

publicaccess@ostp.gov

16th December 2011

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

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

The PA welcomes this opportunity to respond to the Office of Science and Technology Policy request for information regarding public access to peer-reviewed scholarly publications resulting from federally funded research.

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

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

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

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

Question 1

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

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

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

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

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

Question 2

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

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

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

Question 3

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

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

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

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

Question 4

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

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

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

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

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

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

Question 5

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

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

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

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

Question 6

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

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

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

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

Question 7

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

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

Question 8

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

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

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

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

Ends