AIP | American Institute of Physics

One Physics Ellipse, College Park, MD 20740-3843   H. Frederick Dylla, Executive Director and CEO

22 December 2011

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

About AIP
The American Institute of Physics is a 501(c)(3) not-for-profit membership corporation created in 1931 for the purpose of "the advancement and diffusion of knowledge of the science of physics and its applications to human welfare." AIP is an organization of 10 physical sciences societies representing more than 135,000 scientists, engineers, and educators and is one of the largest publishers of scientific information in physics, with activities extending well beyond publishing. AIP delivers valuable resources and expertise in education and student services, science communication, government relations, career services for science and engineering professionals, statistical research, industrial outreach, and the history of physics and other sciences.

As a publisher, AIP plays a central role in the process by which scientific research is developed, communicated, disseminated, and ultimately accepted by the scientific community. AIP publications include 15 journals (three of which are published in partnership with other societies), magazines, including its flagship publication *Physics Today,* and the *AIP Conference Proceedings* series. In addition to its own publication, AIP provides publishing services and expertise to five of its ten Member Societies. To accomplish this, AIP invests millions of dollars annually on peer review, editorial management, production, printing, shipping, distributing, and hosting its archival journals on a fully digital, highly reliable online platform, making the content available at all times to customers around the world in more than 70 countries.

Whether an article is read online or in print, high-quality peer review, page composition (XML), copyediting, and the listing and linking of bibliographic and reference data must be managed, necessitating considerable human capital investment in staff at our Melville, New York, publishing center, in addition to more than 340 editors around the world. Our editors maintain the quality and reputation of our journals, utilizing the well-established system of peer review, whereby independent

AIP

experts review submitted articles. Accepted articles are those that pass muster based on established criteria, including novelty and the substantial nature of the research findings. Managing peer review for approximately 30,000 articles submitted to AIP journals every year is a complex undertaking. It requires a large amount of sophisticated electronic resources, associated support personnel, a staff of professional editors—nearly all PhD physicists—and help from tens of thousands of referees. Each year AIP makes such necessary investments to fulfill its public nonprofit mission, generating an intellectual return through the dissemination of scientific research.

Introduction

AIP's highest goal is to achieve the widest possible dissemination of the research results it publishes, including any pertinent associated data and context information. Enabled by Internet technologies, AIP disseminates more information, more widely and more affordably, than ever before in its history, reaching more authors, subscribers, and users than ever before. This accomplishment requires heavy investments in technology and infrastructure (such as an online platform) and business-model innovation to deliver the option of free or low-cost access: open access, pay-per-view, or article rental, recognizing that the value of the final published article needs to be paid for to remain sustainable.

AIP believes that it would be in the best interest of the United States and its government, as well as in the best interest of all other stakeholders, to strike a balance between public access and sustenance of the scholarly publishing industry because of the impact and value it brings to the progress of science and its contributions to American society and economy. Such a balance can be achieved based on shared principles such as the importance of peer review, the recognition of economic realities through adaptable and viable publishing business models, the need to ensure secure archiving and preservation of scholarly information, and the desirability of broad access. One way to achieve this balance is for government to adopt a sensible, flexible, and cautious approach to drafting public access policies—an approach that engages all concerned parties, including federal agencies, scientists, university administrators, librarians, publishers, and the public.

Consistent with the recognition of economic realities, it is AIP's position that government agencies should develop their public access policies through voluntary collaborations with nongovernmental stakeholders, including researchers and publishers. Any policies should be guided by the need to foster interoperability of information across multiple databases and platforms. Agencies' efforts then could be directed toward facilitating cyberinfrastructure and collaboration programs with and between agencies and the stakeholders to develop robust standards for the structure of full text and metadata, navigation tools, and other applications to achieve interoperability across the scholarly literature. More detail on this is provided later in the document. AIP believes that any scholarly publication access policy needs to be flexible to accommodate agency-specific needs and have the capacity to evolve in response to the rapidly changing nature of scholarly publishing.

AIP

AIP Responses to RFI Questions

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

According to trade association and other industry surveys of US publishers, both the nonprofit and commercial sectors serve a robust, innovative global market for the access and consumption of peer-reviewed publications. Academic, corporate, and governmental research and education communities constitute primary segments of the market. Global revenue from scholarly journal publishing was estimated at $8.0 billion in 2008,[1,2] with approximately $3 billion attributed to the US market. The enterprise employs approximately 110,000 people worldwide, with 30,000 in the United States. New publishers, journals, and business models either evolve or emerge constantly, signaling a healthy, competitive marketplace.

