PRINCIPLES FOR PROMOTING ACCESS TO FEDERAL GOVERNMENT-SUPPORTED SCIENTIFIC DATA AND RESEARCH FINDINGS THROUGH INTERNATIONAL SCIENTIFIC COOPERATION

PRODUCT OF THE
INTERAGENCY WORKING GROUP ON OPEN DATA SHARING POLICY
OF THE SUBCOMMITTEE ON INTERNATIONAL ISSUES
OF THE COMMITTEE ON SCIENCE
OF THE NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

December 2016

About the National Science and Technology Council

The National Science and Technology Council (NSTC) is the principal means by which the Executive Branch coordinates science and technology policy across the diverse entities that make up the Federal research and development (R&D) enterprise. One of the NSTC's primary objectives is establishing clear national goals for Federal science and technology investments. The NSTC prepares R&D packages aimed at accomplishing multiple national goals. The NSTC's work is organized under five committees: Environment, Natural Resources, and Sustainability; Homeland and National Security; Science, Technology, Engineering, and Mathematics (STEM) Education; Science; and Technology. Each of these committees oversees subcommittees and working groups that are focused on different aspects of science and technology. More information is available at www.whitehouse.gov/ostp/nstc.

About the Committees of the NSTC

The purpose of the Committees is to advise and assist the NSTC, under Executive Order 12881, to increase strategic impact, overall effectiveness, and productivity of Federal science and technology (S&T) activities. The Committee on Science (CoS) focuses on Federally supported efforts that develop new knowledge in the sciences, mathematics, and engineering (not including those areas primarily related to the environment and natural resources), and the Committee on Homeland and National Security (CHNS) covers Federal research and development efforts in areas of S&T related to homeland and national security. Both Committees address significant national and international policy, program, and budget matters that cut across agency boundaries and provide a formal mechanism to promote interagency S&T policy development, coordination, collaboration, and information exchange.

About the Subcommittee on International Issues

The purpose of the Subcommittee on International Issues (SII) is to enhance coordination of Federal agencies' international science and technology (S&T) cooperation and partnerships. The SII addresses long-term strategic engagement goals, policy issues related to high-value international collaboration, and short-term country- and issue-specific priorities. The Subcommittee on International Issues was initially charted as the Topics in International Science, Technology, and Innovation (TISTI) under the CHNS.

About the Office of Science and Technology Policy

The Office of Science and Technology Policy (OSTP) was established by the National Science and Technology Policy, Organization, and Priorities Act of 1976. OSTP's responsibilities include advising the President in policy formulation and budget development on questions in which science and technology are important elements; articulating the President's science and technology policy and programs; and fostering strong partnerships among Federal, state, and local governments, and the scientific communities in industry and academia. The Director of OSTP also serves as Assistant to the President for Science and Technology and manages the NSTC. More information is available at www.whitehouse.gov/ostp.

Copyright Information

Printed in the United States of America, December 2016.

Report prepared by

NATIONAL SCIENCE AND TECHNOLOGY COUNCIL
COMMITTEE ON SCIENCE
SUBCOMMITEE ON INTERNATIONAL ISSUES
INTERAGENCY WORKING GROUP ON OPEN DATA SHARING POLICY

National Science and Technology Council

Chair
John P. Holdren
Assistant to the President for Science and Technology and Director, Office of Science and Technology Policy

Staff
Afua Bruce
Executive Director
Office of Science and Technology Policy

Committee on Science

Chairs
Francis Collins
Director
National Institutes of Health

France Córdova
Director
National Science Foundation

Jo Handelsman
Associate Director for Science
Office of Science and Technology Policy

Staff
Sarah Mazur
Executive Secretary
U.S. Environmental Protection Agency

Office of Science and Technology Policy Staff

Mahlet Mesfin
Assistant Director, International Science and Technology

Cori Zarek
Deputy Chief Technology Officer

Jerry Sheehan
Assistant Director, Scientific Data and Information

Interagency Working Group on Open Data Sharing Policy

Mark Paese, Co-Chair
Deputy Assistant Administrator for Satellite and Information Services
National Oceanic and Atmospheric Administration

Jaime Adams, Co-Chair
Senior Advisor for International Affairs, Office of the Chief Scientist
U.S. Department of Agriculture

Interagency Working Group Members and Departments and Agencies Contributing to the Report:

Departments
Department of Agriculture
Department of Defense
Department of Energy
Department of Health and Human Services
Department of State

Agencies
National Aeronautics and Space Administration
National Institute of Standards and Technology
National Oceanic and Atmospheric Administration
National Science Foundation
U.S Agency for International Development
U.S. Patent and Trademark Office

