

Developments on the Creation of a National Repository for Publications and Data Arising Directly from Federal Research Funding

On February 22, 2013, in an action of great potential significance for the research community, John P. Holdren, Director of the White House's Office of Science and Technology Policy, directed that within six months each federal research funding agency with an R&D budget of \$100 million or more "**...develop a plan to support increased public access to the results of research funded by the Federal Government,**" including "**peer-reviewed publications and digital data.**"

Publication Deposit: The OSTP directive essentially spreads to most federal funding agencies the practice begun in 2005 by the National Institutes of Health. NIH-funded researchers are required to place the final edited version of their peer-reviewed papers in NIH's PubMed Central for long-term preservation and access by all. The OSTP directive differs from NIH's policy in that it permits agencies to propose different methods for storage and access other than the central repository chosen by NIH.

Data Deposit: Both NIH and NSF have required principal investigators to file data management and preservation plans, but there are no central federal agency repositories for data and no established universal standards for data. The directive's requirement is that "**To the extent feasible and consistent with applicable law and policy; agency mission; resource constraints; U.S. national, homeland, and economic security; and the objectives listed below, digitally formatted scientific data resulting from unclassified research supported wholly or in part by Federal funding should be stored and publicly accessible to search, retrieve, and analyze.**"

APLU, AAU and ARL Discussions: Because both elements of the OSTP directive have consequences for research, APLU in conjunction with AAU and the Association of Research Libraries (ARL), convened a small group of provosts, research officers and librarians with special expertise from member campuses to discuss various approaches to the directive's publications and data elements. While the discussions entailed many nuances, clear preferences emerged on how publication and data access and preservation were to be accomplished.

On publications, the participants agreed that whether the repository was centralized (like PubMed Central) or distributed (across university libraries, journal publishers or scholarly societies) was not the important issue. The important issue is that the publications be preserved, that free access be permitted to the public, industry and science and that access by humans and harvesting software be permitted.

On data, the consensus was that preservation and sharing of research data was a worthy goal, but because of the varied nature of experiments and the data gathered in different fields, and the metadata necessary to interpret correctly data gathered in a range of representations, commonly accepted standards may not yet exist in all areas covered by the directive. We concluded that a taskforce be empanelled to give operational meaning to this definition and to address the challenges and resolve the issues OSTP enumerated in the directive. The membership of the taskforce should include representatives from research universities, scholarly societies and federal funding agencies.

National Academy Conference: A conference sponsored by the funding agencies subject to the OSTP directive was held at the National Academy of Sciences on May 14 to 17 to permit interested parties to inform the agencies of their preferences concerning the implementation of the directive. The Academy has placed the presentations on their website. Briefly, on the issue of publications there were significant differences among the participants on: whether a repository should be centralized or distributed; the freedom of researchers to access the repository; and on the appropriate length of embargo periods. The data discussion revealed much support for preservation and sharing, but recognition that many matters must be discussed and agreements reached on them before data can usefully be placed in a functioning depository. There was general support for expanding the requirement for the inclusion of data management plans in all research proposals to funding agencies.

Funding agencies are to submit to OSTP their plans to comply with the directive by August 22. The plans will not be made public at that time.

The Emerging University Institutional Repository Option: Most U.S. research university libraries have institutional repositories suitable for deposit and subsequent access of publications arising from federal research funding. Our

library community is now engaged in an implementation study to determine how these repositories might be made sufficiently interoperable such that they could function as a distributed repository that would satisfy the OSTP directive. This option will be presented to the funding agencies for consideration as they develop their responses to OSTP. The research library community intends to develop such an interoperable system in any case.

Publishers have already presented to the agencies an option for a repository distributed across various publishers' websites and some agencies are apparently considering development of centralized repositories of their own.

APLU, AAU and ARL remain fully involved in this important discussion.