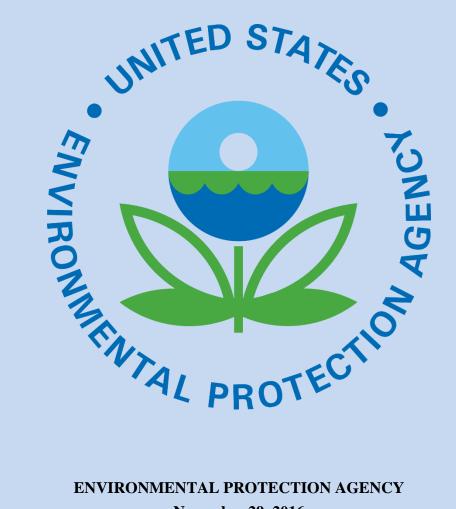# PLAN TO INCREASE ACCESS TO RESULTS OF EPA-FUNDED SCIENTIFIC RESEARCH

ENVIRONMENTAL PROTECTION AGENCY

November 29, 2016

Version 1.1

PLAN TO INCREASE ACCESS TO RESULTS OF
EPA-FUNDED SCIENTIFIC RESEARCH

Table of Contents

PLAN TO INCREASE ACCESS TO RESULTS OF EPA-FUNDED SCIENTIFIC RESEARCH

I.   BACKGROUND AND PURPOSE

On February 22, 2013, the White House Office of Science and Technology Policy (OSTP) issued a memorandum entitled "Increasing Access to the Results of Federally Funded Scientific Research" (OSTP Memo). The memorandum directs Federal departments and agencies that spend more than $100 million per year on research and development (R&D), including the Environmental Protection Agency (EPA), to develop and submit a plan to OSTP to increase public access to peer-reviewed, scientific research publications and research data (herein after referred to as **research publications** and **research data**)[1] resulting from agency-funded (R&D). The fundamental notion underlying this memorandum is that the results of federally funded scientific research should be available to the public, the scientific community, and industry to the greatest extent feasible consistent with applicable law and policy; agency mission; resource constraints; U.S. national, homeland, and economic security; and the specific objectives of the memorandum. In response, this plan describes steps the Agency will take to further increase access to results of EPA-funded scientific research, consistent with objectives of the OSTP Memo.

The mission of the Agency is to protect human health and the environment. To achieve this mission, the Agency generates and assesses information for environmental decision makers

---

[1] Definitions of key terms are included in Appendix A; the word or phrase is bolded the first time it appears in this document.

derived from environmental science, health science, natural resources, and related social science research. Increasing public access to EPA-funded scientific research results can support this mission and enhance innovation and economic competitiveness. The purpose of this plan is to describe steps that EPA will take to increase the availability of the results of EPA-funded research to the scientific community, environmental policy makers, other stakeholders, and the public in order to accelerate scientific breakthroughs that support the Agency's mission and policy making efforts.

EPA has a long history of collaboration in scientific research and is a leader in providing access to environmental information to encourage better decisions and a more informed public. Transparency is a core EPA value. The Agency already makes publicly available much of the Agency's scientific and technical work, including information that supports regulatory decisions. For example, EPA provides all materials and scientific information supporting each regulation in public dockets, which are publicly available for comment at www.regulations.gov. In addition, the Agency maintains an enterprise dataset **metadata** catalog (the Environmental Dataset Gateway (EDG) at https://edg.epa.gov/metadata/catalog/main/home.page) through which thousands of EPA datasets are publicly available. EPA's Enterprise Information Management Policy (EIMP), adopted March 3, 2015, codifies the Agency's approach to facilitating access to data held in EPA information systems.[2] Appendix D describes, in more detail, a small sample of EPA's extensive on-going efforts to be transparent and to increase access to environmental information.

While the Agency strives to increase access to its research results, it recognizes, consistent with the OSTP Memo, that Federal agencies have a responsibility to protect confidentiality and personal privacy, respect proprietary interests and property rights, and balance between the value of providing long-term access and its associated costs. It is important to recognize that some research data cannot be made fully available to the public but instead may need to be made available in more limited ways, e.g., establishing data use agreements with researchers that respect necessary protections. Whether research data are fully available to the public or available

---

[2] Links to the EIMP and associated cataloguing and metadata standards are found at http://www.epa.gov/open/digital-strategy.

to researchers through other means does not affect the validity of the scientific conclusions from **peer-reviewed research publications**.

II. OVERVIEW OF EPA'S PUBLIC ACCESS PLAN

This Plan describes how the Agency intends to increase access to **peer-reviewed scholarly publications** and digital data resulting from EPA-funded scientific research, consistent with applicable law and policy, EPA's mission, and resource constraints. The policies and approaches outlined in this Plan are consistent with the OSTP Memo and with the Office of Management and Budget (OMB) Open Data Policy (M-13-13) issued May 9, 2013.[3]

The goals of this Plan are to:
- Support and expand upon EPA's long-standing commitment to transparency and public access;
- Enhance the value of government-funded scientific research results;
- Support appropriate and effective access to and reliable preservation of EPA-funded scientific research;
- Promote innovation and support opportunities for collaborative, cross-disciplinary scientific research;
- Increase the pace of scientific discovery and promote more efficient and effective use of government funding and resources through effective research data management; and
- Increase public access to research data while protecting proprietary interests, intellectual property, and personal privacy.

