DOE Data Management Plan Requirements

2016 STIP Working Meeting
April 12, 2016

Carly Robinson
OSTI Senior Science Advisor/ Product Strategist

U.S. DEPARTMENT OF ENERGY | Office of Science | Office of Scientific and Technical Information

Brief History – Data

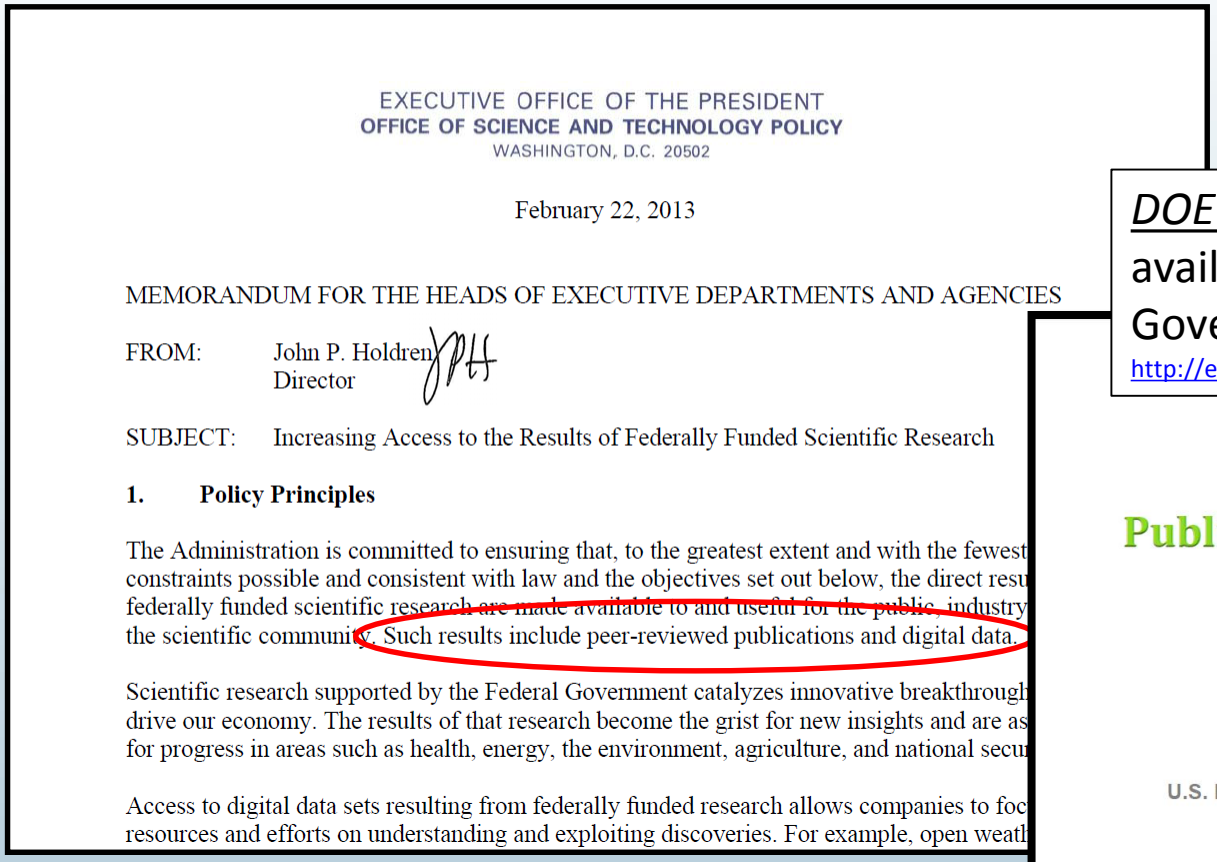
- **COMPETES 2010 "Interagency Public Access Committee"**
- Office of Science Working Group on Digital Data
- Office of Science FACA Reports (2011)
- OSTP Request for Information (2012)
- Office of Science User Facility Input (2013)

- **OSTP Memo "Increasing Access to the Results of Federally Funded Scientific Research" (Feb., 2013)**
- **DOE Public Access Plan and Office of Science Statement on Digital Data Management (July, 2014)**
- **DOE Policy for Digital Research Data Management (Sept, 2015)**

