

The following comments were submitted to the White House Office of Science and Technology Policy on December 28, 2011 in response to the RFI on public-access to peer-reviewed scholarly publications resulting from federally funded research. (<http://federalregister.gov/a/2011-32943>) The comments were written and reviewed by members of the USAIN Executive Council and the USAIN Legislative and Government Relations Committee.

Constance J. Britton / britton.4@osu.edu
United States Agricultural Information Network (USAIN)
Wooster, OH

I am writing on behalf of the United States Agricultural Information Network (USAIN) to respectfully respond to the Request for Information for recommendations related to public access to peer-reviewed scholarly publications resulting from federally funded research.

USAIN (usain.org) is an organization of over 150 agricultural information professionals that provides a forum for discussion of agricultural issues, takes a leadership role in the formation of a national information policy as related to agriculture, makes recommendations to the National Agricultural Library (NAL) on agricultural information matters, and promotes collaboration and communication among its members. USAIN has testified before Congress, played an advisory role in the National Agricultural Text Digitizing Project, written a national agricultural literature preservation plan, served on blue ribbon panels to review NAL services, and participated in the selection process for new NAL Directors. Our members are skilled librarians with knowledge of the modern theories, principles, practices, techniques, and policy issues pertinent to the current practice of librarianship and information science. Many of our members work at Land Grant institutions with extensive federally-funded research programs and are experienced in acquiring, organizing, and preserving scientific and agricultural data. The USAIN Executive Council is privileged to provide the following input related to this important topic of public access to information.

Comment 1. Access to published research information is a fundamental requirement by scientists, students, innovators, entrepreneurs, and other interested citizens in order to further the process of the discovery of knowledge ultimately leading to the creation of new products, jobs, and the growth of the economy. Published research information, created at the public's expense, but that sits behind a subscription pay-wall, may not be equitably available to all who might benefit from its content. Even funded researchers at public universities may have limited access to published literature due to the lack of a library or personal subscription. A system that requires that the publication output of federally-funded research be deposited in one or more publicly-accessible repositories would achieve a necessary level of equitable access.

Comment 2. Nothing in a public-access system should threaten the protection of intellectual property as covered by copyright. In fact, greater access to research information and data will ensure greater visibility and recognition of an author's intellectual achievements.

Comment 3. The national libraries are the obvious leaders in creating a system that supports public access to federally-funded research output. Libraries currently possess the knowledge, expertise, if not the resources, to develop and implement a repository system. In the case of research funded by the United States Department of Agriculture, the National Agricultural Library (NAL) has a key role to play in the acquisition and preservation of research publications and data resulting from USDA funding. NAL librarians and staff possess the knowledge, skills and expertise to manage paper and digital collections,

and by applying metadata, to facilitate the discovery of stored content. The NAL is a demonstrated leader in the development of the metadata and the technical standards needed to insure interoperability of these systems. However, in creating a repository system, it is essential that new funding be appropriated to engage sufficient staff and provide the technical infrastructure necessary to insure success of this program. The best model of a centralized approach is that provided by the NIH-mandated deposit of peer-reviewed research articles in PubMed Central, where the deposit requirement and sufficient program funding have made this repository successful. A comparable repository for USDA-funded research could be managed by the NAL as an expansion of the existing NAL Digital Collections (NALDC). The advantages of a centralized repository are better control of the deposit process, author compliance, and consistent metadata applications. Funding agencies managing a smaller grant portfolio may have a more difficult time supporting a separate repository, so centralization would benefit these agencies. Centralization also minimizes issues of interoperability, consistency and redundancy. Many universities maintain an institutional repository and could help facilitate required deposits within the institutional site or a centralized repository. Even with clearly articulated standards, achieving full interoperability across many repositories may be a challenging goal.

Comment 4. No comment.

Comment 5. Again, the national libraries have the requisite skills, experience and mandate to define and implement the standards that must be put in place to create an interoperable repository system. The minimum metadata elements for describing bibliographic information are currently well-defined by the Dublin Core metadata standard. These elements can be readily derived from publisher data and incorporated as part of the deposit. Adherence to this standard will facilitate the sharing of data from multiple repositories and lead to discovery by the public. Metadata standards are critical for describing publications and data within a repository, but institutions are also faced with the added challenge of increasing access to those resources. Resources must be highly discoverable and understood within a larger context of scientific data and research. For that to happen, several things must occur: 1) the advanced support of author disambiguation initiatives, such as ORCID, which "aims to solve the author/contributor name ambiguity problem in scholarly communications" 2) a general mandate requiring federally funded authors to identify their funding source when submitting publications to a repository and 3) the development and support of Semantic Web technologies that allow for the re-purposing, reuse, and analysis of publication and other data. By design, Semantic Web technologies are machine-readable; continuing to encourage the development and accessibility of these technologies would allow for flexible re-purposing of data, regardless of the model - centralized, decentralized, or mixed-model - chosen by Federal agencies.

Comment 6. The richest benefit with the least burden will be gained by utilizing the expertise that resides with the national library system, including not only our national libraries, but also the university libraries. Libraries already possess extensive knowledge about the development and management of publication repositories. Key to this approach, however, is sufficient funding to implement this program.

Comment 7. While, ideally, the public would have access to all information created as a result of federally-funded research, the peer-reviewed journal articles are the critical body of work to be captured in a repository system. This literature represents the publishable results of the funded research, whereas conference proceedings often report preliminary results and book chapters may synthesize research from multiple sources and often come much later in the scholarly communication process. The journal literature is the most useful report of research outcomes to make available to the public.

Comment 8. No comment.