III. SCOPE AND APPLICABILITY

This Plan prospectively covers peer-reviewed scientific research publications in scholarly journals and digital research data that result from EPA-funded research. The Plan does not apply to research publications or research data generated from scientific research conducted prior to the implementation of the Plan.

---

[3] A list of additional materials, including these resources, is included in Appendix B.

This Plan is subject to law, Agency mission, resource constraints, and the objectives detailed in the Office of Management and Budget (OMB) Open Data Policy (M-13-13) to the extent that research data are collected, stored, or managed by EPA. The Plan aims to increase access to EPA-funded research results while protecting national security, confidentiality, and personal property, recognizing proprietary and intellectual property rights, and preserving the balance between the value of long-term data preservation and access and the associated costs and administrative burdens, consistent with the OSTP Memo and 2 C.F.R. § 200.315. Classified or otherwise protected EPA-funded scientific research will not be made publicly available. This Plan does not apply to scientific research not funded by EPA that may be cited in EPA assessments, rulemakings, or peer-reviewed publications. This Plan and any newly established requirements based upon it applies to **scientific research data** collected after its implementation, and not to scientific research data collected before its implementation. Scientific research data are the digital recorded factual material commonly accepted in the scientific community as necessary to validate research findings (OSTP Memo, 2 C.F.R 200.315).

This Plan complies with existing EPA policies, procedures, and guidance. It does not create or confer legal rights or obligations or impose any legally binding requirements on EPA or any party. It is not legally enforceable, and it does not constitute final Agency action. Nothing in this Plan shall be construed to impair or otherwise affect the authority granted by law to EPA. The validity of scientific conclusions drawn from research publications or their associated research data, or EPA's ability to consider those conclusions and data in its actions, does not depend on compliance with this Plan.

IV. IMPLEMENTATION

EPA will create a Forum on Increasing Public Access to EPA Research (Forum) to implement the Plan. The Forum should be established within four months after approval of the Plan. It will provide cross-Agency coordination to identify and recommend to the Administrator policies, procedures, infrastructure, and training for support of the Plan. Approved recommendations will be implemented by EPA programs and regional offices. Forum members representing each program and the regions will provide oversight to the implementation of the Plan.

Implementation will begin upon approval of this Plan and completed in three phases. Implementation dates listed in this Plan represent targets, rather than deadlines, that take into account the time needed for implementation by different EPA offices. EPA will identify resources within the existing Agency budget to implement the plan. Availability of funding will affect the implementation targets outlined in this Plan.

Phase 1 [2016]: EPA's Office of Research and Development (ORD) **intramural research.**

ORD has used existing infrastructure to implement increased public access to scientific research publications and research data carried out in whole or in part by ORD staff members. ORD has in place a single electronic manuscript clearance system to identify **scientific research** publications generated by intramural researchers. ORD requires intramural investigators to deposit scientific research publications into NIH's PubMed Central (PMC) and make them available no later than 12 months after the date of publication. ORD has developed standards for **Scientific Data Management Plans (SDMP)**. ORD will extend the SDMP standards to include the increasing public access goals established in this Plan. ORD intramural research metadata will be entered into the EPA EDG by the principal investigator, or their delegate, within 30 days of the paper being posted in PMC.

Phase 2 [2017 through 2018] EPA non-ORD Intramural Research.

The Forum will recommend, to program and regional offices, practices to identify and track **final peer-reviewed publications**. Non-ORD intramural researchers will deposit research publications into PMC prospectively once the needed infrastructure and training is established. Implementation is targeted to begin for manuscripts accepted for publication in 2017. Implementation for public access to non-ORD intramural research data is targeted for 2018 pending additional training, policies requiring data management plans, and infrastructure to provide access to non-ORD intramural research data.

Phase 3 [2018]: Extramurally funded research.

Within four months of approval of this Plan, EPA will begin developing additional processes, infrastructure, language, and training needed to increase public access to EPA-funded extramural research publications and data. Once adopted, implementation will begin prospectively with

2018 funded extramural requests for applications (RFAs) for grants, contracts, and cooperative/assistance agreements. EPA-funded extramural research publications will be made available to the public at no charge via NIH's PubMed Central and data management plans will be required for extramural research data.

EPA recognizes that duplicative or conflicting requirements might result when research is subject to public access policies from multiple federal agencies. In the course of implementing this Plan, EPA will consider how, when, and whether to apply the EPA policy to research that is subject to public access policies from other agencies. Interagency public access implementation working groups could provide a forum to consider and address these issues government-wide.