Table of Contents

Executive Summary

Increasing access to scientific data and research findings generated by Federal agencies or resulting from Federally funded research is a U.S. policy priority. The Interagency Working Group on Open Data Sharing Policy (IWGODSP) encourages Federal departments and agencies ("Federal agencies") to rely on the following principles when promoting U.S. policies and positions on open data, data sharing, and increased access to scientific data in the context of U.S. international science and technology cooperation. These principles apply to unclassified scientific data generated by Federal agencies or resulting from Federally funded research ("government-supported scientific data"), which can include primary data (e.g., observations and measurements) and derived data (resulting from computations performed on the primary data), together with accompanying metadata that exist in digital form, as well as publications that result from Federally funded scientific research. As appropriate, the seven principles that may be used are as follows:

1. Scientific progress and collaboration benefit from an early and continuing commitment to the establishment, description, curation, maintenance, validation, discoverability, accessibility, and distribution of scientific data.
2. Scientific data should be made openly accessible to the extent permitted by law and subject to privacy, confidentiality, security, and other appropriate restrictions (e.g., recognizing proprietary interests, business confidential information, and intellectual property rights).
3. Government-supported scientific data should be available without charge whenever feasible.
4. Partners in international science and technology cooperative activities should establish a data management plan at project initiation that considers the full data lifecycle for scientific data.
5. Federal agencies should encourage technical and legal interoperability to facilitate international sharing of government-supported scientific data, using compatible, publicly available and open source formats.
6. Government-supported scientific data (and publications) should be made available as early as possible, with the timing of release and the duration of any exclusive-use period explicitly defined.
7. Federal agencies should work with international science and technology partners to adopt policies and data standards that encourage open sharing of data for collaborative activities.

Introduction

The Interagency Working Group on Open Data Sharing Policy (IWGODSP) of the Subcommittee on International Issues, Committee on Science, National Science and Technology Council recommends seven principles to inform and improve consistency among Federal departments and agencies ("Federal agencies") on open scientific data sharing positions to support scientific cooperation between the United States and international partners. These principles do not establish new policy positions, but seek to summarize principles that have been articulated in U.S. Government policy documents on data sharing.

The U.S. Government recognizes that openly-accessible data contribute to scientific progress and international collaboration, and are a catalyst for innovation. The principles set forth in this report demonstrate the United States' commitment to increasing access to unclassified scientific data generated by Federal agencies or resulting from Federally funded research ("government-supported scientific data") to promote science and technology and to further international cooperation in science and technology to address global challenges.

The seven principles summarize policy positions regarding open data that may be used by Federal agencies to advance scientific cooperation with international partners, although additional national laws, regulations, and policies may apply. For the purposes of these principles, scientific data may include primary data (e.g., observations and measurements) and derived data (resulting from computations performed on the primary data), together with accompanying metadata, that exist in a digital form. Where applicable, these principles also summarize how U.S. policies apply to publications that result from Federally funded scientific research.

Methodology and Limitations

To develop the principles, the IWGODSP conducted an inventory and review of relevant existing U.S. government and international and intergovernmental data sharing policies. This review identified common statements of policy, which are the basis for the principles in this report. These principles are consistent with, but not limited, to the:

- Office of Management and Budget's (OMB) *Open Government Directive* (2009),[1]
- Open Data Executive Order (2013),[2]
- OMB Memorandum M-13-13, *Open Data Policy – Managing Information as an Asset* (2013),[3]
- Project Open Data,[4]
- OMB Circular A-130 (revised 2016),[5]
- OSTP Memorandum - *Increasing Access to the Results of Federally Funded Scientific Research* (2013),[6]
- *U.S. Open Data Action Plan* (2014),[7]
- Bromley Principles (1991),[8] and
- Agency-specific open government plans and public access plans.[9]

This report reinforces the principles adopted by the broader international scientific community including, but not limited to, the G8 Open Data Charter;[10] the Organization for Economic Cooperation and Development Principles and Guidelines for Access to Research Data from Public Funding;[11] and the Group on Earth Observation System of Systems Data Sharing Principles.[12]

This report is written to be consistent with existing requirements in the Paperwork Reduction Act, the E-Government Act of 2002, the Privacy Act of 1974, the Federal Information Security Management Act of 2002 (FISMA), the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA), the Freedom of Information Act, the Information Quality Act, the Federal Records Act, the Bayh-Dole Act (Patent and Trademark Amendments Act of 1980), and existing OMB and OSTP guidance. Nothing in this report shall be construed to reduce the protection of information whose release would threaten national security, invade personal privacy, breach confidentiality or contractual terms, violate any other statutory requirement, or have significant negative impact on intellectual property rights or other compelling interests.