The combination of investments in digital and online technology (by publishers as well as others) and the formation of library consortia (assisted by publishers in many cases) in the United States and around the world has accelerated and broadened access to peer-reviewed literature and dramatically decreased cost of such access. AIP serves approximately 2,000 research institutions, and every person affiliated with these institutions has instant access to AIP journal content.

There is a growing presence and diversity of business models in the scholarly market. It is our belief that the government should support and encourage this diversity through its actions and policies through sustainable partnerships with publishers that would contribute to the US economy and maximize the productivity of the scientific enterprise. (For AIP's suggestions of partnerships and pilot projects that would meet mutually beneficial goals and conserve precious federal research funds for the agencies' primary mission of funding research, please see the responses to Questions 4 and 5. These recommendations for partnerships and pilot projects with federal agencies were developed in collaboration with a number of scientific publishers as we engaged over the last year in productive discussions with subject matter experts within the NSF and DOE, two US federal agencies that fund substantial research in the physical and biological sciences and engineering.)

As stated in the 2010 *Scholarly Publishing Roundtable* report,[3] many publishers have made the decision to move toward increasingly open structures and archives,[4] as enabled by open access business models

---

[1] Cox, J. and L. Cox, *Scholarly Publishing Practice: Academic Journals Publisher's Policies and Practices in Online Publishing*, 3rd ed., ALPSP (2008),
http://www.alpsp.org/ngen_public/article.asp?id=200&did=47&aid=24781&st=&oaid=-1.
[2] Outsell, "An Open Access Primer-Market Size and Trends" (2009),
http://www.outsellinc.com/contact_us/open_access_primer_2009.
[3] Report and Recommendations of the Scholarly Publishing Roundtable, January 2010, available at www.aau.edu/WorkArea/showcontent.aspx?id=10044. Referred to throughout this document as the *Roundtable Report.*
[4] Morris, S., *Journal Authors' Rights: Perception and Reality* (London: Publishing Research Consortium, 2009), http://www.publishingresearch.net/documents/JournalAuthorsRights.pdf.

AIP

and new solutions to associated permissions such as Creative Commons[5] licenses. These licenses provide a means for exercising certain rights regarding the re-use of an item. For example, these licenses could provide re-use rights if the resulting new works are also made available to the public. The *Roundtable Report* also notes that the number of journals making a change in business model is appreciable but small within the universe of more than 25,000 scholarly peer-reviewed journals.[6] AIP echoes the *Roundtable Report* assertion that no existing digital business model has demonstrated its viability to the satisfaction of all, and cautions against government endorsement of any single approach.

As part of the market's evolution and scholarly publishers' commitment to community and distribution of results, an increasing number of all types of journal publishers are electing to make their articles freely available to academics and others in 100 or more developing countries. Some well-known programs include the United Nation's HINARI, AGORA, and OARE Research4Life programs, HighWire's Developing Economies Program, and JSTOR's Developing Nations Initiative. Additional programs include those of EIFL, INASP, and TEEAL. For descriptions of these and more, see www.library.yale.edu/~llicense/develop.shtml.

To meet the market's increasing demand for easily accessible, quality information, AIP invests considerably in new technologies for viewing and sharing its journals. Within just the past two years, AIP developed a mobile phone reader for journals, a professional (and freely available) social networking site for physical scientists (www.aipuniphy.org), and an electronic book platform. AIP also launched a multimedia journal on renewable energy (http://jrse.aip.org) and one of the first community-style journals in the physical sciences (http://aipadvances.aip.org).

Such ongoing investments in existing products and services and the development costs for new products are funded through subscription fees or author payments. AIP and most other scholarly publishers offer an open access option for authors, no matter what type of journal they decided to publish in. Through AIP's Author Select, authors have the option to choose open access for their published article. Less than one percent of authors choose to do so. *AIP Advances*, AIP's newest journal, is an initiative based on community-style review, rapid publication, is fully open access, and employs a Creative Commons license.