V.   PUBLIC ACCESS TO PEER-REVIEWED PUBLICATIONS FROM EPA-FUNDED SCIENTIFIC RESEARCH

Once implemented, EPA will require all funded investigators (intramural and extramural) to ensure that any publications resulting from EPA-funded research are publicly accessible on PMC no later than 12-months after the date of publication. Researchers are responsible for depositing research manuscripts or publications into PMC so that the public may read, download, and analyze research publications in digital form. EPA may base its actions on valid research publications and associated research data during the 12-month embargo period. This section of the Plan describes how EPA will enhance access to covered peer-reviewed publications to the public without charge.

Although the regulatory licenses for **extramural awards** allow the Agency to provide the public with access to peer-reviewed publications resulting from EPA-funded scientific research, there may be some limitations to allowed uses. Copyright holders retain rights for reproduction, redistribution, and reuse. Users of any peer-reviewed publications resulting from EPA-funded scientific research are directly and solely responsible for compliance with copyright restrictions, and expected to adhere to the terms and conditions defined by the copyright holder.

Described below are the publication repository functionalities, implementation approaches, and timelines for each peer-reviewed publication category, as well as the metrics to evaluate compliance with the plan.

A.   Repository for Publications

EPA will use NIH's PMC as the repository allowing the public to read, download, and analyze in digital form publications arising from EPA-funded research.

Prior to selecting PMC, EPA conferred with other federal agencies, and carefully considered and evaluated several options for a publications repository. The decision to employ PMC as the designated final repository for peer-reviewed publications reflects the Agency's commitment to collaborate with other agencies and the private sector in pursuing its public access policies. PMC is designed to preserve and make public full-text articles published in scientific journals. NIH actively engages with publishers to proactively make their content available in PMC, without charge to the public, and to assist researchers in complying with the statutory public access requirements applicable to their work. According to NIH, approximately 1,500 journals contribute full content of their journals to PMC regardless of whether any particular article is required to be submitted to PMC pursuant to the NIH Public Access policy.[4]

EPA's decision to have peer-reviewed articles included in PMC expands the scope and importance of the article repository by including additional environmental health publications. By supporting an existing, centralized repository, this selection will help facilitate the public's ability to locate publications. It should also further encourage publishers to use PMC to make their articles publicly available and contribute to the PMC public-private collaboration. Furthermore, EPA's decision to contribute its scientific publications to the existing and robust PMC repository amplifies and maximizes the Agency's research investment and the investments of NIH and other agencies in the PMC repository.

---

[4] https://publicaccess.nih.gov/policy.htm

Use of PMC's NIH Manuscript Submission System (NIHMS)-like system and the association of PMC with the PubMed citation catalog will satisfy the publication repository directives outlined in the OSTP Memo[5] including:

- *Facilitate easy public search and access* – support full and free access for users in determining the existence, description, location, and availability of information stored in the archive.
- *Input* - accept submissions and prepare the contents for storage and management within the archive.
- *Archival Storage* – store, maintain, and retrieve publications and metadata information.
- *Data Management* - populate, maintain, and access both publications and metadata.
- *Administration* - the overall operation of the archive system, including auditing submissions to ensure that they meet archive standards, and maintaining configuration management of system hardware and software.
- *Help prevent unauthorized mass redistribution* of scientific research publications.
- *Ensure proper author, journal, and publisher attribution.*
- *Preservation Planning* - ensure that the information stored in the archive remains accessible to the designated user community over the long-term
- *Uses widely available and nonproprietary archival formats* for text and associated content.
- *Provides access to materials in compliance with Section 508 of the Rehabilitation Act (29 U.S.C. §798).*

EPA intramural and extramural researchers will be responsible for ensuring that peer-reviewed publications resulting from EPA-funded scientific research are placed into PMC. EPA will develop standard extramural award terms and provide guidance to future research award recipients regarding public access to publications, including refraining from signing any agreements with publishers that restrict EPA's license rights and depositing publications in PMC. EPA will use a template developed by PMC similar to the NIHMS as a mechanism for

---

[5] For details on the way in which PMC and NIHMS satisfy the requirements of the OSTP Memo, please refer to Section II of the NIH's "Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research."

allowing researchers to submit the final, peer-reviewed version of their manuscripts for inclusion in PMC.

B.　　Embargo Period

EPA will establish a maximum 12-month post-publication **embargo period** for making publications publicly available. Stakeholders may petition the Agency's Science Advisor to change the embargo period for publications in a specified scientific field. A petition must demonstrate that the existing embargo period for certain fields of scientific research would be inconsistent with EPA's mission or the objectives articulated in the OSTP Memo. When considering changes to the embargo period, EPA will consult with other agencies that fund related areas of scientific research.

