



U.S. DEPARTMENT OF
ENERGY

Office of
Science

Data Management Plan Requirements

CENDI

January 8, 2011

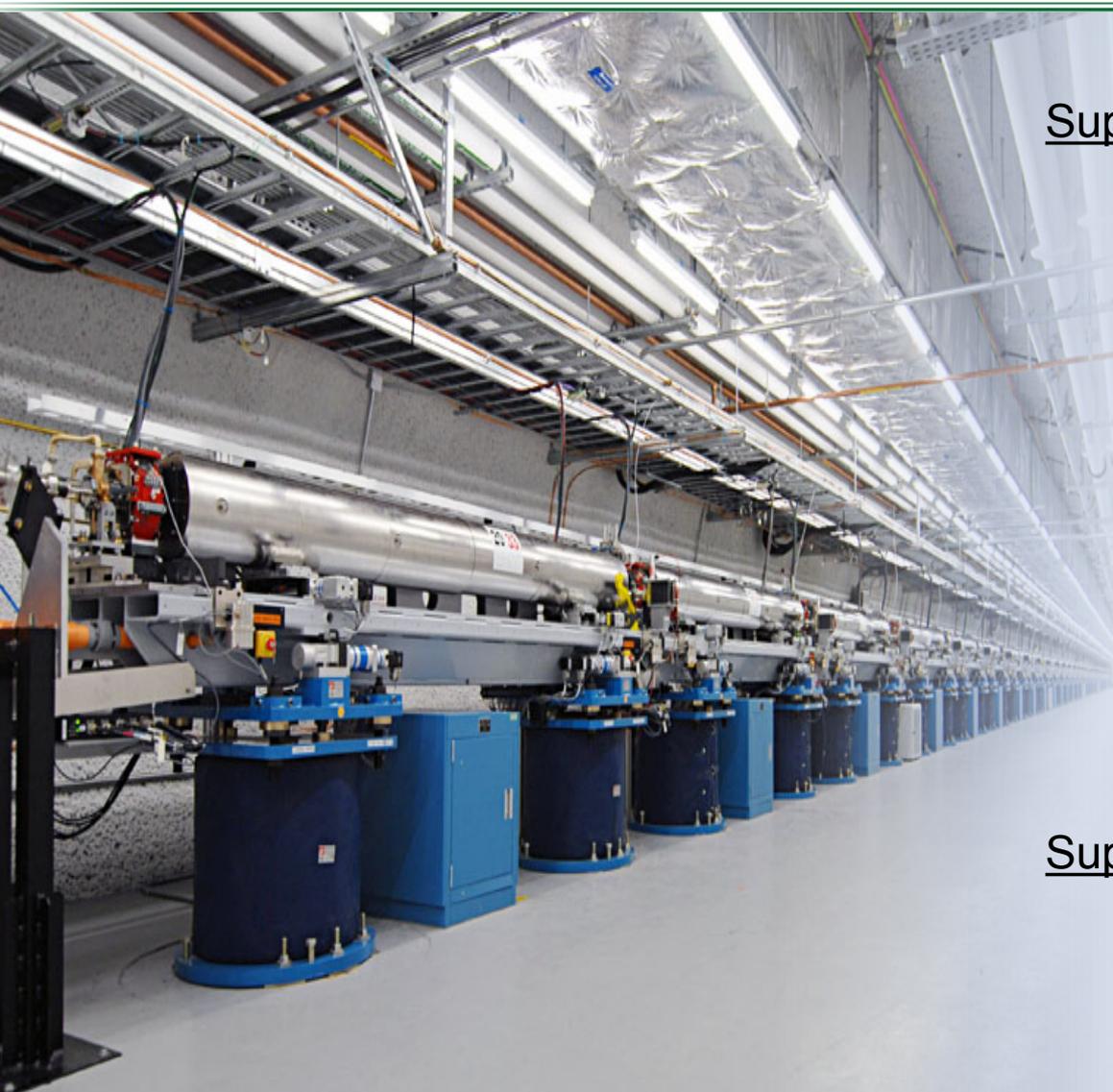
Laura Biven, PhD
Senior Science and Technology Advisor
Office of the Deputy Director for Science Programs (SC-2)
U.S. Department of Energy
Laura.Biven@science.doe.gov

