

Adrian Pohl
Thu 1/12/2012 4:20 PM
RFI on Public Access to Peer-Reviewed Scholarly Publication

Dear people at the OSTP,

below are my answers to your questions on Open Access. I am responding as an individual working in an institution which provides information (research tools as well as licensed content) to academic libraries. Also, I am coordinating the Open Knowledge Foundation's "Working Group on Open Bibliographic Data". I make all my academic publications accessible on the web under a CC-BY license.

All the best
Adrian

(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?

The most efficient way to translate basic research findings into innovations that grow the economy is to allow as many innovators access to these findings as easily as possible. Thus, making scientific content openly available under open licenses such as CC-BY is exactly the kind of measure to enable and further research-driven innovation. (For a clear definition of the term "open" see <http://opendefinition.org/okd>.) The traditional entities to archive and make accessible scholarly works have always been (university) libraries. As such, federal agencies can and should support libraries to resume this task which has temporarily been outsourced to commercial publishers, which have prevented access by a subscription and copyright model which prevents innovators from accessing the latest research findings and has generated a rise in cost manifold beyond inflation as measured by the consumer price index. Thus, it is imperative to reduce the costs of public access to publicly funded research, while allowing as many innovators access to research findings as possible. Diverting those funds who are currently being accumulated with the shareholders of commercial publishers, towards libraries will cut publishing costs by orders of magnitude (through eliminating many middle-men) and enlarge the circle of potential innovators by orders of magnitudes, thus potentiating the current cost/benefit ratio exponentially.

(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?

I would dispute any actual intellectual property rights of publishers over publicly funded research. I am aware that legally there are such rights as researchers most often hand over their copyrights to these

publishers. However, since commercial publishers do usually add little to no value to the published research results (even peer-review is performed pro-bono by researchers themselves), this practice needs to end. Publicly funded research has been bought by the public and belongs to the public. Ensuring open publication licenses such as CC-BY for literature about publicly funded research is one of the possibilities to ensure the public retains its intellectual property on the research it funded.

(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?

Multiple private sources are always a suboptimal choice for long-term archiving: few private entities survive 'long-term'. Some libraries, in contrast, have been around for centuries, significantly longer than most private entities. Centralized access, however, suffers from as many cons as any monopoly. Ensuring the libraries of every research institution are sufficiently equipped to maintain long-term archiving of scholarly literature allows for a federated, decentralized archive of scholarly literature beyond any short-term financial fluctuations and allows for international collaboration for maximum safety through world-wide redundancy (following the motto "Lots of copies keep stuff safe"). A small fraction of the current subscription costs paid for by libraries would ensure such a long-term, publicly accessible archive.

(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?

Not to my knowledge. Given the instability of private sources and the decades-long history of price-gouging, I would argue that this would not be a good idea, either.

(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?

What is required is an evolving metadata and protocol standard that grows with the scientific enterprise and is under the control of scientists and developed in close cooperation with librarians. A recent innovation in this direction is BibJSON (<<http://bibserver.okfn.org/bibjson/>>). At best, this standard would build on Linked Open Data technologies, as this technology for exposing (meta)data on the web allows interlinking and thus makes the aggregation of citation data and indication of publication usage in university courses and analysis of this data technically easy. The actual location of the publications is, of course, irrelevant, as long as proper long-term archiving is ensured (see above). It is crucial that metadata for all scientific publications is made fully accessible on the web under an open license. See the Principles on Open Bibliographic Data for more detail: <<http://openbiblio.net/principles/>>.

(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?

The benefit is maximized by minimizing the costs associated with access. The costs are minimized by preventing third parties from adding costs to the process. One way to establish a short and thus cheap supply line is to have scholars deposit their work directly at their libraries, avoiding the costs of intermediaries such as publishers. The process of this deposition would still be identical to the current process (i.e., peer-review), albeit without intervening entities which withdraw funds but add little to no value.

(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?

Most definitely, yes, all of them. The public bought them and thus owns them. In some ways, scholarly work is nothing but commissioned by the public.

(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?

As the public already own these scholarly works, it is difficult to understand why there should be any embargo period to allow private entities, which have not contributed to the work, to profit from it. Every scholarly work that has been paid for by the public should be available to the public for immediate re-use and enter into the economy. It is hard to understand why there should be a waiting period for innovation to enter the market.