VI.　PUBLIC ACCESS TO EPA-FUNDED RESEARCH DATA

This Plan aims to maximize access, by the general public and without charge, to digitally formatted data resulting from EPA funded research, while protecting confidentiality and personal privacy, recognizing proprietary interests, business confidential information and intellectual property rights, and preserving the balance between the relative benefits and costs of long-term preservation and access. This Plan encourages public release of research data supporting peer-reviewed, research publications. It does not cover research data shared with EPA but owned by other organizations. Inclusion of costs for data management and public access may be included in intramural and extramural research proposals. EPA will require research data underlying a publication are posted to publicly accessible data repositories within 30 days of posting the paper in PMC, unless: a) the dataset has already been made available to the public via public release or another sharing mechanism, or b) the research data cannot be released due to one or more of constraints, such as requirements to protect confidentiality, personal privacy, proprietary interest, or property rights.

A. Scientific Data Management Plans

Upon full implementation, researchers are required to establish a SDMP addressing public access, for all covered, EPA-funded, scientific research (intramural and extramural). SDMPs describe all collected or created research data and metadata, as well as plans for providing long-term preservation of, and access to, the research data, as appropriate. The SDMP should indicate that researchers will share, at minimum, research data funded by EPA associated with any scientific publication. The SDMP should also indicate if research data are accessible from a publicly accessible repository, a data use agreement, or research data centers under controlled conditions.   SDMPs should also describe the location, primary contacts, organization, and access restrictions for the research data.

This Plan requires that data security be included in SDMPs that are consistent with applicable laws, regulations, rules, and policies. Data security is an important aspect of protecting privacy, government assets, and EPA policies include provisions related to data security, and/or refer to regulatory requirements, as applicable. Additionally, data that belong to the Federal government are subject to the Federal Information Security Modernization Act (FISMA), 44 U.S.C. § 3541 *et seq*.[6]

EPA released a SDMP Guidance for ORD intramural research projects in August 2015. Future guidance will be developed for non-ORD intramural research and extramural research. New RFAs for extramurally funded EPA research will require an SDMP be included in submitted proposals.

To ensure appropriate evaluation of intramural SDMPs, program offices will review or evaluate staff plans. Applicants responding to a new RFA for EPA-funded scientific research will include a SDMP in their application, which will be evaluated as part of the proposal review process.

ORD will work with non-ORD programs and regions to support training for research data management through policies and procedures, webinars, data stewards and community of practice meetings /training sessions, and online training. For example, ORD maintains a variety

---

[6] http://csrc.nist.gov/groups/SMA/fisma/overview.html

of procedures, standard operating procedures, and best practices for research data management. The Scientific Data Management Community of Practice (SDMCoP) supports ORD's research community in sharing information and ideas about managing scientific research data. This group is intended to grow over time into a long-term, self-sustaining, collaborative community where members learn about SDM topics, help each other solve research data management problems, and provide input to enhance SDM practices across ORD.

In addition, the SDM Development Team coordinates with ORD research staff on real-world testing of SDM processes and tools. Participating researchers receive training on how to write an effective SDM plan. Training emphasizes how to tailor the SDM plan to the unique needs of each research effort. Research teams then write an SDM plan and provide feedback on the plan development process that ORD can then use to improve its SDM methods, guidance, training, and resources. ORD has created a SDM Kit, which is a collection of templates, guidance, and training materials that helps EPA researchers develop SDMPs.

B.   Access, Maintenance and Evaluation

To support data preservation and access, EPA will promote the deposit of data in publicly accessible databases, where appropriate and available. In coordination with other agencies and the private sector, EPA will assess preservation needs of scientific research data funded by EPA. Discovery of research data described in the SDMPs will be achieved by researchers registering the metadata in EPA's EDG, consistent with the EIMP Cataloging Information Procedure[7] (https://www.epa.gov/open). The metadata will include a field that points the user to both the data repository in which the research data resides and from which it can be downloaded, and web services and /or application programming interfaces available to access the data (www.epa.gov/open). The EDG is EPA's framework for making data more open and accessible, via a metadata catalog, data downloads, and web service integration. EDG provides users with a central access point to publicly available data sets and geospatial tools created and/or paid for by EPA program offices, regions, and laboratories. The EDG supports implementation of EPA's Open Government Plan with metadata within the system being harvested to the joint data.gov

---

[7] EPA EIMP Cataloging Information Procedure (CIO Transmittal No.: 15-004).

and the National Geospatial Platform metadata catalog to facilitate interagency and public data sharing. EPA will develop approaches for using EDG metadata for identifying and providing attribution to publications and datasets available under this Plan.

VII. EVALUATION AND COMPLIANCE

The Forum will develop and recommend to the Administrator specific procedures, guidelines, and strategies to implement this Plan. To ensure compliance, EPA will train current Agency researchers on the new requirements and will provide data management orientation to new Agency researchers. As described above, SDMPs for intramural research will be reviewed or evaluated by program offices. Extramural researchers will submit SDMPs that will be evaluated as part of the proposal review process. EPA will track research publications and establish metrics to monitor publications in PMC, SDMPs, entry of research data sets and metadata into EDG, as well as the interconnectivity of the publications in PMC, research data sets, and metadata in EDG.

