing the wheel but rather deploying a good search engine within a government portal as well as exposing the data to all major search engines.

Most publishers have long-term archiving strategies. We understand the need and our institutional customers expect it. For example, SPIE participates in both Portico and CLOCKSS because we believe it's our responsibility to ensure long-term ("perpetual") availability of our publications. A large number of STM publishers already participate in one or both of these programs. There is a nontrivial cost for these archiving solutions, which is a reminder that publishing needs to be more than an altruistic activity. It would not be efficient for the U.S. Government to assume yet another archiving role that would ensure preservation of only a portion of the published literature. Publishers take this responsibility seriously and have implemented strategies to preserve all of it.

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

There are many examples of cooperative ventures in publishing. Publishers strive to standardize their data using protocols such as the NLM DTD so that content can be easily shared and rendered in a common format. CrossRef has perhaps been the most visible and impactful publisher collaboration. Our data can be easily aggregated in key index databases such as Summon, Web of Science, and Scopus. The OpenURL standard and CrossRef are fully interoperable, and new initiatives such as ORCID continue to reflect a common need for interoperability across publisher databases. COUNTER is another example of publishers assuming responsibility to develop industry-wide standards. None of these focuses specifically on federally funded research but all federally funded content benefits from these ongoing publisher initiatives to improve the interoperability and discoverability of STM information.

In the end, putting information derived from federally funded research and research funded by other governments and organizations in different delivery systems will only fragment these efforts to improve standardization and interoperability, at significant cost.

(5) What steps can be taken by Federal agencies, publishers, and/or scholarly and professional societies to encourage interoperable search, discovery, and analysis capacity across disciplines and archives? What are the minimum core metadata for scholarly publications that must be made available to the public to allow such capabilities? How should Federal agencies make certain that such minimum core metadata associated with peer-reviewed publications resulting from federally funded scientific research are publicly available to ensure that these publications can be easily found and linked to Federal science funding?

The response to question (4) above addresses this in part. Publishers have been mostly self-regulating about metadata standards creation and DTD uniformity and the situation is steadily improving. Even though publishers more or less compete, we recognize the importance of interoperable content elements to our customers and consequently to our businesses. Through CrossRef and other alliances, publishers have taken the initiative to fund and develop protocols that benefit not only the publishers themselves but also customers and users. We would hope that government input would focus on how information sharing might be improved to better serve our common interests. However, we don't believe that any sort of independent government standards or initiatives are warranted or desirable.

(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?

Peer-reviewed papers are only some fraction of the research funded by the federal government, and even those that are published generally include only a portion of the data from that research. In our view, the best means to maximize public access to all the information is to make the reports quickly available in a comprehensive database and to standardize metadata and formats by providing templates to researchers. Journal peer review and publication is a step beyond the research itself and is undertaken by researchers for a variety of reasons. It might be out of social responsibility to the community or it might be for career advancement and recognition. Regardless of the incentive, preparing and publishing a journal article requires a significant amount of additional effort, time, and cost on the part of the author and the publisher.

Awardee institutions already have IP and access rights to the work of their employed researchers. Most publishers permit authors to deposit their papers in a local repository and/or post their work on their own website. For most federally funded research, the U.S. Government already retains the right to use the information for government purposes and in fact holds the primary information from that research

in the form of research reports. Derivative publication in a peer-reviewed journal does not change the results, only the articulation of the work and its impact. Libraries have been the traditional conduit for this information. They are also publishers' primary revenue source. It's an economic model that has worked for the scholarly community for a long while. It is true that costs have escalated and in a few cases publisher profits may seem excessive, mainly in the commercial publishing world. Libraries are beginning to push back and have had some influence on pricing. How libraries themselves are ultimately funded is a complex question in itself. Ultimately someone has to bear the cost of journal publishing. So, what is the most fair and least disruptive model, and the one that best ensures access to primary users and to authors? It could be argued that access has never been greater than it is today, either through the auspices of an institution's library, personal access through membership or low-cost subscription, free availability on the web, or simply requesting something from an author. The current model provides a good balance so that most, if not all, interested readers either have access or can obtain what they need for a modest cost or for no cost. It also provides the needed revenue to publishers to produce high-quality products and to encourage and fund innovation and technical progress.

A federal open-access mandate will ultimately work against the interest of the U.S. taxpayer and U.S. science and technology enterprise because the knowledge flow is unbalanced, with the outbound flow of U.S. research and innovation exceeding the inbound flow and providing advantage to non-U.S. competitors. The current means of disseminating research creates a more level playing field.

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

We believe the same arguments regarding public access apply to other publication formats. SPIE publishes both proceedings and books. Proceedings are often preliminary reports of work in progress and their primary value is to other researchers. In a few fields, proceedings are the preferred format. Books and book chapters typically serve an educational purpose and typically do not report on specific research projects. We believe that the report approach would include much of this information and be far more comprehensive and complete.

(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?

Since we are recommending against adopting a policy that requires journal articles to be open access either on publishers' site or in a system developed and managed by the U.S. Government, we cannot recommend an appropriate embargo period. No embargo would be necessary or even desirable for a government managed report database.