



This document is scheduled to be published in the Federal Register on 12/23/2011 and available online at <http://federalregister.gov/a/2011-32947>, and on FDsys.gov

OFFICE OF SCIENCE AND TECHNOLOGY POLICY
REQUEST FOR INFORMATION: PUBLIC ACCESS TO DIGITAL DATA
RESULTING FROM FEDERALLY FUNDED SCIENTIFIC RESEARCH

ACTION: Notice of Request for Information (RFI)

SUMMARY:

The RFI is being extended to change the response date to January 12, 2012. The RFI was published in the Federal Register, Volume 76, Number 214, on November 4, 2011, pages 68517-68518.

In accordance with Section 103(b)(6) of the America COMPETES Reauthorization Act of 2010 (ACRA; Public Law 111-358), this Request for Information (RFI) offers the opportunity for interested individuals and organizations to provide recommendations on approaches for ensuring long-term stewardship and encouraging broad public access to unclassified digital data that result from federally funded scientific research. The public input provided through this Notice will inform deliberations of the National Science and Technology Council's Interagency Working Group on Digital Data.

RELEASE DATE: November 3, 2011

RESPONSE DATE: January 12, 2012

ADDRESS: digitaldata@ostp.gov

ISSUED BY: Office of Science and Technology Policy (OSTP) on behalf of the National Science and Technology Council (NSTC)

SUPPLEMENTARY INFORMATION:

Purpose

In accordance with Section 103(b)(6) of the America COMPETES Reauthorization Act of 2010 (ACRA; Public Law 111-358), this Request for Information (RFI) offers the opportunity for interested individuals and organizations to provide recommendations on approaches for ensuring long-term stewardship and encouraging broad public access to unclassified digital data that result from federally funded scientific research. The public input provided through this Notice will inform deliberations of the National Science and Technology Council's Interagency Working Group on Digital Data.

Background

The multi-agency Interagency Working Group on Digital Data (Working Group), established under the National Science and Technology Council (NSTC) Committee on Science (CoS), has been tasked with developing options for implementing the digital data policy and standards requirements of Section 103 of ACRA. OSTP will issue a report to Congress, in accordance with Section 103(e) of ACRA, describing priorities for the development of agency policies for ensuring broad public access to the results of federally funded unclassified research, the status of agency policies for public access to digital data resulting from federally funded research, and a summary of public input collected from this RFI and other mechanisms. The Working Group is considering steps that can be taken by Federal agencies to encourage and coordinate the development of agency policies and standards to promote long-term preservation of and access to digital data resulting from federally funded scientific research. Ideally, such policies would harmonize, to the extent practicable and feasible, data management plans for digital data that are collected or otherwise produced either by the agency itself or extramurally with Federal funds. The 2009 report of the Interagency Working Group on Digital Data of the National Science and

Technology Council, “Harnessing the Power of Digital Data,” recommended that agencies lay the foundations for digital scientific data policy and make their policies publicly available. It also recommended that agencies consider requiring data management plans for projects that will generate “preservation data” – those data for which the benefits of preservation exceed the costs. Federal science agencies already have some experience with policies to promote long-term preservation and access to scientific data. Indeed current Federal policies promote and in many cases require Federal agencies to make the digital data generated by Federal agencies more publically accessible. However, such policies do not routinely cover data generated through Federal grants, cooperative agreements, and some other types of funding mechanism. Exceptions include, the National Institutes of Health’s (NIH) Data Sharing Policy, which requires all investigator-initiated applications with direct costs greater than \$500,000 in any single year provide a data management plan. In addition, NIH has more specific data management and data sharing requirements for specific types of projects, such as genome-wide association studies.

In January 2011, the National Science Foundation (NSF) reaffirmed its data management policy requirement, indicating that proposals must include a Data Management Plan that describes how funded researchers will conform to NSF policy on the dissemination and sharing of research results. The NSF policy is clear that “Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants.” Such models may not necessarily be appropriate for all types of federally sponsored research.

