Data Management Policy

The Gulf Research Program funds research and other activities that generate and disperse knowledge. Most funded activities produce information products. Information products include databases and data sets; documents (scholarly publications, reports, workshop summaries, etc.); images (maps, data visualizations, photographs, etc.); digital models, simulations, and/or software code; multi-media curricula for education and training (video and/or online tutorials, manuals and handbooks, etc.); and other media and communication platforms. Data are one of the most common information products.

Guided by its mission and three goals, the Gulf Research Program (GRP) promotes the responsible management of scientific information that maximizes the value of data. The program requires management practices that curate information products for future use and that make information widely available for public discovery and access. For example, the program encourages practices that enhance information discovery and accessibility (e.g., sharing). Sharing research data enables data verification and reproduction of results, as well as the re-use of those data and results. Moreover, sharing allows scientists, engineers, and health professionals to expedite the translation of research results into new knowledge, products, and processes that can benefit society.

In developing its own Data Management Policy, the Gulf Research Program looks to the White House’s Open Data Policy and the data management policies of other major funding agencies and institutions. Grantees are expected to make information products generated by program-funded projects, initiatives, and activities publicly discoverable and widely accessible in a timely manner, subject to any Institutional Review Board (IRB) or legal restrictions. The program requires that program-funded projects, initiatives, and activities create detailed, machine-readable metadata for information products and deposit metadata in a digital repository or data center to enable discovery and sharing. In addition, program-funded projects, initiatives, and activities must deposit information products, where appropriate, in a digital repository, data center, and/or other suitable curation facility that facilitates access to these products and ensures long-term curation. The Data Management Policy will be updated as the program grows and evolves.

Data: Definition and Rights

The Gulf Research Program follows the federal government’s definition of data in the Office of Management and Budget (OMB) 2 Code of Federal Regulations (CFR) Section 200.315: “…the recorded factual material commonly accepted in the scientific community as necessary to validate research findings.” This definition includes both original data (observations, measurements etc.) and metadata (title, date(s), collection methods, experimental protocols, statistical methods, etc.), as well as software or computer code that is required for replication, etc.

Although the grantee shall retain all rights in research data, the grantee shall provide timely and unrestricted
access to the data to the Gulf Research Program and the U.S. Government. Without limitation, the U.S. Government and the Gulf Research Program shall have the right to (1) obtain, reproduce, publish, or otherwise use the research data first produced under this Grant, and (2) authorize others to receive, reproduce, publish, or otherwise use such data for the grantor’s or U.S. Government purposes. Unless agreed upon between the Gulf Research Program and grantee, the Program expects the timely release and sharing of data to be no later than the acceptance for publication of the main findings from the final dataset or 1 year after the project end date, whichever comes first.

Requirements

All full proposals submitted to the Gulf Research Program must include a data management plan that does not exceed 1,500 words. In the unlikely case in which a proposed project anticipates no collection, production, or generation of information products (i.e., no data, data sets, or other information products), the full proposal must contain a management plan that states, “This project will not result in the production of data or other information products” with a short description as to why no data will be created.

In collaborative proposals or proposals involving sub-awards, the lead applicant is responsible for the information or data management plan for the entire project (i.e., The plan must cover all data types that collaborators plan to collect and all other information products generated over the course of the project.). The lead applicant is also responsible for reporting on the management of all project information products and the accessibility of these products in the Annual and Final Grant Reports.

The data management plan should describe how the applicant will manage and disseminate program-funded information products (data and other products) in sufficient detail to enable evaluation of the plan during the merit review process. The Gulf Research Program staff will monitor adherence to the proposed data management plan.

Data Management Plans

The Data Management Policy strongly encourages applicants and their project personnel to practice information and data management as a vital part of the entire research project process, from project design to completion and dissemination. The policy directs applicants and project personnel to manage project information and data through all data life-cycle stages: planning, collecting and assuring (quality and security), describing (i.e., creating metadata and documentation throughout all life-cycle stages), processing and analyzing, preserving, and publishing, curating, and sharing data.