EPA will ensure compliance with extramural data management requirements by requiring, as a term and condition of the grant or contract award, periodic reporting to contracting officer representatives and project officers as a part of regular grants and contract management. Extramural researchers will periodically report the status of publications and research data collection and preservation, including any deviation from the approved SDMP required by the extramural award. Non-compliance with the terms and conditions of the award regarding public access may be considered as a negative indicator of past performance and may result in withholding of extramural funding.

For all intramural research, EPA will collect information on publications from our internal clearance processes, relevant research data repositories, metadata directories, other reference sources, and grant, cooperative agreement, and contract reports to assess compliance with this Plan's requirements. This information will also be a useful component for evaluating overall program success.

VIII. STAKEHOLDER CONSULTATION AND INTERAGENCY COORDINATION

EPA will coordinate with other agencies and the private sector to improve research data access and support for training, education and professional development related to data management, analysis, and/or long-term preservation of scientific data in areas of science supported by the Agency.

EPA participates in CENDI, an interagency working group of senior managers from major federal agencies who create, manage, aggregate, organize, and provide access to scientific and technical information. Member organizations represent a cross-section of federal data and publication providers, including libraries, data centers, aggregators, information technology developers, and content management providers. CENDI agencies play an important role in addressing science- and technology-based national priorities and strengthening U.S. competitiveness. CENDI meetings, workshops, and conferences provide training, education and a showcase for best practices in data management by its members.

EPA intends to work collaboratively with the STEM communities and information specialists, including those at the National Archives and Records Administration (NARA) to identify emerging challenges and technologies that affect stewardship and long-term preservation, and to define practices that ensure access to peer-reviewed publications into the future. For example, EPA has participated in recent National Academy of Science (NAS) meetings on Federal Research Regulations and Reporting Requirements: A New Framework for Research Universities in the 21st Century, as well as engaged in discussions on the impact of public access with our federal partners as part of the Interagency Working Group on Research Integrity.

Further, EPA will explore strategies, including outreach programs and online guidance in the form of frequently asked questions (FAQs), for ensuring the relevant scientific communities are aware of public access requirements. EPA will collaborate with our federal partners on the most effective ways to increase awareness within the scientific community.

IX. PUBLIC NOTICE

Once finalized, EPA will post this Plan on its Open Government website (www.epa.gov/open).

X.   UPDATE AND RE-EVALUATION OF THE PLAN

This Plan will be reviewed and revised as needed, including reviews as part of the periodic reporting requirements to OSTP and OMB. EPA will amend this plan, as appropriate, in consultation with OSTP and OMB. EPA will provide updates to OSTP and OMB on implementation of this plan in January and July of each year for the first two years after the plan is approved.

XI.   TIMELINE FOR IMPLEMENTATION

Action	Target Date
Convene Agency Scientific Publication Access Data and Management Working Group	COMPLETE
Select ORD intramural scientific article and manuscript repository submission system	COMPLETE
Identify and test repository compatibility with EPA's EDG	COMPLETE
Participate in EPA efforts regarding identification of data management needs and strategies for leveraging data management resources	ONGOING
Participate in interagency data and publication access plan implementation working groups	ONGOING
Activate prospective publication submission system in PMC for ORD intramural research efforts covered by the Plan	2016
Begin depositing of prospective ORD intramural peer-reviewed research publication in PMC	2016
Begin submission and review of SDMPs as part of ORD intramural research efforts	2016
Begin depositing prospective ORD intramural research data into the designated EPA data repository	2016
Complete EPA's Plan for Public Access to EPA-Funded Scientific Research Publications and Associated Digital Data	As Soon As Possible
Establish the Forum on Increasing Public Access to EPA Research.	Within four months after the approval of the plan
Begin depositing of prospective EPA (non-ORD) intramural peer-reviewed research manuscripts in PMC	2017

Action	Target Date
Identify and work towards developing processes, infrastructure, language, and training needed to implement future EPA-funded extramural research.	Within four months of approval of the Plan.
Begin depositing prospective non-ORD intramural research data into designated EPA data repository	2018
Begin implementing Phase 3 by including relevant language in EPA-funded extramural solicitations.	2018
Evaluate progress of access to peer-reviewed publications and research data	Every 6 months upon implementation

Appendix A. DEFINITIONS

Assistance agreement is an EPA grant, cooperative agreement, interagency agreement, or fellowship.

Embargo period is the time after an article's peer-reviewed publication during which EPA does not itself provide for public access to the article.

Extramural award means financial assistance to an external entity that provides support or stimulation to accomplish a public purpose. Awards include grants and other agreements in the form of money or property in lieu of money by the U.S. Government to an eligible recipient.

Extramural research is research that is done by external entities, e.g., through grants, contracts, and cooperative agreements.

Final manuscript is an author's or authors' final version of a peer-reviewed paper accepted for journal peer-reviewed publication, including all modifications resulting from the peer-review process.

Final peer-reviewed publication is a publisher's authoritative copy of the paper (version of record), including all modifications from the publishing peer-review process, copyediting, stylistic edits, and formatting changes.

Intramural research is research that is done by EPA employees as part of their official duties.

Manuscripts are non-published documents.