Outline

- **DOE Office of Science 101**
- **Context and History for DMP requirements**
- **DOE Public Access Plan**
- **Office of Science Statement on Digital Data Management**



Office of Science



Support for Researchers

- Supports about 22,000 Ph.D. scientists, graduate students, undergraduates, engineers, and support staff at more than 300 institutions.
- Provides 47% of Federal support of basic research in the physical sciences and key components of the Nation's basic research in biology and computing.
- Supports research that led to over 100 Nobel Prizes during the past 6 decades—more than 20 in the past 10 years.

Support for Scientific User Facilities

- Provides the world's largest collection of scientific user facilities to nearly 28,000 users each year.

Quick-Facts about the DOE Office of Science

**Advanced Scientific Computing
Research**

Basic Energy Sciences

**Biological and Environmental
Research**

Fusion Energy Sciences

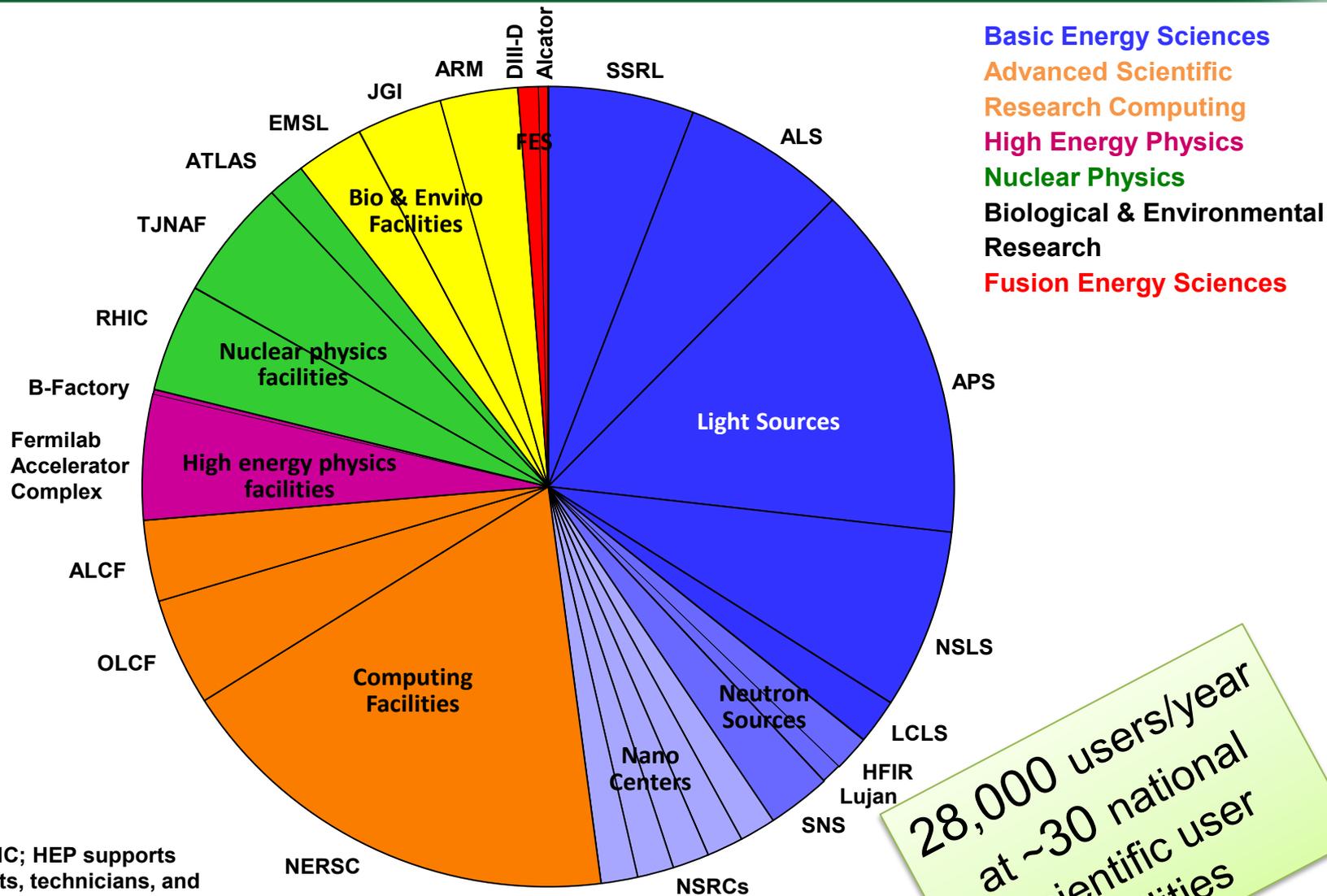
High Energy Physics

Nuclear Physics

Very diverse portfolio with truly big data



Office of Science User Facilities 2013



Basic Energy Sciences
Advanced Scientific Research Computing
High Energy Physics
Nuclear Physics
Biological & Environmental Research
Fusion Energy Sciences

28,000 users/year at ~30 national scientific user facilities

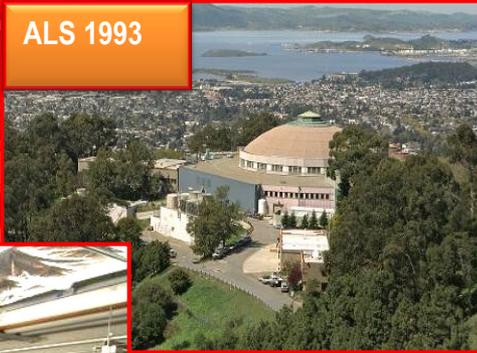
Does not include LHC; HEP supports about 1,700 scientists, technicians, and engineers at the LHC.