This ability for scientific publishers to experiment with different publication, business, and access models is paramount and assures the vitality, diversity, and effectiveness of scholarly communication, leading to scientific and technological advances. Rather than mandate business models and de-incentivize market efficiencies, a more effective approach by government would be to incentivize the continued growth and vitality of the scholarly communication market for the benefit of the scholarly community. To that end, working with publishers, libraries, and other stakeholder communities, research agencies should identify specific needs of particular user groups and collaborate with publishers to meet those needs most effectively. Obviously, researchers, professionals, funders, and various segments of the general public (e.g., patients) have different information needs. AIP is collaborating with other scholarly publishers to identify and address any existing access gaps through

---

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

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

AIP

initiatives such as the low-cost article rental scheme pioneered by DeepDyve, the Research4Life consortium for developing countries (mentioned above), the patientINFORM portal for patients or their caregivers, the Emergency Access Initiative offered to communities affected by natural disasters, and free or substantially discounted access for public libraries, journalists, and high schools.

Nevertheless, based on our experimentation with a modest-cost article rental model (through DeepDyve), AIP remains unconvinced that there is a large unmet demand for public access: only a few thousand members of the general public attempted to access our scholarly content over a year's time, compared to the nearly eight million visitors to AIP content on our online platform.

To maximize the effectiveness of its efforts, government has an important convener role to play in developing standards for data and metadata, and making research more readily searchable and discoverable. Publishers are already working in partnership to develop standardized information and collections through initiatives such as CrossRef.[7] (For more detail on this, please see response to Question 5.)

With a relatively straightforward implementation of existing policy, government could make the funder-collected and maintained outputs of taxpayer-funded research, such as grant reports and research progress reports, freely available to the public.[8] Furthermore, to incentivize open access publishing, funds could be made available specifically to support payment for open access to published articles as pilot projects. Several research funders already do this (Howard Hughes Medical Institute, Wellcome Trust, and Max-Planck Institutes).

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

AIP has been a leading participant in organizing working groups that are proposing and planning partnerships with NSF and DOE on access, linking of grantee reports to publications, data mining across agency-publisher databases, tools and methodology for identifying publicly funded work, and potential pilot projects in the above areas. (More detail on this can be found in response to Question 5.)

Government mandates for public access come at a significant cost to the US economy and to the scientific enterprise. The National Institute of Health's (NIH) PubMed Central (PMC) data indicates that two-thirds of its users are from overseas. This suggests that critical export opportunities for the industry may be compromised, resulting in loss of US jobs.[9] Significant economic value added generated by the publishing industry could be wasted if revenue derived from sales in the global market is compromised

---

[7] CrossRef (www.crossref.org) is a not-for-profit group founded by publishers in 2002 and maintains 50 million items. Almost 1000 publishers participate, assigning Digital Object Identifiers (DOIs) to published content items. Development of the CrossRef service has resulted in seamless navigation of the research literature by users so that researchers using the bibliography in one article can link from a reference to the full text of the referenced article.
[8] This would ensure readability to the broadest audience. NSF is already pursuing such a policy, see http://www.nsf.gov/pubs/policydocs/porfaqs.jsp, and DOE through its Office of Scientific and Technical Information provides public access to nearly 300,000 DOE-funded research reports, see http://www.osti.gov/bridge/.

AIP

or eliminated. Furthermore, mandates often result in additional costs for publishers. For example, although only a very small portion of AIP's content is subject to the NIH public access mandate (AIP is primarily a physical science publisher), AIP had to incur costs to modify formats and procedures in order to deposit manuscripts into PMC. AIP remains concerned that PMC is shifting readers from the publishers' sites to PMC despite linking arrangements, thus undermining the value of the publishers' investments.

AIP has concerns about any government policy affecting global trade balance. The number of papers submitted to AIP journals from China exceeded the US submissions two years ago. In response, AIP opened an editorial and marketing office in China to help promote established AIP journals in the physical sciences in China, rather than see China develop competing international journals. Free and unimpeded access to US journal content, even if one were to factor in a short embargo period, will undermine our and other US publishers' needed revenue to establish business relationships in potentially lucrative and large global markets such as China.

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

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

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

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

AIP

AIP encourages agency policies and actions that work to ensure copyrighted materials are protected from unauthorized dissemination and piracy. Copyright is an essential ingredient in promoting creativity, innovation, and the continued integrity and reliability of the scholarly record. There is some evidence that the NIH policy undermines intellectual property rights and promotes piracy of intellectual property. As noted in response to Question 1, the NIH public access policy and availability of articles through NIH's database, PMC, undermine an important US export market. Furthermore, copyrighted material downloaded from PMC appears on rogue Internet sites, resulting in millions of dollars in annual losses to US publishers.