Metadata are structured information that describe content, data resources, or objects and help locate, use, understand, share, and manage these information sources. Metadata should answer questions about information (content and/or data) such as its purpose, means of creation or collection, date of creation or collection, name of the author/creator, peer-reviewed publisher; and, for data in particular, characteristics of the resource (i.e., raw data/in original form or how the data were changed), what they were used for, and their use limitations.

Peer-reviewed publication is the full text document, and any associated supplementary text materials, posted or published by a peer-review journal. This term describes specific types of

articles or research data deemed published when the version of record of scientific research results appears in a peer-reviewed journal (on-line or hard copy). EPA interprets "publication" to include articles in journals specializing in natural and physical sciences (e.g., biology, chemistry, physics, health science, geology, and engineering), social sciences (e.g., economics, psychology, and sociology), mathematics, statistics, and computer science.

Scientific data management plan (SDMP) is a document that accompanies a research proposal and includes information on the scope, costs, and process of making scientific research results available, considering protected data, data use, and preservation.

Scientific research is the systematic inquiry directed toward fuller scientific knowledge or understanding of the subject studied. As discussed in this Plan, the *results of scientific research* refer to both the research publications in peer-reviewed journals and the associated digital research data that support the scientific results of the research publications.

Scientific research data are defined, consistent with the OSTP Memo and 2 C.F.R 200.315 as the digital recorded factual material commonly accepted in the scientific community as necessary to validate research findings.[8] *Research data* as used in this Plan are the digital scientific research data resulting from EPA-funded scientific research.

---

[8] Consistent with the definition in 2 C.F.R. § 200.315(e)(3), research data does not include:
- Preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues;
- Physical objects (e.g., laboratory samples);
- Trade secrets and commercial information;
- Materials necessary to be held confidential by a researcher until publication of results in a peer-reviewed journal; and
- Personnel, medical, and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.

The following specific examples of scientific research are excluded from this Plan:
- Interim results or other preliminary scientific research data not used to generate the results in the final peer-reviewed publication;
- Preliminary scientific research documentation beyond the article, supplementary materials, and metadata regarding preliminary research plans, including preliminary study protocols and other preliminary *a priori* decisions (recognizing that preliminary plans may have changed during the research project);
- Information that may disclose intellectual property rights;
- National security and other classified information.

Appendix B. ADDITIONAL MATERIAL

U.S. law and policy promote access to peer-reviewed scientific research publications produced using federal funds. The following is a non-exhaustive list of such laws and policies.

- Office of Science and Technology Policy (OSTP) Memorandum, "Increasing Access to the Results of Federally Funded Scientific Research," February 22, 2013: http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo _2013.pdf
- Office of Management and Budget (OMB) Memorandum, "Open Data Policy—Managing Information as an Asset," May 9, 2013 (M-13-13): http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf
- EPA's Enterprise Information Management Policy is a policy designed to establish a standard framework for information management across EPA. March 2015: https://www.epa.gov/open/digital-strategy
- EPA's Scientific Integrity Policy recognizes EPA's commitment to the free flow of scientific information, stating that the Agency "will continue to expand and promote access to scientific information by making it available online in open formats in a timely manner, including access to data and non-proprietary models underlying Agency policy decisions." The Scientific Integrity Policy "is intended to outline the Agency's expectations for developing and communicating scientific information to the public … by further providing for and protecting EPA's longstanding commitment to the timely and unfiltered dissemination of its scientific information." February 2012: https://www.epa.gov/sites/production/files/2014-_02/documents/scientific_integrity _policy_2012.pdf
- Section 102(G) of National Environmental Policy Act, 42 U.S.C. § 4332(2) (G), authorizes federal agencies to "make available … advice and information useful in restoring, maintaining, and enhancing the quality of the environment."
- Section 104 of the Clean Water Act, 33 U.S.C. § 1254, authorizes the Administrator to, among other things, "collect and make available" results of certain research related to pollution.
- Section 103 of the Clean Air Act, 42 U.S.C. § 7403, authorizes the Administrator to, among other things, "collect and make available" results of certain research peer-reviewed publications related to air pollution.

- Section 8001 of the Solid Waste Disposal Act, 42 U.S.C. § 6891, authorizes EPA to "conduct, and encourage, cooperate with, and render financial and other assistance to appropriate public … authorities, agencies, and institutions, private agencies and institutions, and individuals in the conduct of, and promote the coordination of, research, investigations, experiments, training, demonstrations, surveys, public education programs, and studies relating to [solid waste issues]."

- Section 10 of the Toxic Substances Control Act, 15 U.S.C. § 2609, authorizes the exchange of research and development results regarding toxic chemical substances and mixtures between EPA and other federal, state, and local authorities.

- Section 20 of the Federal Insecticide, Fungicide, and Rodenticide Act, 7 U.S.C. § 136r, authorizes EPA to undertake research and monitoring activities, including monitoring humans and animals for pesticide exposure and research into integrated pest management.