Light Sources

SSRL 1974&2004



ALS 1993



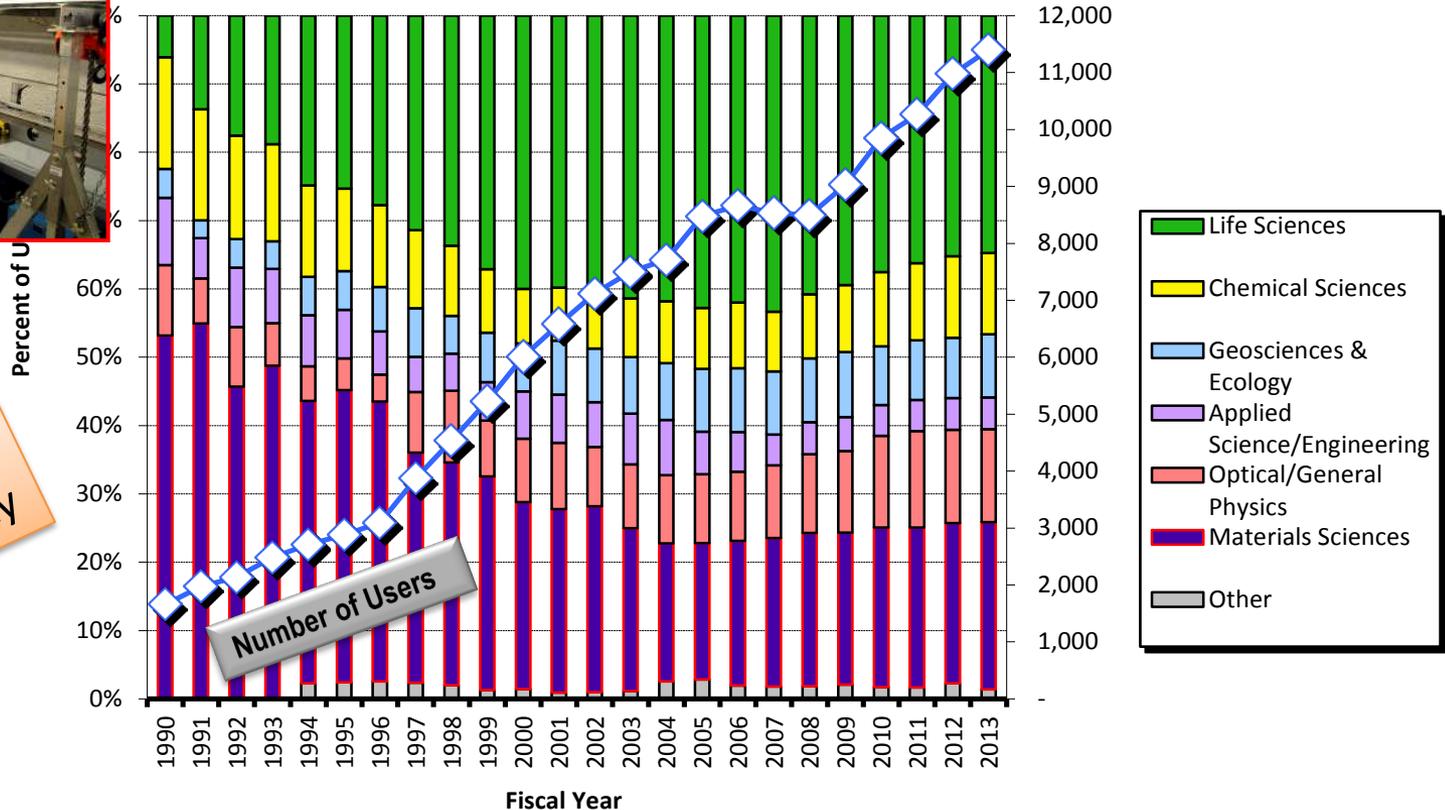
APS 1996