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

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

Although a centralized data platform has some obvious advantages of simplicity of operation, the use of a centralized, government-controlled platform for a large corpus of scholarly content has significant downsides, including increased costs to the government. A centralized approach discourages innovation by driving traffic away from innovators, including publishers, thus minimizing scientific and commercial opportunities. However, an important role for government in this arena would be to drive and fund the development of interoperability standards and promote the widespread use of such standards.

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

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

AIP

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

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

What constitutes a publication and the nature of publication is changing with technology. A publication is no longer just a chunk of text fixed in time forever but a fluid representation. Publications can include supplemental material, multimedia files, software, links to resources on the web, and can be revised and corrected over time by the authors and publishers, hence the emergence of new community initiatives such as CrossRef's CrossMark[9] service, which electronically watermarks an article's Version of Record (VoR), and DataCite,[10] which extends the CrossRef-promoted Digital Object Identifier (DOI) to datasets. Any plan for the future should recognize that the static aggregation/library model is not likely to hold up well in the distributed and dynamic Internet milieu.

AIP believes that it is unlikely that one optimal procedure for preservation and stewardship would emerge to become applicable across all of scholarly publishing. For now, AIP strongly recommends that agency policies embrace diversity, decentralization, and interoperability. In the long term, systematic collaborations among stakeholders (government, publishers, universities and their libraries, and other not-for-profit participants in the scholarly publishing system) will be necessary to achieve maximum benefit. We note that libraries, in partnership with publishers, have established entities for preservation of digital documents that are already in wide use, for example, Portico[11] and CLOCKSS.[12]

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

---

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

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

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

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

AIP

equal partners to their mutual benefit in strengthening the scholarly publishing system and expanding public access to its outputs.

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

Yes, please see detailed response to Question 5 below.

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

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

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

Specifics on these items are discussed below.

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

Many scholarly publishing organizations, such as AIP, were founded by scientists for scientists and fully embrace providing publishing and other services as their primary mission. As part of this, AIP's CEO was an active member the Scholarly Publishing Roundtable and subsequently helped organize working groups of nonprofit and commercial publishers to propose and implement joint projects with both the DOE and NSF with a mutually agreed-upon goals. Moreover, AIP is a cofounder of CrossRef and participates in a number of standards organizations such as the National Information Standards Organization (NISO—www.niso.org), National Federation of Advanced Information Services (NFAIS—www.nfais.org), and the newly formed consortium Open Researcher and Contributor ID (ORCID—www.orcid.org), with a purpose to develop unique researcher identifiers.

AIP

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

This proposal has been endorsed by CrossRef and the major scientific, technical, and medical (STM) publishing trade associations: the Professional and Scholarly Publications Division of the American Association of Publishers (PSP-AAP) and the International Association of Scientific Technical and Medical Publishers. Related to this proposal, the DOE's Office of Scientific and Technical Information (OSTI) has agreed to maintain a registry of standard nomenclature for funding agencies and the associated naming and numbering system for grants. OSTI already houses technical reports and data sets for more than 40 federal and international funding organizations.

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

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

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

AIP

interdisciplinary scientific approaches. For supporting information, publishers would draw on their involvement with the joint NISO/NFAIS Working Group on Supplementary Journal Information (see www.niso.org).

Standards and Identifiers: Author Disambiguation

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

Such a standard will not only enhance the scientific discovery process but also improve the efficiency of funding and collaboration. Participation in ORCID is open to any organization that has an interest in scholarly communications. All profile data contributed to ORCID by researchers or claimed by them will be available in standard formats for free download (subject to the researchers' own privacy settings) that is updated once a year and released under a Creative Commons license. All software developed by ORCID will be publicly released under an open-source software license approved by the Open Source Initiative (OSI). For the software it adopts, ORCID will prefer open source. ORCID is governed by representatives from a broad cross section of stakeholders, including publishers, library organizations, research institutions, and funding agencies (see http://orcid.org/board-of-directors).