Principles

The following principles derive from U.S. Government policies that promote access to government supported scientific data and research findings:

1. **Scientific progress and collaboration benefit from an early and continuing commitment to the establishment, description, curation, maintenance, validation, discoverability, accessibility, and distribution of scientific data.**[6,8,11]

Science is increasingly collaborative and multi-disciplinary, and openly accessible scientific data can be powerful catalysts for international collaboration and the advancement of scientific research and innovation. Proper stewardship of scientific data and research findings requires effective frameworks to maximize their potential. Partners to international science and technology cooperative activities should agree prior to the beginning of the formal research cycle on policies and strategies to facilitate the management, preservation, and sharing of scientific data resulting from such activities. Doing so early is beneficial because data will be produced much earlier than formal publications. Research partners should employ appropriate data management strategies across the lifecycle of the scientific data, and should consider which data should be preserved and made accessible considering the relative value and costs associated with different types of data and long-term data preservation.

2. **Scientific data should be made openly accessible to the extent permitted by law and subject to privacy, confidentiality, security, and other appropriate restrictions (e.g., recognizing proprietary interests, business confidential information, and intellectual property rights).**[1,2,3,5,6,8]

Federal agencies should adopt a presumption in favor of openness with the fewest constraints possible while protecting privacy, confidentiality, and national security,[5,6] and recognizing proprietary interests, business confidential information, and intellectual property rights that may be associated with scientific data. The U.S. Government's open data policy promotes access to scientific government data developed or managed by Federal agencies.[1,2,3,6,13]

The U.S. Government's public access policy also calls for agencies to maximize public access to scientific data resulting from Federal R&D funding.[6]

3. **Government-supported scientific data should be made available without charge whenever feasible.**[3,5,6,8,10,12,14]

OSTP's February 22, 2013 *Memorandum on Increasing Access to the Results of Federally Funded Research*[6] directs Federal agencies to maximize access, by the general public and without charge, to digitally formatted scientific data created with Federal funds. OMB's *Open Government Directive*[1] instructs Federal agencies to publish such information online and make it freely accessible. In cases in which scientific

data can be made available but not without charge, charges should not, in general, exceed the costs associated with distributing the data to the public. To the extent permitted by law, open data formats should be used that are non-proprietary and publicly available, with minimal necessary restrictions placed upon their use.

For publications, OSTP's February 22, 2013 *Memorandum on Increasing Access to the Results of Federally Funded Research* [6] directs Federal agencies to develop policies that ensure that a version (i.e., the published version or the final accepted manuscript) of any peer-reviewed scholarly publication resulting from Federally funded research is freely available for the public to read, download, and analyze in digital form not later than one year after publication, and that publication metadata is freely available upon first publication.

4. **Partners in international science and technology cooperative activities should establish a data management plan at project initiation that considers the full data lifecycle for scientific data.**[2,3,5,6,9]

As stewards of public information, Federal agencies should work with international partners to ensure management of scientific data throughout the scientific data's lifecycle from creation or collection, through processing, dissemination, use, storage, and disposition. This includes determining which scientific data should be preserved and made accessible, and ensuring scientific data are sufficiently described to enable their use.

Describing data fully ensures that users can discover the data and understand its strengths, weaknesses, analytical limitations, and security requirements. Best practices for describing data include the use of robust, granular metadata, thorough documentation of data elements, data dictionaries, and, if applicable, additional descriptions of the purpose of the collection, the population of interest, the characteristics of the sample, and the method of data collection. Scientific data should be released at a level of granularity that maximizes their utility as permitted by law and other requirements.

In addition, for data resulting from government-funded research, Federal agencies must ensure that researchers develop a data management plan describing how the data will be shared and preserved, or provide a justification for why the data will not be shared or preserved. Data management plans must preserve the balance between the relative value of long-term preservation and access and the associated cost and administrative burdens. Agencies must have procedures in place to evaluate the merits of the data management plan and must allow the inclusion of appropriate costs for data management and access in proposals for Federal funding for scientific research.[5]

5. **Federal agencies should encourage technical and legal interoperability to facilitate international sharing of government-supported scientific data, using compatible, publicly available and open source formats.**[1,2,3,5,6,8]

Use and reuse of data and metadata are fostered when the data are accessible, discoverable, and interoperable. Technical interoperability refers to the data and metadata standards and data architecture that enable seamless discovery, sharing, and use of data across various systems and boundaries. Technical interoperability allows users to better find data, compare data, merge and combine data from different sources, and enhance the value of the data. Machine-readable (i.e., data and metadata reasonably structured to allow automated processing) and open formats, data standards, and use of common core metadata all enhance technical interoperability. Data and metadata made available in convenient, modifiable, and open formats can be more easily retrieved, downloaded, indexed, and searched, ensuring that the data are available to the widest range of users for the widest range of purposes.