The Gulf Research Program recognizes that different disciplines may have their own best practices and standards for management of information products, particularly data. Moreover, the program understands that accepted norms will likely change as collaborations among disciplines increase. Therefore, each data management plan should be appropriate for the information products that a project expects to generate. The management plan should adhere to 1) data management best practices that are widely adopted across many disciplines, or 2) those data management standards and best practices that are specific to the area(s) of research or activities proposed.
1. Data and Other Information Products: Planning

   a. Describe the kinds of data that the proposed project will create or capture (e.g., environmental or ecological data, oceanographic data, climatic data, health or social well-being data, oil and gas safety management, risk assessment data, quantitative modeling data).

   b. Identify the different formats for data capture and storage (e.g., tabular data sets, relational databases, geospatial or visual media [i.e., maps], simulations and related data, software code).

   c. If the proposed project involves the use of existing data, please identify the original source of the data. Describe when and how the existing data were collected. Describe any relevant sharing arrangements.

   d. If the proposed project involves the use of proprietary data, please describe the data and the permissions obtained to use these data.

   e. If the proposed project involves the creation or use of confidential or other sensitive information, please describe the data and reasons for protection. When working with sensitive information, the applicant must adhere to IRB policies, institutional guidelines, and other professional policies and best practices as applicable. For projects involving human subjects, the applicant is responsible for providing evidence of IRB approval or justification for exemption from human subjects regulation [see 45 CFR 46.101(b)].

   f. Describe the project’s quality control and data validation procedures.

   g. Identify other information products that the proposed project will create (e.g., documents, images, software and software code).

2. Short-term Management: Collection and Processing

   a. Identify the hardware, physical facilities, and cyberinfrastructure that project personnel will use to capture and store data and other information products (e.g., images, audio and video recordings) during data collection and processing.

   b. Identify software that project personnel will use to capture, store, and process data and other information products.
c. Describe procedures and protocols for managing and storing data and other information products during the project (e.g., organizing and aggregating data, adding newly collected data, altering and correcting data, version control, daily back-ups on and off site).

d. If the proposed project involves confidential or other sensitive information, describe security procedures and protocols. Describe provisions for protection of privacy, intellectual property rights, other rights, and other security issues as appropriate.

e. Identify the persons responsible for maintaining and managing data and other information products during collection and processing. If the proposed project involves confidential or other sensitive information, identify persons who will have access to these data and persons responsible for maintaining security.

3. Metadata: Describe the Data, Data Collection, and Data Processing

a. Identify the metadata that project personnel will create to document the project, the project’s data and other information products.

b. Identify the metadata and any other documentation (e.g., guides and manuals) that project personnel will create to describe data collection methods and data processing/analysis.

c. Wherever possible, ensure that metadata are created and stored in machine-readable formats (e.g., XML and JSON).

d. Describe the management and storage of metadata.

4. Data Sharing

a. Identify policies and procedures that will govern the timely release of and access to project data, other information products, and related metadata.

b. Describe plans for making project data, other information products, and associated metadata discoverable and available to others (e.g., other researchers, decision makers, and the interested public).

c. Describe the format for citing data products in publications (e.g. [data originator last name], [data originator first name], [data co-originators]. [title of the dataset] [year dataset registered/published]. Distributed by: [data repository name]. doi#) so they can be more easily discovered by others.

d. If the project involves confidential or other sensitive information, refer to IRB and/or other professional policies and guidance as applicable. Describe the short-term and long-term management of restricted access to sensitive data and other information products.

5. Long-term Management: Curation and Accessibility

a. Identify a curation facility or facilities (e.g., digital repository and/or data center) in which project data, other information products and metadata will be deposited. The facility or
facilities should ensure the long-term curation and wide accessibility of project data, other appropriate information products, and associated metadata.

b. Describe the timeline and process for transitioning from short-term management to long-term management of data and other information products.

c. Describe procedures that will govern the continued, long-term management of data, other information products, and metadata.

d. Describe policies and procedures that will govern access to project data and other appropriate information products.

6. Data Management Budget

a. Include an allocation in your budget for data management activities.