Discovery Tools: Content Mining

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

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

AIP

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

Pilot Projects: Sponsored Access to Published Research

The "Gold Open Access" dissemination model, which includes an article processing charge paid by the author or their institution, delivers immediate and unrestricted online access to the final published article (defined by NISO as the Version of Record).

AIP suggests that agencies could work with publishers to set up experiments to answer the following questions dealing with the cost, benefits, and sustainability of the Gold Open Access model, as well as investigate how such a model should be funded and administered:

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

Pilot Projects: Linking to/from Research Reports

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

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

Possible outcomes of the pilot could include:

- The ability to identify all agency-funded research within publisher offerings and the ability to deliver associated metadata to agencies,
- The ability to establish mechanisms and approaches that could be implemented (for all research funders) across the industry,

AIP

- A capability to report to major funders on the impact of the research they fund, e.g., through bibliometric and other tools,
- A "research dashboard" capability or the ability to contribute to one already in existence, e.g., http://rd-dashboard.nitrd.gov/,
- A mechanism for low-cost content rental access to published articles (Versions of Record) and a mechanism to explore its impact,
- Subject area content portfolios of agency-funded research articles for internal agency use (e.g., study sections),
- The possibility to use the DOE-OSTI platform (the http://www.science.gov/) to extend this pilot to other federal funding agencies, and
- Models to illustrate how traditional publishing systems can coexist with self-archiving.

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

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

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

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

In addition, please see the response to Question 5 above for specific agency initiatives.

AIP

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

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

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

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

AIP analyzed related industry data using the "cited half-life" metric as a relative indicator for how long journal titles within scientific categories are being accessed and cited, thus reflecting economic viability. The findings could help inform considerations related to embargo periods. Based on the evidence related to AIP journals and to journals covering physics and related sciences, significant economic uncertainty remains with the assignment of minimum embargo periods. In looking at a sample of several physics and related topics and AIP journals within those categories, AIP found that physics journals have a longer cited half-life compared to some other scientific disciplines, and furthermore, AIP Journals have a longer cited half-life than their respective physics category averages.

The chart on the following page provides are some examples.

AIP

Cited Half-Life of AIP Journals/Physics vs. Other Scientific Disciplines
The cited half-life for the journal is the median age of its items cited in the current year. Half of the citations to the journal are to items published within the cited half-life.

Scientific Categories	Average Cited Half-Life within Sub-Category (Years)	Average Cited Half-Life of AIP Journals within Sub-Category (Years)	AIP Journals in Category
Applied Physics	5.6	7.8	*Applied Physics Letters, Journal of Applied Physics, Journal of Low Temperature Physics, Review of Scientific Instruments*
Chemical Physics	7.1	>10	*The Journal of Chemical Physics*
Physics—Fluids & Plasmas	6.6	7.6	*Physics of Fluids, Physics of Plasmas*
Mathematical Physics	6.5	>10	*Journal of Mathematical Physics*
Medicine, Research and Experimentation	5.4		
Cardiac & Cardio Systems	4.9		
Emergency Medicine	5.7		
Robotics	5.5		

Source: Thomson Reuters, ISI Web of Knowledge, Journal Citation Reports, Year 2010

In lieu of trying to solve the public access problem by imposing a one-size-fits-all solution with a fixed embargo length for all articles that have some component of federal funding or introducing a complicated scheme for varying embargo lengths (as necessary to address field-specific conditions), AIP proposes a simpler system that allows government to accomplish public access in a way that is not only effective, efficient, and sustainable, but also keeps the US scientific enterprise thriving as it moves into the future.

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

1. Scholarly publishers as a group have proposed modifications to their author submission software so that all journal articles written after the implementation date would include funding agency information along with the standard metadata that is already being deposited in CrossRef and other standard bibliographic databases. This new metadata, which specifically tags the funding agency(s) responsible for the research leading to the journal article, would be deposited in the CrossRef database. (The CrossRef database has been developed and

AIP

maintained by this nonprofit consortium for the past 12 years and now contains the metadata for more than 50 million scholarly articles and related content.) Funding agencies can procure a license to this database at modest cost—many already have. Such a license provides access to the article metadata, including the critical article identifier (the DOI).

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

3. Scholarly publishers have proposed and initiated pilot projects with funding agencies to link agency research reports and related content on agency sites to publisher content tagged with the same funding information, thus expanding interoperability between agency and publisher databases and access to the linked content.

AIP