NSLS-II 2014



LCLS 2009



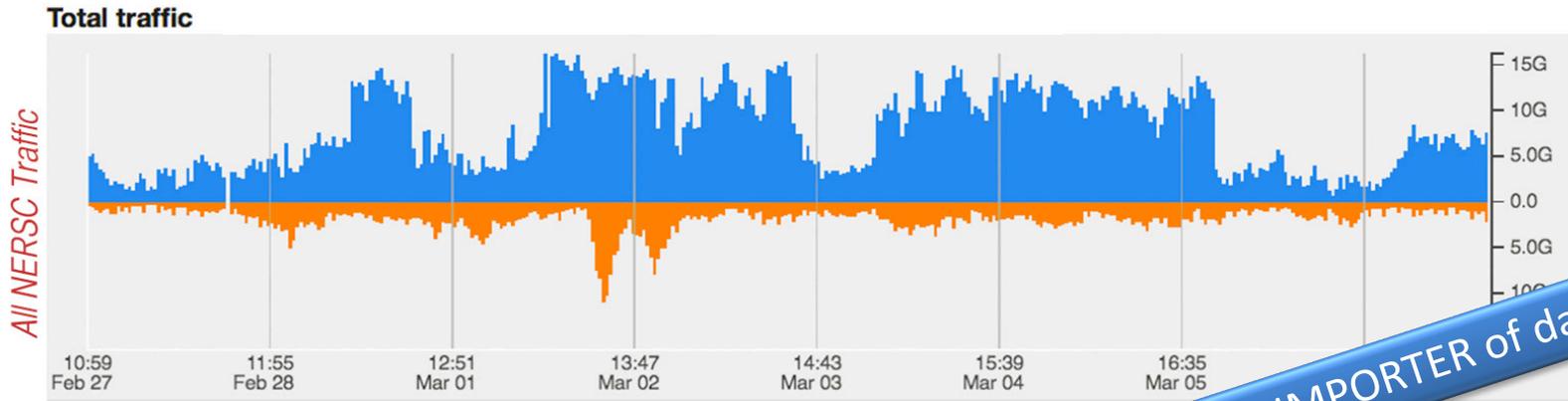
Researchers may use more than one facility