OSTP Memo and the DOE Response

Increasing Access to the Results of Federally Funded Scientific Research

http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

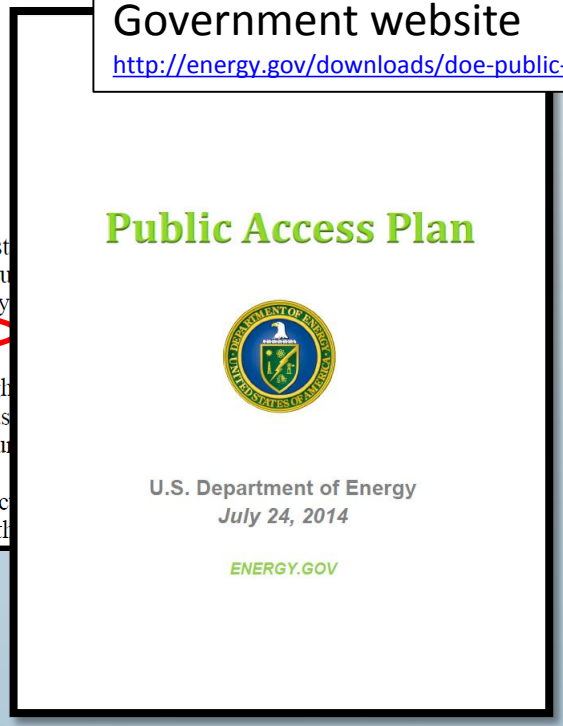
1. Policy Principles

The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, the scientific community. Such results include peer-reviewed publications and digital data.

Scientific research supported by the Federal Government catalyzes innovative breakthroughs drive our economy. The results of that research become the grist for new insights and are as for progress in areas such as health, energy, the environment, agriculture, and national secu

Access to digital data sets resulting from federally funded research allows companies to foc resources and efforts on understanding and exploiting discoveries. For example, open weath

DOE Public Access Plan is available on the DOE Open Government website
http://energy.gov/downloads/doe-public-access-plan



Public Access Plan

U.S. Department of Energy
July 24, 2014

ENERGY.GOV

DOE Public Access Plan: A Department-wide policy

The Office of Science intends to publish its data management plan requirements on July 28, 2014. Starting October 1, 2014, the requirements will be included in all invitations and solicitations for research funding issued by the Office of Science.

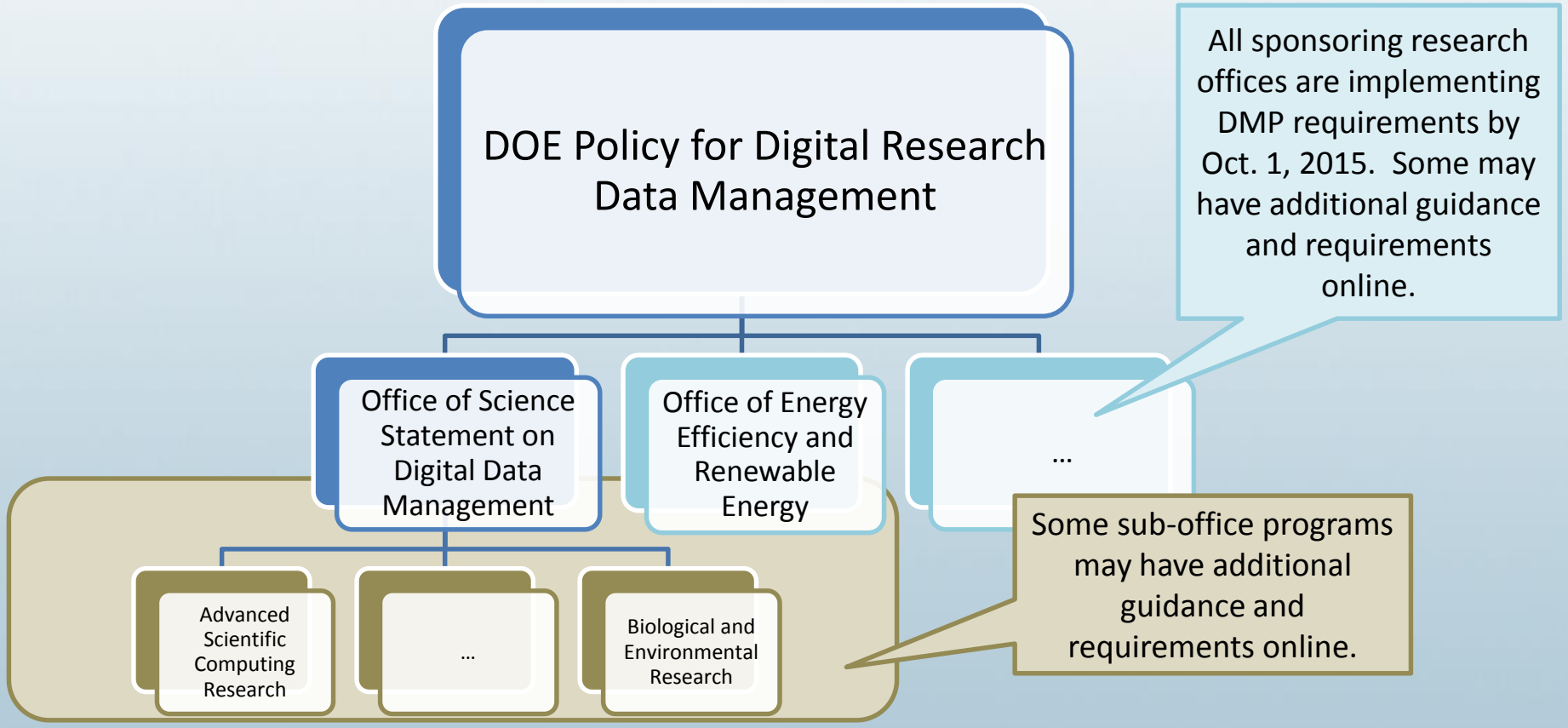
Other DOE Offices and elements will implement data management plan requirements no later than October 1, 2015. The result will be a Department-wide policy. Should it be necessary, additional supplementary guidance and requirements addressing specific needs would be issued by each Office or element and coordinated centrally.