As agencies consider how to further develop digital data policies, it is important to note that all policies for increasing accountability and access to digital data must follow statutory requirements and follow best practices for protecting confidentiality, personal privacy, proprietary interests, intellectual property rights, author attribution, and for ensuring that homeland and national security interests are not compromised.

The Working Group is now seeking additional insight from “non-Federal stakeholders, including the public, universities, nonprofit and for-profit publishers, libraries, federally funded and non-federally funded research scientists, and other organizations and institutions with an interest in long-term stewardship and improved public access to the results of federally funded research,” as described in Section 103(b)(6) of ACRA. Specifically the Working Group seeks further public comment on the questions listed below:

Preservation, discoverability, and access

- (1) What specific Federal policies would encourage public access to and the preservation of broadly valuable digital data resulting from federally funded scientific research, to grow the U.S. economy and improve the productivity of the American scientific enterprise?
- (2) What specific steps can be taken to protect the intellectual property interests of publishers, scientists, Federal agencies, and other stakeholders, with respect to any existing or proposed policies for encouraging public access to and preservation of digital data resulting from federally funded scientific research?
- (3) How could Federal agencies take into account inherent differences between scientific disciplines and different types of digital data when developing policies on the management of data?

- (4) How could agency policies consider differences in the relative costs and benefits of long-term stewardship and dissemination of different types of data resulting from federally funded research?
- (5) How can stakeholders (e.g., research communities, universities, research institutions, libraries, scientific publishers) best contribute to the implementation of data management plans?
- (6) How could funding mechanisms be improved to better address the real costs of preserving and making digital data accessible?
- (7) What approaches could agencies take to measure, verify, and improve compliance with Federal data stewardship and access policies for scientific research? How can the burden of compliance and verification be minimized?
- (8) What additional steps could agencies take to stimulate innovative use of publicly accessible research data in new and existing markets and industries to create jobs and grow the economy?
- (9) What mechanisms could be developed to assure that those who produced the data are given appropriate attribution and credit when secondary results are reported?

Standards for interoperability, re-use and re-purposing

- (10) What digital data standards would enable interoperability, reuse, and repurposing of digital scientific data? For example, MIAME (minimum information about a microarray experiment; see Brazma et al., 2001, Nature Genetics 29, 371) is an example of a community-driven data standards effort.
- (11) What are other examples of standards development processes that were successful in producing effective standards and what characteristics of the process made these efforts successful?

(12) How could Federal agencies promote effective coordination on digital data standards with other nations and international communities?

(13) What policies, practices, and standards are needed to support linking between publications and associated data?

Response to this RFI is voluntary. Responders are free to address any or all the above items, as well as provide additional information that they think is relevant to developing policies consistent with increased preservation and dissemination of broadly useful digital data resulting from federally funded research. Please note that the Government will not pay for response preparation or for the use of any information contained in the response.

How to Submit a Response

All comments must be submitted electronically to: digitaldata@ostp.gov

Responses to this RFI will be accepted through January 12, 2012. You will receive an electronic confirmation acknowledging receipt of your response, but will not receive individualized feedback on any suggestions. No basis for claims against the U.S. Government shall arise as a result of a response to this request for information or from the Government's use of such information.

Inquiries

Specific questions about this RFI should be directed to the following e-mail address: digitaldata@ostp.gov.

Form should include:

[Assigned ID #]

[Assigned Entry date]

Name / Email

Affiliation/Organization

City, State

Comment 1

Comment 2

Comment 3

Comment 4

Comment 5

Comment 6

Comment 7

Comment 8

Comment 9

Comment 10

Comment 11

In addition, please identify any other items the Working Group might consider for Federal policies related to public access to peer-reviewed scholarly publications resulting from federally supported research.

Please attach any documents that support your comments to the questions.

Ted Wackler
Deputy Chief of Staff
Billing Code – 3710-W12