Single Light Source Data Flow Triples NERSC's Network Usage

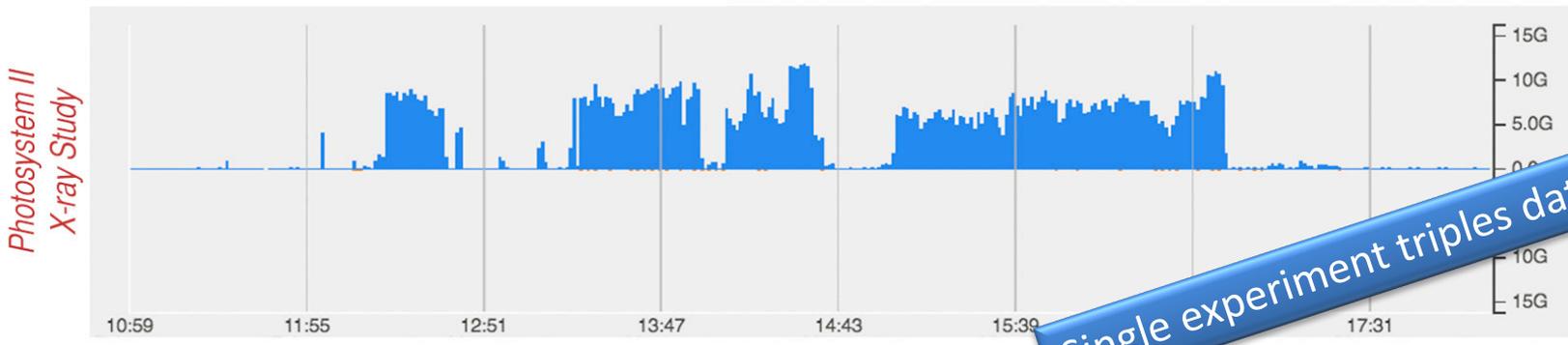


From : Wed Feb 27 10:59:00 2013 To : Thu Mar 7 10:59:00 2013

■ To site ■ From site



Net IMPORTER of data



Single experiment triples data rate

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)**



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, and the scientific community. Such results include peer-reviewed publications and digital data.

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

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

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



U.S. DEPARTMENT OF
ENERGY

Office of
Science

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.





Funding Opportunities

[Grants & Contracts Support](#)[Award Search](#)[Find Funding](#)[Early Career Research Program](#)[Statement on Digital Data Management](#)[Suggested Elements for a Data Management Plan](#)[Frequently Asked Questions](#)[Resources at the Office of Science User Facilities](#)[Acknowledgements of Federal Support](#)

CONTACT INFORMATION

Office of Science

U.S. Department of Energy
1000 Independence Ave., SW
Washington, DC 20585

Statement on Digital Data Management

[Print](#) | [Text Size: A A A](#)[Feedback \[+\]](#) | [Share Page ▼](#)

The Office of Science mission is to deliver the scientific discoveries and major scientific tools that transform our understanding of nature and advance the energy, economic, and national security of the United States. The Office of Science Statement on Digital Data Management has been developed with input from a variety of stakeholders in this mission¹.

Here, data management involves all stages of the digital data life cycle including capture, analysis, sharing, and preservation. The focus of this statement is [sharing](#) and [preservation](#) of [digital research data](#).

Table of Contents

- [Principles](#)
- [Requirements](#)
- [Additional Guidance \(including suggested elements for Data Management Plan\)](#)
- [Additional Requirements and Guidance from Office of Science Program Offices](#)
- [Information about Data Management Resources at Office of Science User Facilities](#)
- [Glossary](#)
- [FAQs](#)
- [References](#)

Principles

The Office of Science affirms that the following principles related to the management of [digital research data](#) directly support fulfillment of its mission.

Office of Science Statement on Digital Data Management

<http://science.energy.gov/funding-opportunities/digital-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.