DOE Policy for Digital Research Data Management

http://www.energy.gov/datamanagement/

DOE Policy on Digital Research Data Management

- Office of Science Statement on Digital Data Management, effective Oct. 1, 2014
- DOE Policy for Digital Research Data Management, effective Oct. 1, 2015



DOE Policy on Digital Research Data Management

REQUIREMENTS AND GUIDANCE FROM DOE SPONSORING OFFICES

REQUIREMENTS

All DMPs submitted to any DOE sponsoring office should meet the following requirements:

- DMPs should describe whether and how data generated in the course of the proposed research will be **shared** and **preserved** and, at a minimum, describe how data sharing and preservation will enable **validation** of results, or how results could be validated if data are not shared or preserved.
- DMPs should provide a plan for making all research data displayed in publications resulting from the proposed research open, machine-readable, and digitally accessible to the public at the time of publication. This includes data that are displayed in charts, figures, images, etc. In addition, the underlying digital research data used to generate the displayed data should be made as accessible as possible to the public in accordance with the **Principles** stated above. The published article should indicate how these data can be accessed.
- DMPs should consult and reference available information about data management resources to be used in the course of the proposed research. In particular, DMPs that explicitly or implicitly commit data management resources at a facility beyond what is conventionally made available to approved users should be accompanied by written approval from that facility. In determining the resources available for data management at DOE Scientific User Facilities, researchers should consult the **published description of data management resources** and practices at that facility and reference it in the DMP.
- DMPs must protect confidentiality, personal privacy, **Personally Identifiable Information**, a security; recognize proprietary interests, business confidential information, and intellectual pr on innovation and U.S. competitiveness; and otherwise be consistent with all applicable laws and DOE orders and policies.

In instances where the Department intends to collect digital data resulting from the supported research, additional requirements for data management may be necessary to ensure the Department meets the requirements of the **Open D… olicy**.

All sponsoring research offices will have DMP requirements by Oct 1, 2015. Some may have additional guidance and requirements online.

Additional requirements for the DMP may be identified by the sponsoring office, program, sub-program, or in the solicitation. Some sponsoring research offices have provided additional requirements and guidance as detailed below:

- **The Office of Science Statement on Digital Data Management**
- **DOE SBIR and STTR Programs Office**

DMP Requirements

DOE Policy for Digital Research Data Management	SC Statement on Digital Data Management
DMPs may be required as part of research proposal or later in the award process	DMPs required as part of research proposals
DMP requirements come into effect Oct 1, 2015	DMP requirements came into effect Oct 1, 2014
Common principles and requirements for DMPs for all DOE offices	
Common “Suggested Elements” for DMPs for all DOE offices	
Common definitions for all DOE offices	

DOE Policy for Digital Research Data Management

Principles

- Effective data management has the potential to increase the pace of scientific discovery and promote more efficient and effective use of government funding and resources. Data management planning should be an integral part of research planning.