- Section 203 of the Marine Research, Protection, and Sanctuaries Act, 33 U.S.C. § 1443, authorizes EPA to "conduct research, investigations, experiments, training, demonstrations, surveys, and studies" to address dumping into ocean waters. This section also authorizes the Administrator to "encourage, cooperate with, promote the coordination of, and render financial and other assistance" to public and private agencies and institutions and individuals.

- EPA's regulation, under 2 C.F.R. § 200.315 makes data produced under an award made available to the public through procedures established under the Freedom of Information Act."

Appendix C. ROLES AND RESPONSIBILITIES

The roles and responsibilities for carrying out this Plan cross a number of offices and groups within EPA. Four categories of tasks will be carried out by these organizations:

- Policies, procedures, and standards development
- Business processes, procedures, and resources development
- Information systems and automation development
- Outreach, education, and training

EPA's Science Advisor, in consultation with the Science and Technology Policy Council

- Assumes EPA-wide authority and management responsibility for this Plan.
- Coordinates development of a sustainable funding strategy to implement the Plan.
- Coordinates collaboration and cooperation on public access with other federal agencies.
- Issues guidance, as necessary, for implementation of the Plan.
- Develops metrics to evaluate compliance with the Plan.
- Considers petitions to alter an embargo period for a disciplinary field.
- In collaboration with other federal agencies, develops approaches for identifying and providing appropriate attribution to scientific data sets that are made available under the Plan.

Office of Administration and Resources Management

- In consultation with the Office of General Counsel and other relevant offices, develops language for contracts and **assistance agreements** (e.g., for Request for Proposals and terms and conditions) that will enable funding recipients to comply with this Plan.
- Reviews general contracts and assistance agreement language and revises it as needed to incorporate the requirements of this Plan.
- With assistance of the Office of Environmental Information and Office of Research and Development, and in collaboration with other federal agencies, identifies training, education, and workforce development available within the government or in the private sector related to the skills needed for information management, storage, preservation, and stewardship.

Office of Environmental Information

- Reviews the Plan for conformance with EPA CIO Policy 2130.1 Section 508: Accessible Electronic and Information Technology (to ensure accessibility to people with disabilities) and makes any needed revisions to this Plan.
- Provides ongoing support for issues on metadata standards, federal data requirements and initiatives, enterprise information catalogs, outreach to libraries and librarians, and others.
- Chief Information Officer, working with the Quality Information Council, ensures that the principles of this Plan are coordinated with the Agency's Information Policy and Procedures framework.

Office of Research and Development
- Establishes a memorandum of understanding (MOU) with Pub Med Central.
- Develops, disseminates, and provides training on Data Management Guidelines.

All EPA Headquarters and Regional Offices
- Ensure compliance with this Plan in coordination with the Office of the Science Advisor.
- Ensure that necessary staff receives the appropriate training to carry out the requirements of this Plan.

Appendix D. EXAMPLE OF EXISTING MECHANISMS FOR PUBLIC ACCESS TO EPA-GENERATED DATA SETS AND VISUALIZATION TOOLS

EPA has a long history of making environmental data and information, especially information supporting regulatory actions, publicly accessible. Currently EPA provides public access to a wide range of environmental data and information through websites, libraries, data analytics, and presentation tools. Ensuring this access includes delivering high-quality data discovery, scientific, analytical, and statistical services to support research and decision making in the environmental and health arenas.

The EPA Toxics Release Inventory (TRI) Program (https://www.epa.gov/toxics-release-inventory-tri-program), established in 1986, is a centerpiece of the community right-to-know approach for environmental information. Through this program, EPA annually provides information on toxic chemical releases and other waste management data to the public. The TRI *Explorer (http://catalog.data.gov/dataset/toxics-release-inventory-tri-explorer-widget)* allows the public to generate reports on releases, transfers, and waste management that can be displayed by facility, chemical, geographic area, industry (North American Industry Classification System code), reporting years, and maps.

Envirofacts went public in 1995, providing a variety of data from program systems across EPA to the public via the Envirofacts website (https://www3.epa.gov/enviro/), including data about air, land, water, waste, toxics, radiation, regulated facilities, compliance grants, and regulated substances. The EnviroMapper (https://catalog.data.gov/dataset/enviromapper) released in 1998 enabled the public to access an interactive web mapping application to map EPA-regulated facilities and obtain associated information from Envirofacts in a geographic area of their choice such as a city, town, or zip code. Over the years, this application evolved into MyEnvironment (https://catalog.data.gov/dataset/myenvironment), currently EPA's main public access mapping tool. Users can view Envirofacts data in conjunction with links to web sites that contain information on air, land, water, and changes in their environment. This includes allowing users to map impaired water bodies and streams and view information about the area they mapped, e.g., watersheds.