Office of Science Statement on Digital 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 apply to proposals submitted in response to solicitations and invitations issued after Oct. 1, 2014
- Requirements do *not* apply to applications to use Office of Science user facilities.

Office of Science Statement on Digital Data Management

Requirements

All proposals submitted to the Office of Science for research funding must include 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 Requirement #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.

Office of Science Statement on Digital Data Management

Requirement 2 of 4

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. 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.



Office of Science Statement on Digital Data Management

Requirement 3 of 4

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 Office of Science User Facilities, researchers should consult the published description of data management resources and practices at that facility and reference it in the DMP. Information about other Office of Science facilities can be found in the additional guidance from the sponsoring program.





Funding Opportunities

Grants & Contracts Support

Award Search

Find Funding

Early Career Research Program

Statement on Digital Data Management

Suggested Elements for a Data Management Plan

Frequently Asked Questions

Resources at the Office of Science User Facilities

Acknowledgements of Federal Support

CONTACT INFORMATION

Office of Science
U.S. Department of Energy
1000 Independence Ave., SW
Washington, DC 20585

Statement on Digital Data Management

Data Management Resources at the Office of Science User Facilities

Print | Text Size: A A A

Feedback [+] | Share Page ▼

Follow the links to learn more about data management resources at each of the Office of Science User Facilities. If you do not see your facility of interest listed here, please consult the appropriate [Office of Science Program](#) page.

Advanced Scientific Computing Research (ASCR)

Facility	Host Institution	Data Management Resources
National Energy Research Scientific Computing Center (NERSC)	LBL	Link
Argonne Leadership Computing Facility (ALCF)	ANL	Link
Oak Ridge Leadership Computing Facility (OLCF)	ORNL	Link
Energy Sciences Network (ESnet)	LBL	Link

Basic Energy Sciences (BES)

Facility	Host Institution	Data Management Resources
<i>Light Sources</i>		
Advanced Light Source (ALS)	LBL	Link
Advanced Photon Source (APS)	ANL	Link
Linac Coherent Light Source (LCLS)	SLAC	Link
National Synchrotron Light Source (NSLS)	BNL	Link

Office of Science Statement on Digital Data Management

Requirement 4 of 4

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, and DOE orders and policies. There is no requirement to share proprietary data.
- DMPs will be reviewed as part of the overall Office of Science research proposal merit review process.
 - Additional requirements and review criteria for the DMP may be identified by the sponsoring program or sub-program, or in the solicitation.

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.

Suggested Elements for a Data Management Plan



Funding Opportunities

Grants & Contracts
Support

Award Search

Find Funding

Early Career Research
Program

Statement on Digital Data
Management

**Suggested Elements for a
Data Management Plan**

Frequently Asked Questions

Resources at the Office of
Science User Facilities

Acknowledgements of
Federal Support

CONTACT INFORMATION

Statement on Digital Data Management

Suggested Elements for a Data Management Plan

Print | Text Size: A A A

Feedback (+) | Share Page ▾

The Principal Investigator should determine which data should be the subject of the DMP and, in the DMP, propose which data should be shared and/or preserved in accordance with the [Requirements for Digital Data Management](#).

The following list of elements for a DMP provides suggestions regarding the structure of the DMP:

- **Data Types and Sources.** A brief, high-level description of the data proposed research and which of these are considered [digital research data](#).
- **Content and Format.** A statement of plans for data and metadata description of documentation plans, annotation of relevant software, standards. (Existing, accepted community standards should be used where possible. Where community standards are missing or inadequate, the DMP could propose alternate strategies that facilitate sharing, and should advise the sponsoring program of any need to develop or generalize standards.)
- **Sharing and Preservation.** A description of the plans for data sharing and preservation. This should include, when appropriate:
 - the anticipated means for sharing and the rationale for any restrictions on who may access the data and under what conditions;
 - a timeline for sharing and preservation that addresses both the minimum length of time the data will be available and

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

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

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?