In addition to technical interoperability, legal interoperability enhances sharing of data. Legal interoperability means that data among multiple datasets and from different sources can lawfully be used, combined, and enhanced without restrictions and without seeking additional authorization from the data creator on a case-by-case basis. Federal agencies should work with international partners to ensure international legal interoperability of the scientific data. Federal agencies can facilitate international sharing of scientific data by expressly stating that they are available for use globally through mechanisms such as worldwide public domain dedications and permissive open licenses, as appropriate, if there are any proprietary or intellectual property rights associated with the data.

6. **Government-supported scientific data (and publications) should be made accessible as early as possible, with the timing of release and duration of any exclusive-use period explicitly defined.**[3,6,8,10,11]

The purpose of this principle is to promote timely access to scientific data, recognizing that scientific data may not be fully documented or very useful during initial data collection. Agencies should establish when datasets are best made available. Federal agencies may allow exclusive use of the data for a reasonable period of time. After any period of exclusivity ends, the data should be made openly accessible to the extent permitted by law and subject to privacy, confidentiality, security, and other valid legal restrictions.

In addition, as specified in OSTP's February 22, 2013 *Memorandum on Increasing Access to the Results of Federally Funded Research*[6]*,* agencies must develop policies that ensure that a version (i.e., the published version or the final accepted manuscript) of a peer-reviewed scholarly publication resulting from Federally funded research is publicly accessible within one year of publication. Agencies must also provide a mechanism for stakeholders to petition to change the embargo period for a specific field by demonstrating that the existing embargo period is inconsistent with the objectives of the memorandum.

Agencies should encourage international partners to agree to make publicly available scientific data and published research findings from cooperative activities. The funding agencies should agree on when the data should be made publicly available as well as the duration of any exclusive use period.

7. **Federal agencies should work with international science and technology partners to adopt policies and data standards that encourage open sharing of data for collaborative activities.** [1,8,10,11,13,14]

Science stands to benefit greatly from optimal, international exchanges of data, information, and knowledge. The advancement of scientific research and innovation is best supported by policies that promote transparency and open government to enhance the efficiency and effectiveness of government.

While the specific data sharing needs and opportunities will vary among countries and scientific communities, Federal agencies should raise awareness of U.S. data sharing policies, and principles adopted by the broader international community in the:

- G8 Open Data Charter;[10]
- Organization for Economic Cooperation and Development Principles and Guidelines for Access to Research Data from Public Funding;[11]
- Group on Earth Observation System of Systems Data Sharing Principles;[12] and
- Open Government Partnership.[14]

Agencies should promote the use of these open scientific data sharing principles in international scientific collaborations. Both bilateral and multilateral scientific conferences and meetings provide an opportunity to encourage international partners to understand the value of these principles.

References

[1] OMB Memorandum M-10-06, Open Government Directive (Dec. 8, 2009), available at https://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf
[2] Executive Order13642: Making Open and Machine Readable the New Default for Government Information, available at https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-
[3] Memorandum Open Data Policy – Managing Information as an Asset (M-13-13), available at https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf
[4] Project Open Data, available at https://project-open-data.cio.gov/
[5] OMB Circular A-130, as updated in 2016, is available at https://www.whitehouse.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf.
[6] OSTP Memo "Increasing Access to the Results of Federally Funded Scientific Research, February 22, 2013, available at http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf
[7] U.S. Open Data Action Plan (May 9, 2014), available at https://www.whitehouse.gov/sites/default/files/microsites/ostp/us_open_data_action_plan.pdf
[8]The "Bromley Principles" Regarding Full and Open Access to "Global Change" Data. By Allan Bromley, published in Policy Statements on Data Management for Global Change Research from the U.S. Office of Science and Technology Policy, July 2, 1991.
[9] The latest agency open government and public access plans are available on their respective Open Government Webpages, located at http://www.[agency].gov/open
[10] G8 Open Data Charter (June 18, 2013), available at: https://www.gov.uk/government/publications/open-data-charter
[11] OECD Principles and Guidelines for Access to Research Data from Public Funding available at http://www.oecd.org/sti/sci-tech/38500813.pdf
[12] GEOSS Data Sharing Principles available in the 10-Year Implementation Plan available at https://www.earthobservations.org/documents/10-Year%20Implementation%20Plan.pdf.
[13] White House Transparency and Open Government Memorandum for the Heads of Executive Departments and Agencies, available at http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment.
[14] Open Government Partnership, http://www.opengovpartnership.org/about/mission-and-goals.