

## **Request for Information: Public Access to Digital Data Resulting From Federally Funded Scientific Research**

The Zoological Society of London is pleased to respond to the Office of Science and Technology Policy (OSTP) request for information concerning long-term stewardship and broad public access to peer-reviewed scholarly publications resulting from federally funded scientific research.

The Zoological Society of London was founded in 1826 and has published scholarly zoological science since 1830. ZSL publishes three peer-reviewed journals, in partnership with Wiley-Blackwell. ZSL is a charity and learned society: revenue generated from publishing contributes to the science and conservation activities undertaken by the Society. Specifically, income generated from publishing funds ZSL's annual programme of science and conservation events, including conferences, seminars and workshops, which provide opportunities for knowledge sharing, mentoring and priority setting.

We support sustainable models of access. ZSL's journals use a subscription-based model, with an option for open access, where an author (via the institution or funder) provides payment to fund publication.

We appreciate the opportunity to participate in the Administration's consultation.

**(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?**

Publishers invest significant resources to establish new markets in developing economies to extend access to research journals: exploring these new markets is part of the 'journals business'. Publishers have pioneered new subscription initiatives that optimise access to academic literature, and support programmes that provide free or very low cost access to universities, research institutes, schools, hospitals, governmental offices and national libraries in countries with the lowest gross national income per capita through the Research4Life, eIFL and PERii schemes.

The US public should have access to research data resulting from federally-funded research. Project reports and summaries are the appropriate conduit for disseminating the outcome of federally funded research. These reports should be publically available via funding organisation databases, and can be linked to grant applications and agency auditing systems.

The process of scholarly publishing adds value (through editor expertise and peer-review). The journal article is a product in its own right, and not one that should be expropriated without compensation. Free access to papers resulting from federally funded research undermines copyright, intellectual property rights, jobs and exports. The value-added activities provided by publishers are not paid for by federal agencies and they are not free. In a rapidly changing publishing environment research communication needs to be operated under strict quality controls. Considerable investment is made to develop software (i.e. to manage the peer-review process) and the costs associated with supporting editorial groups are often met by the publisher. Policies that starve the system of resources will negatively impact the quality of published journal articles. Publishers should be able to develop and use appropriate business models to recover their investment.

Policies for archiving and public access that are tied to growth in the US economy may well serve to undermine the competitive advantage of the US scientific enterprise. A free access policy will allow the benefits of federally funded scientific research to be disseminated globally, giving an advantage to economies that have not made an investment in it. It could be argued that this is a poor return for US taxpayers. How can this be managed to promote growth in the US economy?

Government agencies should work with publishers to establish the most appropriate ways of meeting access requirements. This will provide an opportunity for all parties to develop systems for information provision which preserve the integrity of the scientific publishing process and give US taxpayers a return on the investment in research.

**(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?**

Providing global public access to the peer-reviewed published output from federally funded research can expose content to piracy and other unauthorized dissemination, which would undermine the income that scholarly publishers require in order to continue their investment for the benefit of the scholarly community.

The PEER project <http://www.peerproject.eu/reports/> into the behaviour of researchers has shown that the publisher-created Version of Record (VoR) from a peer-reviewed journal is considered by researchers to be the authoritative, definitive version (over versions in subject or institutional repositories). Publishers take seriously their role as the stewards of the research literature and version control is important to the integrity of the scientific record. The intellectual property interests of the publisher who have invested in the VoR should be protected.

The most efficient way to ensure the protection of intellectual property interests of all stakeholders would be to make the final research report freely available. This would permit broad dissemination of the research results obtained from federal funding. Subsequent research articles present far more than research data: they are a synthesis of the researchers' knowledge to date, and should be protected by appropriate intellectual property rights.

**(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?**

With advances in technology, centralization is not required. Any move towards centralization would require unnecessary duplication of effort and expense. Publishers have developed tools to ensure interoperability between different access systems. One example is CrossRef <http://www.crossref.org/>, an independent membership association, founded and directed by publishers. The mandate of CrossRef is to connect users to primary research content, by enabling publishers to work collectively. CrossRef is the official DOI® link registration agency for scholarly and professional publications and its citation-linking network covers tens of millions of articles and other content items from thousands of scholarly and professional publishers.

Publishers have an excellent record of developing discipline-specific tools to meet user requirements and have invested heavily in the development of tools to achieve interoperability between different access systems. Can this responsibility plausibly be the remit of government, or a good use of government funds?

