

U.S. Geological Survey (USGS): U.S. Geological Survey (USGS)

1. Project and Contact Information

What is the name of the project?

Include any identifiers related to the project (e.g. Project ID, Funding ID etc).

What is the name of the USGS Center/Program that oversees the project?

Include contact information (email, phone, address).

Summary description of the project.

Include reason why the data is being collected.

What is the project start date?

Start date.

What is the project's expected end date?

Estimated end date. This field can be updated as needed.

Are there additional information available?

Include any web links with more information related to the project, if applicable.

Who is the main point of contact for the project and its data?

Also list any alternate points of contact, if any.

Are there collaborating/funding agencies and organizations?

Who are they and who are the main points of contact?

2. Plan and Acquire

How will the data be acquired?

Are they newly collected or using existing datasets?

If acquiring existing datasets, include more information.

Include the name, format, a persistent identifier, and source citation, if any. Are there any restrictions or agreements such as Memorandum of Understandings (MOUs) for use and storage?

If collecting new data, include more information.

Are there special processes or procedures for collecting the data (e.g. licenses, permissions, equipment, software)?

What is the estimated volume of the data collected, transformed, and/or generated?

For example in megabyte (MB), GB, TB, or PB.

Will the data be static or frequently updated?

If frequent updates, describe how frequent (e.g. Continuously, weekly, annually, irregular etc)

Are the appropriate equipment and staff resources accounted for in the budget?

Include estimated time and cost for such data management activities.

3. Describe/Metadata and Manage Quality

How many new datasets will be created?

List the anticipated title of each dataset.

What are the data types and formats, in which the data will be maintained?

Open data formats such as csv, tiff, mp3, are required.

Briefly describe the data processing steps or provide the scientific workflow.

Also identify any software or technology needs where applicable.

How will the metadata for each dataset be created?

Who will be responsible for the metadata creation and update? Include their contact information.

Which metadata standard will be used to describe each dataset?

For USGS, FGDC-CSDGM or ISO 19115 series are required.

What procedures will be used for ensuring data quality (QA/QC)?

If using a known standard or protocol, include the citation source.

4. Backup/Secure and Preserve

Where will the data be stored in the short-term?
Is it properly secured, backed up, and environmentally controlled?

What will be the approach for routine backup of the data?
Include the frequency, duration, software, and media information. Will the data be stored in multiple places and on different media types (recommended minimum of 3 copies with 1 stored in an offsite location)?

Describe any potential access restrictions.
For example if the data contain Personally Identifiable Information (PII). Please include any practices to ensure access will be restricted.

What will be the final format of the data product?
Will there be any software needs? Will the data format be appropriate for long-term preservation? Open data formats such as csv, tiff, mp3, are required.

Where will the data and metadata be preserved in the long-term?
And which funding Program if in collaboration, will be responsible for the preservation of the data? Who will be the point of contact?

If costs are associated with long-term storage, how will they be provided for?
Are there agreements made for the preservation of the data and metadata?

5. Publish and Share

How will the data be shared and made available to the public?
For example a web page, system or application, data portal, repository, USGS Data Series, etc. Are there data release policies that need to be followed?

Will there be access or use restrictions on the data?
For example for sensitive data, restricted data, privacy, software with license restrictions, etc. Provide justification for the restriction citing any policies or legal reasons.

How can someone overcome any access restrictions?
For example are the following required? Fees, non-disclosure statements, special authorization, data embargo or hold, MOUs/MOAs.

Identify any anticipated publications or electronic outlets resulting from the data.
For example, peer-reviewed articles, information/fact sheets, web pages. If a USGS publication, indicate type (e.g. Open File Report, Provisional Release etc).

Where will metadata be stored to enable data discovery by the public?
USGS requires that your metadata must be available for harvest by the USGS Science Data Catalog. Contact sciencedatacatalog@usgs.gov for more information.

How and where will you obtain a digital object identifier (DOI) for the data?
USGS provides a Digital Object Identifier Creation Tool available at <https://www1.usgs.gov/csas/doi/>