

Department of Energy: Office of Science

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. 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.

Detailed instructions can be found on the [Office of Science Statement on Digital Management](#) page.

Check with your specific program to see if there are [specific program requirements](#).

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 \(e\)](#)), continuing the definition from 2 CFR 215 ([OMB Circular A-110](#)) 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 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.

View information about the [data management resources](#) available at the Office of Science User Facilities.

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.

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.

Personally Identifiable Information

For proposals with Human Subjects Research (HSR), including research involving Personally Identifiable Information (PII), an appropriate research protocol will need to be approved by the appropriate DOE Institutional Review Board (IRB) or an external IRB with an approved federal wide assurance. Follow the instructions of the research solicitation to determine whether or not the data management aspects of this protocol should be included in the DMP. At a minimum, the DMP should acknowledge the type of HSR and/or PII involved and give a projected timeline for IRB approval. Information regarding DOE requirements for HSR and research involving PII, including how to obtain IRB approval, can be found at this link: [DOE Human Subjects FAQ](#).

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.

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.

Software & Codes

Both the Advanced Scientific Computing Research and Fusion Energy Sciences program areas address software and codes. Program specifics are listed below.

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. [See full definition](#)

Advanced Scientific Computing Research

ASCR considers software to be a data artifact that is covered by the Office of Science Statement on Digital Data Management. Research funded by ASCR may produce several different types of

software (e.g., those that encode a mathematical algorithm, implement a specific tool or service, become part of a base Operating System). Other funded research may result in the creation or consumption of data that could be used to analyze application or system behavior. Software and data created by funded research must be released with sufficient descriptions to facilitate the validation of research results.

*Requirements - The DMP must briefly discuss the applicable issues listed below:

- * The Open Source License to be used, if applicable;
- * If executable versions of the software will also be released, and if so what format will be used;
- * How software can be found and accessed and the length of time the software will be publicly available;
- * How any proprietary 3rd party software or libraries are used in the creation of this software;

Office of Fusion Energy Sciences

The Office of Fusion Energy Sciences does not have any additional requirements, but they do provide additional guidance of software and codes.

Data management plans for codes should address the elements outlined in the SC Statement. With respect to codes, FES specifically:

- - Encourages that codes developed with FES funding and used to analyze experimental data and/or to perform simulations be made available to the broader community via Open Source licensing.
- - Recognizes the right of the code developers to establish a reasonable exclusive use period for the developed software; in addition, recognizes that sharing may be limited to stable versions of the code(s) and not to the currently-under-development “research” versions of the code(s).
- - Recognizes that there may be code sharing limitations driven by proprietary interests, export control issues, intellectual property rights, and issues affecting US competitiveness.
- - Acknowledges that code sharing does not create obligations for the code developers to provide software support, unless there is a collaboration agreement between the parties.
- - Understands that code sharing should respect existing user license agreements established by the code developers