# Department of Energy (DOE)

## Data types and sources

### The sections in the template outline are based on Suggested Elements of a DMP (see Links tab) provided by DOE, but DMPs are not required to follow this template. Consult the funding solicitation for guidance about when and how to submit a DMP. For the data types and sources suggested element, a brief, high-level description of the data to be generated or used through the course of the proposed research and which of these are considered digital research data necessary to validate the research findings may be included.

*\*Requirement #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 the Protection section). 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.*

*Principles related to the management of digital research data:*

* *- 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.*

*Detailed instructions can be found on the* [***DOE Policy for Digital Research Data Management***](http://www.energy.gov/datamanagement/) *page.*

*Check with your specific office or program to see if there are* [***Requirements and Guidance from DOE Sponsoring Offices****.*](http://www.energy.gov/datamanagement/doe-policy-digital-research-data-management#Requirements%20and%20Guidance)

*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.* ***Research data*** *are defined in regulation (*[***2 CFR 200.315 (e)***](http://www.ecfr.gov/cgi-bin/text-idx?node=2:1.1.2.2.1.4.30.16&rgn=div8)*, continuing the definition from 2 CFR 215 (*[***OMB Circular A-110***](http://www.whitehouse.gov/omb/circulars_a110#36)*) 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.”*

## Content and format

### A statement of plans for data and metadata content and format including, where applicable, a description of documentation plans, annotation of relevant software, and the rationale for the selection of appropriate 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.)

DMPs should reflect relevant standards and community best practices for data and metadata, and make use of community accepted repositories whenever practicable.

## Sharing and preservation

### *[A description of the plans for data sharing and preservation.*

*\*Requirement #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.*

*\*Requirement #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.*

*The section on sharing and preservation 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 any anticipated 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);*
* *- cost/benefit considerations to support whether/where the data will be preserved after direct project funding ends and any plans for the transfer of responsibilities for sharing and preservation;*
* *- whether, when, or under what conditions the management responsibility for the research data will be transferred to a third party (e.g. institutional, or community repository);*
* *- any other future decision points regarding the management of the research data including plans to reevaluate the costs and benefits of data sharing and preservation]*

The Principal Investigator will manage data access and use during active project work. Upon completion of the project, the Principal Investigator will evaluate, distill, and select the data for archiving. These decisions will be based on publication usage, storage location, space availability, quality, uniqueness, broader need/usage, relationship to other data sources, reproducibility, sponsor requirements, privacy, confidentiality and intellectual property rights.

Upon completion of the project, the Principal Investigator will evaluate, distill, and select the data for archiving. These decisions will be based on publication usage, storage location, space availability, quality, uniqueness, broader need/usage, relationship to other data sources, reproducibility, sponsor requirements, privacy, confidentiality and intellectual property rights.

The Principal Investigator will provide metadata typically associated with the kind of data generation techniques described above. Some metadata may be generated as part of the data generation process. Other metadata may be generated by programmatic or manual means. The Mines Research Support Services group will advise the PI on descriptive metadata requirements. At a minimum, for each subset of data archived, metadata descriptions and keywords will be provided to the Mines institutional repository in order to provide discovery of discrete subsets of data or for the data as a collective/project whole. Access is then granted per archiving and sharing mechanisms described above.

For those data being archived, *[Choose a data archive mechanism below and include one and/or all of the following paragraphs about data archiving as appropriate.]*

*[****OPTION******1****) If archiving data to the Mines Institutional Repository….]* electronic data will be archived to the Mines Institutional Repository (<http://publish.mines.edu>). The repository is a service jointly provided, operated, and managed by Mountain Scholar, which is a consortium of Colorado and Wyoming academic institutions and is housed at Colorado State University. That storage facility will (a) store data in redundant, multiple, secure locations, (b) provide retrieve only access to the data to ensure integrity, (c) include data, metadata and related materials (graphs, pictures, links to dissertations and publications, etc.), and (d) be accessible by Mine’s Research Support Services to facilitate access. The storage media will conform to industry-standards for stability and accessibility.

The Mines Institutional Repository will store data and make it accessible for a minimum of 3 years *[(or as contractually required)]* after the project or final publication has ended and will conform to academic standards for stability and accessibility. *[Please note: This option requires you to pay for archival storage at $1000/TB and is best for small to medium datasets.*

*[****OPTION 2****) If archiving data on your own local lab server, department server, Mines campus server or purchased cloud-based solution….]* electronic data will be archived to a computer server in *[choose all that are appropriate: local lab, department, Mines campus computing or purchased cloud-based solution]*.

