

Data Management Overarching Policy Working Group

This focus group is no longer active, it was active in 2013. See the [DMWG Focus Groups](#) page to access current and past Data Management Focus groups.

Purpose

This Working Group will lead the development of Survey-wide policies for scientific data management that will become part of the Survey Manual. These policies will clarify the responsibilities of all USGS employees and identify relevant guidance, processes, and resources.

USGS Champions

Kevin Gallagher, Associate Director Core Science Systems

Alan Thornhill, Associate Director Office of Science Quality and Integrity

Linda Gundersen, USGS Emeritus

Background

Data management is an essential part of doing science. Documented procedures and the preservation of data ensure research integrity and enable reproducibility of research results. Some scientific disciplines and Agencies have developed their own data management standards, technologies, and best practices to meet these requirements, but these are not universally known or applicable.

The Federal government has data management requirements that are consistent with good science. These include:

- Federal Records Management policies addressing documentation of activities and preservation of data.
- Federal Information Security Management Act (FISMA) requirements to mitigate IT risks that might impact the security, accessibility, and long-term preservation of data.
- Open Government Initiative policies that facilitate accessibility of scientific data as well as interpretive results.
- Data Quality Act requirements for rigorous and documented quality assurance of scientific data.

In the academic research sector, technological innovation and data-intensive scientific discovery have motivated new data policies. Some journals now require public release of data that are the basis of published scientific papers. The National Science Foundation (NSF) has established policies that require proposals for funding to include a data management plan that states arrangements for data release and long-term storage. NSF has also begun funding data centers (such as the Biological and Chemical Oceanography Data Management Office) and data management systems (such as Rolling Deck to Repository) to support effective data management.

Multiple groups within the USGS have roles in scientific data management:

- Scientists who assemble or create data and data products in the course of their research activities
- Core Science Systems, including: Core Science Analytics and Synthesis; National Geospatial Program; USGS Libraries, especially the Publications Warehouse; and National Geological and Geophysical Data Preservation Program
- Records Management Program
- Fundamental Science Practices unit within the Office of Science Quality and Integrity
- Community for Data Integration, and especially its Data Management Working Group
- Science Data Coordinator Network
- Mission areas, programs, and regions that maintain employees, activities, and systems for data management

While many of these groups are formally and informally working together on different aspects of data management, their work is not guided by a set of common policies.

Scope

The Working Group will consider USGS policies for data management activities related to USGS scientific data, where these terms have the following meanings:

- **USGS scientific data** includes all non-interpreted observations and measurements that are made or used in the course of USGS research, and also derived, integrated, or interpretive products that are expressed in machine-readable formats, such as GIS layers, spreadsheets, and netCDF files.

- **Data management** includes all stages of the USGS data management lifecycle^[1] (plan, acquire, process, analyze, preserve, publish/share) and related activities (data documentation, quality management, and data protection).
- **Policies** are the guidelines developed by an organization to govern its actions, including definitions of roles and responsibilities, the scope or spheres in which employees are free to make decisions, the results that must be achieved, and the resources available.

Operations

The Working Group will engage in these activities:

- 1 Articulate the foundational principles for USGS scientific data management.
- 2 Map the existing policies and the relationships among them.
- 3 Plan for Survey Manual Chapters that need to be written or revised to keep pace with changes in technology, science practices, and the development of common standards.
- 4 Convene teams to write and revise Survey Manual Chapters.
- 5 Articulate the resource requirements of the new policies to USGS leadership.
- 6 Monitor the review and approval of the chapters.

Criteria for success

- All USGS employees who work with scientific data:

Are aware that there are USGS policies for how we manage scientific data.

Know where to find these policies and understand their relevance and applicability.

Have access to guidance, processes, and resources that support compliance with these policies.

- USGS research projects are successful in creating and applying scientific data to achieve their goals.
- Scientists and other customers consider USGS scientific data to be accessible, trustworthy, and useful.
- Managers of USGS programs and regions:

Are aware of their responsibilities for maintaining capabilities to manage scientific data.

Can plan staffing and funding appropriate to these capabilities.

Can evaluate the success of their capabilities toward meeting USGS policy.

- USGS is transparently in compliance with laws and regulations related to scientific data management.

^[1] See <http://www.usgs.gov/datamanagement/why-dm/lifecycleoverview.php>