**(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?**

Proposals have been put to NSF for collaborative projects to enhance the public access, utility and preservation of publications resulting from federally funded research. These initiatives include standardizing the collection, display and use of metadata to acknowledge the federal grant supporting the research from which a scholarly publication derived and linking back to the federal agency website. Also proposed is a project to understand the requirements for and benefits derived from content mining and to establish a methodology for overcoming current barriers, so that publishers can facilitate content mining with sustainable business models.

Other public-private partnerships that take advantages of existing publisher archives include:

**Author disambiguation.** STM publishers are working to eliminate author ambiguity through the Open Researcher & Contributor ID (ORCID) project. ORCID will allow researchers to create, edit, and maintain an ORCID ID and profile free of charge. Participation in ORCID is open to any organization that has an interest in scholarly communications. All software developed by ORCID will be publicly released under an open-source software license approved by the Open Source Initiative (OSI). ORCID is governed by representatives from a broad cross-section of stakeholders including publishers, societies, libraries, and other institutions.

**Funding agency information.** Acknowledgment is often given to the research funding organisations in the published journal article but this is not standardised. Funding organisations are tasked with tracking publications resulting from their funded research. A community-wide

solution to this problem is being pursued. This proposal has been endorsed by CrossRef and major STM publishing trade associations.

**Content mining.** Content mining has the potential to be useful to the scientific community by supporting new areas of scientific discovery. The aim is for publishers, their society partners and research funders to work together to develop pilot projects for journal content mining that would identify, organize, and perform analysis to create conceptual links within and between content that are not obvious to initial human inspection.

**(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?**

Publishers are involved in a project with CrossRef and the Department of Energy (DoE) to standardize the way funding information is collected and included in article metadata. This will allow agencies to easily obtain information about publications resulting from federally funded research.

Core metadata can allow users to find related information without the requirement of accessing the full text. The Dublin Core Metadata Initiative provides key specifications and best practice regarding the use of metadata for the description of various digital resources (including books and journal articles). It facilitates interoperability of different applications and vocabularies and optimizes the metadata for searching. In addition, CrossRef provides a cross-publisher linking network, allowing readers to easily link to other resources of interest on other publisher platforms. This works seamlessly through DOIs and metadata which are embedded in articles and other content as part of the value-added publication process.

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

Authors should have the freedom to submit the outcome of their research the most appropriate journal, which has the greatest impact and relevance to their field. This results in research papers being reviewed and read by research peers and furthers the advancement of the science. Research funding organisations could provide funds to publishers to cover the fees associated with the publishing process (Gold Open Access). The publication of research reports from funding agencies would provide open access to the research without compromising the integrity of the published journal article. Funds retained in the funding agency account could be credited on a paper-by-paper basis, which would give federal agencies and taxpayers an ongoing return on their

investment. Research published long after the termination of the agency grant can still be credited to the awarding agency.

**(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 add value to all types of content they produce. This investment is of enormous benefit to community and any government policy that mandates free access compromise future investment in high-quality publication, dissemination and preservation of the research.

**(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?**

Federal agencies should not impose inappropriate embargo periods on non-federally funded businesses. Any embargo period shortens the period of copyright protection and peer-reviewed papers should not be made public within the duration of the article's copyright without the copyright holder's permission. The embargo period established for each journal is evaluated according to the needs of the market. The journal half-life is an important means to establish the length of usage of an article within a given discipline. Foreshortening the time a publisher is able to recoup their investment has the potential to seriously damage publishers and the overall economy.

Publishers have invested considerably to optimise the speed and functionality of the online publication of research articles. This digital revolution provides rapid access and complex searching, access to datasets, supplementary information, and links to cited material. Furthermore, SMT publishing benefits from publisher activities which direct readers to content, alert readers to new research and marketing initiatives, Government agencies should collaborate closely with publishers, scholarly associations and universities to achieve the full potential of publicly accessible databases. Primary data and supplementary materials are increasingly being made available to the research community and publishers are at the forefront of promoting best practice in this area, for example, by presenting and repurposing data in formats to increase cross-referencing and reuse, and by linking datasets to primary research articles.

Policies which mandate free access will hamper future investment in these areas. In a rapidly changing publishing environment, continued development in functionality of the research article is critical. A co-ordinated response to the needs of the academic community and funding bodies can be met by publishers. The aim should be to produce a climate of equitable access while protecting appropriate intellectual property rights.