The Principal Investigator (PI) will continue to use devices and methods similar to those described during the active project storage phase. The Principal Investigator will ensure archiving, preservation and accessibility of the data for a minimum of 3 years *[(or as contractually required)]* after the project has ended/final publication has occurred.

Access will be provided to Mines faculty, students and, as appropriate, researchers at other institutions via *[describe how: email request, website* (*If a website is also a result at the end of the project, describe how it may be used to help with data archiving, sharing and public accessibility*)].

As an additional means of public access and dissemination and as an institutional record, appropriate sample data, a data product or overall project description will be provided to the Mines Institutional Repository (<http://publish.mines.edu>). The repository is a service jointly provided, operated, and managed by Mountain Scholar, which is a consortium of Colorado and Wyoming academic institutions and is housed at Colorado State University. That storage facility will (a) store data in redundant, multiple, secure locations, (b) provide retrieve only access to the data to ensure integrity, (c) include data, metadata and related materials (graphs, pictures, links to dissertations and publications, etc.), and (d) be accessible by Mine’s Research Support Services to facilitate access. The storage media will conform to industry-standards for stability and accessibility.

 *[Please note: this option requires you to perform and manage the archival data storage and budget consideration if buying your own lab server. And it expects that a data sample/ data product/ project description be deposited in the Mines Institutional Repository (there is no cost for samples less than 10GB)]*

*[****OPTION 3****) If you will be using another institution as the data archive (e.g. NCAR HPSS) write your own paragraph about how that institution archives and preserves data for future research and public accessibility. Provide appropriate web links to their policies or repository information. And include the paragraph below]*

As an additional means of public access and dissemination and as an institutional record, appropriate sample data, a data product or an overall project description will be provided to the Mines Institutional Repository (<http://publish.mines.edu>). This data sample will have associated metadata and contact information for getting further access to the broader dataset. This record at the institutional level allows Mines to keep the Mines Data Inventory up to date. *[Please note: this option requires you to perform and manage the archival data storage and budget consideration for it if purchasing cloud services or using an external repository. And it expects that a data sample/ data product/ project description be deposited in the Mines Institutional Repository (there is no cost for samples less than 10GB)]*

The existence and availability of the data sets will be communicated to sponsors in progress reports and to research peers at conferences, and they will be referenced in publications. Availability may be subject to appropriate embargo periods determined by copyright, sponsor agreements, patent and licensing provisions. *[Describe any sensitive data that may need additional confidentiality restrictions, and specify the details of those restrictions.]*

## Protection

### A statement of plans, where appropriate and necessary, to 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; and avoid significant negative impact on innovation, and U.S. competitiveness.

*\*Requirement #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.*

Availability of the data may be subject to appropriate embargo periods determined by copyright, sponsor agreements, patent and licensing provisions. Under no conditions will the proposed project disseminate data not directly collected or generated by the project. There is no intent to collect sensitive data as part of this project.

*[If there are products produced or utilized during the project that cannot be publically released, they should be addressed in this section]*

## Rationale

### A discussion of the rationale or justification for the proposed data management plan including, for example, the potential impact of the data within the immediate field and in other fields, and any broader societal impact.

\*Requirement #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 the Protection section). 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.

## Software & Codes

### For office and program specific guidance, refer to the Requirements and Guidance from DOE Sponsoring Offices webpage.

***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.*

**Data budget**
The total amount of electronic data to be archived is estimated to be *[Estimate size in GB or TB]*. To cover the cost of archival storage and access, electronic materials will be stored and made accessible at a total cost of *[choose an appropriate cost scenario]*

*[****1****) If using the Mines Institutional Repository as the data archive mechanism (does not include the use of the repository when it holds only a data sample/product under 10GB): use the multiplier of $1000 per TB (****ensure that the proposal budget reflects this cost as an incidental direct cost)***

*2) If using an external repository: determine if the external repository has costs that must be included*

*3) If buying other storage, cloud services or local lab servers, be sure your budget form reflects those costs in appropriate areas*

*If you need additional data management staff expertise listed within your proposal, you may include and budget for Mines staff member from Research Support Services.]*

*If you are a large data user (>10GB), contact Research Support Services to discuss the potential cost of additional active storage.*

*\*Archival storage will be charged in the final year of your project, and covers the cost of data preparation, administration, and archival storage. Storage architecture is compliant with industry standards (primary storage + replication +backup)*

***Need Additional Help?*** *If you have questions about writing data management plans, please contact Torey Battelle at Research Support Services, extension 3393 or* *battelle@mines.edu*