- Sharing and preserving data are central to protecting the integrity of science by facilitating validation of results and to advancing science by broadening the value of research data to disciplines other than the originating one and to society at large. To the greatest extent and with the fewest constraints possible, and consistent with the requirements and other principles of this Statement, data sharing should make digital research data available to and useful for the scientific community, industry, and the public.

- Not all data need to be shared or preserved. The costs and benefits of doing so should be considered in data management planning.

DOE Policy for Digital Research Data Management

- Requirements apply to proposals for research funding
- Requirements apply to proposals submitted for new, renewal, and some supplemental research funding
- Requirements apply to proposals regardless of the PI's institution
- Requirements do *not* apply to applications to use DOE scientific user facilities.
- DOE sponsored research activities at the DOE National Laboratories for which a DOE-approved DMP does not already exist will be required to develop a DMP. In most cases, the DMP will be requested as part of the next peer review organized by the DOE sponsoring research office.

DOE Policy for Digital Research Data Management

Requirements

1. DMPs should describe whether and how data generated in the course of the proposed research will be shared and preserved and, at a minimum, describe how data sharing and preservation will enable validation of results, or how results could be validated if data are not shared or preserved.

DOE Policy for Digital Research Data Management

2. DMPs should provide a plan for making all research data displayed in publications resulting from the proposed research open, machine-readable, and digitally accessible to the public at the time of publication. This includes data that are displayed in charts, figures, images, etc. In addition, the underlying digital research data used to generate the displayed data should be made as accessible as possible to the public in accordance with the Principles stated above. The published article should indicate how these data can be accessed.

DOE Policy for Digital Research Data Management

3. DMPs should consult and reference available information about data management resources to be used in the course of the proposed research. In particular, DMPs that explicitly or implicitly commit data management resources at a facility beyond what is conventionally made available to approved users should be accompanied by written approval from that facility. In determining the resources available for data management at DOE Scientific User Facilities, researchers should consult the published description of data management resources and practices at that facility and reference it in the DMP.

DOE POLICY FOR DIGITAL RESEARCH DATA MANAGEMENT: RESOURCES AT DOE SCIENTIFIC USER FACILITIES

DOE Policy for Digital Research Data Management
Resources at DOE Scientific User Facilities
Suggested Elements for a Data Management Plan
FAQs
Glossary

OFFICE OF SCIENCE

ADVANCED SCIENTIFIC COMPUTING RESEARCH (ASCR)

FACILITY	HOST INSTITUTION	DATA MANAGEMENT RESOURCES
National Energy Research Scientific Computing Center (NERSC)	LBNL	Link
Argonne Leadership Computing Facility (ALCF)	ANL	Link
Oak Ridge Leadership Computing Facility (OLCF)	ORNL	Link
Energy Sciences Network (ESnet)	LBNL	Link

BASIC ENERGY SCIENCES (BES)

FACILITY	TYPE	HOST INSTITUTION	DATA MANAGEMENT RESOURCES
Advanced Light Source (ALS)	Light Source	LBNL	Link
Advanced Photon Source (APS)	Light Source	ANL	Link
Linac Coherent Light Source (LCLS)	Light Source	SLAC	Link
National Synchrotron Light Source II (NSLS-II)	Light Source	BNL	Link
Stanford Synchrotron Radiation Light Source (SSRL)	Light Source	SLAC	Link
High Flux Isotope Reactor (HFIR)	Neutron Source	ORNL	Link
Spallation Neutron Source (SNS)	Neutron Source	ORNL	Link
Center for Functional Nanomaterials	Nanoscale Science Research Center	BNL	Link

DOE Policy for Digital Research Data Management

4. DMPs must protect confidentiality, personal privacy, Personally Identifiable Information, and U.S. national, homeland, and economic security; recognize proprietary interests, business confidential information, and intellectual property rights; avoid significant negative impact on innovation and U.S. competitiveness; and otherwise be consistent with all applicable laws, regulations, agreement terms and conditions, and DOE orders and policies.

Suggested Elements for a Data Management Plan

ENERGY.GOV

Search Energy.gov

PUBLIC SERVICES | SCIENCE & INNOVATION | ENERGY SAVER | ABOUT ENERGY.GOV | OFFICES

Home » DOE Policy for Digital Research Data Management: Suggested Elements for a Data Management Plan

DOE POLICY FOR DIGITAL RESEARCH DATA MANAGEMENT: SUGGESTED ELEMENTS FOR A DATA MANAGEMENT PLAN

DOE Policy for Digital Research Data Management

Resources at DOE Scientific User Facilities

Suggested Elements for a Data Management Plan

FAQs

Glossary

The Principal Investigator or other appropriate research lead should determine which data should be the subject of the Data Management Plan (DMP) and, in the DMP, propose which data should be shared and/or preserved in accordance with the Requirements of this policy.

