January 12, 2012

To: Office of Science and Technology Policy, Executive Office of the President
From: Ginger Strader, Director, Smithsonian Institution Scholarly Press
Re: Response to OSTP RFI on Public Access

In response to the Office of Science and Technology Policy request for information on "Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research," I offer a compilation of opinions from individual staff members. These responses do not represent official responses or the formal position of the Smithsonian Institution.

The Smithsonian is not an executive branch agency, and many of its publications result from both federal and private funding and incorporate the work product of federal and non-federal employees. As a result of its unique classification, when considering the question of public access, the Smithsonian is mindful of its obligation to protect intellectual property rights of authors and other contributors when publications are not in the public domain.

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?

- Include funds in all research grants that will support reasonable authors' fees for online publication in open access journals (or equivalents for different types of publications). If the grant is the only source of funding, grantees could be required to publish in open access venues.

- Continue to fund conversion and tagging of metadata and content so it may be harvested by existing, online databases and portals (e.g., Encyclopedia of Life).

- Researchers should take advantage of open repository options that may exist within their library systems (e.g., National Agriculture Library, Smithsonian Libraries) that offer archiving options for digitized publications. By exposing these repositories to standard harvesting tools and methods (e.g., OAI-PMH, Google Scholar), the publications will have higher visibility.

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?

- Encourage conspicuous identification of copyrighted elements.
- Do not require agencies to identify all content with general statements such as "government publication," which may inadvertently mislead users to believe the entire publication is public domain and may lead to unpermitted reproduction and /or sale of intellectual property.
- Allow agencies to impose embargo periods on results prior to publication, during which grantees may protect their intellectual property.
- Restrict or limit assignment of rights to publishers that did not contribute to the creation of the IP, only to production and dissemination of the final product. Impose time limits to be applied to publishers' exclusive rights to disseminate, so publicly funded research results can be re-issued in an open access journal after first publication.

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?

- Centralization of published scientific findings would allow an extension of simple search and download functionality. For example it would allow natural language processing and text mining software to inform research and guide discoveries that scientists cannot achieve by simple, linear reading of publications. This could be used to better inform hypothesis creation and to prevent duplication of effort.
- Encourage decentralized but open and standard metadata that can be aggregated into larger tools (e.g., the defunct National Science Digital Library).
- Unless a publisher distributes through an open access platform, restrict or limit assignment of rights to the publisher to ensure non-exclusivity, with a clause in publishing agreements for federally funded research that allows authors and/or funding agencies to reproduce and disseminate the same research independently through its own open access outlets.

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?

- In addition to Google Scholar, a good model that provides storage as well as access and social media activity is Mendeley (www.mendeley.com).
- Subscription models such as JStor.

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?

- Metadata: OAI-PMH, OAI-ORE, ONIX, Dublin Core, MODS, BibTeX, RIS, OPDS.

- Market forces seem to be taking care of this, led by open access journals, which are making content discoverable. In response, research communities are building systems for mining and processing that content. Some basic research on this area of information science is being funded. Federal agencies may want to follow suit.

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?

- Book chapters and essays in conference proceedings that are fully federally funded and, in isolation, are in the public domain should be included.

- Primary datasets should be included.

--
Ginger Strader
Director, Smithsonian Institution Scholarly Press
P.O. Box 37012, MRC 957
Washington, D.C. 20013-7012