In 2003, EPA developed the Geospatial Data Index (GDI), the first Agency data registry, which provided EPA staff with an extensive index of geospatial data available to support the Agency's programmatic and regulatory responsibilities. By 2007, the GDI evolved into

the Geospatial Data Gateway (GDG), providing a central access point for EPA's geospatial resources with metadata links to data, applications, services and other resources contributed by EPA regions, programs, and laboratories. "Unrestricted" metadata in GDG were for the first time shared with both the public and external catalogs such Geospatial One-Stop (GOS) and Data.gov (Geodata Catalog). In 2011, the GDG was expanded to include metadata for all types of datasets and renamed the Environmental Data Gateway (EDG). The EDG is EPA's framework for making data more open and accessible, via a metadata catalog, data downloads, and web service integration. The EDG supports implementation of EPA's Open Government Plan and adherence to the principles of transparency, participation, and collaboration by helping EPA identify and publish high-value resources.

The Agency continues to expand its public access efforts through the EPA Geospatial Platform released in May 2012. This platform is a shared infrastructure, which contains a suite of geospatial tools, data, and services that are closely linked with the EDG from which users can access data. During the last year, several tools and maps such as EJSCREEN, an environmental justice screening and mapping tool with data visualized via maps, have been made available through the Geoplatform. In the future, EPA hopes to publish more public maps on the platform, because visualization of the data makes it easier to understand the wide array of associated environmental and health issues.

With the release of the EPA EIMP in March 2015, the Agency entered the next phase of data management. It is now EPA policy to manage information as a strategic asset critical for meeting its mission and open government goals. This includes ensuring that information held or cataloged in information management systems by EPA is:

- Managed via a defined information life cycle process (appropriate for the information type)
- Cataloged and/or labeled with metadata, in EPA and federal-wide registries, repositories, andother information systems
- Developed, maintained, and preserved in open and machine readable formats using established standards; and
- Publicly available compliant with National Security Information (NSI) and Controlled Unclassified Information and other pertinent statues, regulations, and policies that have requirements associated with sensitive information.

ORD is now partnering with the Office of Environmental Information to build enterprise data management approaches, governance, and platforms to help steward research data, make it

publicly available, and ensure it is available for reuse, analysis, and further study. ORD is also implementing a science data management program that includes requirements to develop data management plans, tag research data sets with metadata, register them in the EDG, and file and maintain the data plan and associated records centrally within a platform called the Science Hub. This effort relates to the broader issue of data management within EPA to protect, publish, and reuse these valuable data assets. Its initial focus is on addressing the various Open Data Policies, including the OSTP memo.

Other examples of data made publicly available by the Agency include:

- The Health and Environmental Research Online (HERO) database provides an easy way to access and influence the scientific literature used in EPA science assessments. The database includes more than 600,000 scientific references and data from the peer-reviewed literature used by EPA for the following: Integrated Science Assessments (ISA) that are part of the reviews of the National Ambient Air Quality Standards (NAAQS); Provisional Peer Reviewed Toxicity Values (PPRTV) that represent human health toxicity values for the Superfund program; and the Integrated Risk Information System (IRIS), a database that supports critical Agency policymaking for chemicals. HERO is an EVERGREEN database; i.e., new studies are continuously added, so scientists can keep abreast of current research. Imported references are systematically sorted, classified, and made available for search and citation. (https://hero.epa.gov/hero/)
- The Aggregated Computational Toxicology Resource (ACTOR) is a database on environmental chemicals that is searchable by chemical name (and other identifiers) and by chemical structure. This information is consolidated from more than 200 publicly available sources of data. ToxCast, a part of ACTOR, is used as a cost-effective approach for efficiently prioritizing the toxicity testing of thousands of chemicals. It uses data from state-of-the-art high-throughput screening bioassays and builds computational models to forecast potential chemical toxicity in humans. ToxRefDB stores the data related to ToxCast. (https://www.epa.gov/aboutepa/about-national-center-computational-toxicology-ncct)
- Storage and Retrieval for Water Quality Data is a repository for water quality, biological, and physical data. (https://www.epa.gov/waterdata/storage-and-retrieval-and-water-quality-exchange)
- The Ecotoxicology Database provides information on effects of single chemicals to ecologically relevant species. (https://cfpub.epa.gov/ecotox/)
- EnviroAtlas was produced by EPA and allows those interested to interact with a web-based, easy-to-use, mapping application to view and analyze multiple ecosystem services

for the contiguous United States. The dataset is available as downloadable data (https://edg.epa.gov/data/Public/ORD/EnviroAtlas) or as an EnviroAtlas map service. Additional descriptive information about each attribute in this dataset can be found in its associated EnviroAtlas Fact Sheet. (https://www.epa.gov/enviroatlas)

- The Air Quality System database contains measurements of air pollutant concentrations throughout the United States and its territories. The measurements include both criteria air pollutants and hazardous air pollutants. (http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html)

- The Air Quality System Data Mart is a storehouse of air quality information that allows users to make queries of of data. The Data Mart also includes information from EPA's substance and facility registry systems. (https://www3.epa.gov/airdata/ad_data.html)

Plan to Increase Access to Results of EPA-Funded Scientific Research

U.S. Environmental Protection Agency
1200 Pennsylvania Ave NW
Washington, DC 20460

601-R-16-005