The following list of elements for a DMP provides suggestions regarding the data management planning process and the structure of the DMP:

- **Data Types and Sources**. A brief, high-level description of the data to be gen
  and which of these are considered Digital Research Data necessary to Valid
- **Content and Format**. A statement of plans for data and metadata content and
  documentation plans, annotation of relevant software, and the rationale for the
  community standards should be used where possible. Where community stand
  alternate strategies that facilitate sharing, and should advise the sponsoring pr
- Data Sharing and Data Preservation. A description of the plans for data sha
  - the anticipated means for sharing and the rationale for any restrictions on
  - a timeline for sharing and preservation that addresses both the minimum le
    delay to data access after research findings are published;
  - any special requirements for data sharing, for example, proprietary software needed to access or interpret data, applicable policies, provisions, and licenses for re-use and re-distribution, and for the production of derivatives, including guidance for how data and data products should be cited;
  - any resources and capabilities (equipment, connections, systems, software, expertise, etc.) requested in the research proposal that are needed to meet the stated goals for sharing and preservation. (This could reference the relevant section of the associated research proposal and budget request);

- **Data Types and Sources**
- **Content and Format**
- **Sharing and Preservation**
- **Protection**
- **Rationale**

Definitions

Digital Research Data:

The term *digital data* encompasses a wide variety of information stored in digital form including: experimental, observational, and simulation data; codes, software and algorithms; text; numeric information; images; video; audio; and associated metadata. It also encompasses information in a variety of different forms including raw, processed, and analyzed data, published and archived data.

This statement focuses on *digital research data,* which are *research data* that can be stored digitally and accessed electronically. OMB Circular A110 defines *research data* as follows:

> "Research data is defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This 'recorded' material excludes physical objects (e.g., laboratory samples). Research data also do not include:
>
> (A) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and
>
> (B) Personnel and medical information and similar information 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."

Definitions

Data Preservation:

Data preservation means providing for the usability of data beyond the lifetime of the research activity that generated them.

Data Sharing:

Data sharing means making data available to people other than those who have generated them. Examples of data sharing range from bilateral communications with colleagues, to providing free, unrestricted access to the public through, for example, a web-based platform.

Validate:

In the context of this statement, *validate* means to support, corroborate, verify, or otherwise determine the legitimacy of the research findings. Validation of research findings could be accomplished by reproducing the original experiment or analyses; comparing and contrasting the results against those of a new experiment or analyses; or by some other means.

SC Funding Opportunity Announcement Language

APPENDIX 6: DATA MANAGEMENT PLAN

Provide a Data Management Plan (DMP) that addresses the following requirements:

1. DMPs should describe whether and how data generated in the course of the proposed research will be shared and preserved. If the plan is not to share and/or preserve certain data, then the plan must explain the basis of the decision (for example, cost/benefit considerations, other parameters of feasibility, scientific appropriateness, or limitations discussed in #4). At a minimum, DMPs must describe how data sharing and preservation will enable validation of results, or how results could be validated if data are not shared or preserved.
2. DMPs should provide a plan for making all research data displayed in publications resulting from the proposed research open, machine-readable, and digitally accessible to the public at the time of publication. This includes data that are displayed in charts, figures, images, etc. In addition, the underlying digital research data used to generate the displayed data should be made as accessible as possible to the public in accordance with the principles stated in the Office of Science Statement on Digital Data Management (http://science.energy.gov/funding-opportunities/digital-data-management/). This requirement could be met by including the data as supplementary information to the published article, or through other means. The published article should indicate how these data can be accessed.
3. DMPs should consult and reference available information about data management resources to be used in the course of the proposed research. In particular, DMPs that explicitly or implicitly commit data management resources at a facility beyond what is conventionally made available to approved users should be accompanied by written

SC Funding Opportunity Announcement Language

Merit Review Criteria:

SCIENTIFIC AND/OR TECHNICAL MERIT OF THE PROPOSED RESEARCH

- What is the scientific innovation of proposed effort?
- How does the proposed work compare with other efforts in its field, both in terms of scientific and/or technical merit and originality?
- How might the results of the proposed work impact the direction, progress, and thinking in relevant scientific fields of research?
- What is the likelihood of achieving influential results?
- Is the Data Management Plan suitable for the proposed research and to what extent does it support the validation of research results?

Thank you!

Questions?