NASA Center for Climate Simulation
Data Management Plan (DMP)

Instructions are available at the end of this form.

Project Description:

Project Name _____

Principal Investigator _____

Curator(s) _____

Computational Project ID _____ Organization _____

Project Status (On-going/Directed Funding or Finite/Grant Funding) _____

Brief description of work:

Brief description of input, intermediate, and final data:

Input data are:	
Intermediate files are:	
Final products are:	

A workflow diagram can be attached below. This is optional but encouraged.

Data Storage

Where is the data located now? (Discover, Explore, CSS, or an external location)

Input Data

Estimated volume (GB, TB)	
Number of files/directories (inodes)	

Intermediate Data

Estimated volume (GB, TB)	
Number of files/directories (inodes)	

Final Data

Estimated volume (GB, TB)	
Number of files/directories (inodes)	

What is the expected growth rate for the data, either percentage or volume (est. by month or year)?

You may detail additional storage requirements here:

Access and Sharing

Please classify the following data types as either Public, Private, or Controlled (e.g., ITAR, EAR):

Input Data _____ Intermediate Data _____ Final Data Products _____

Project Space Information

Please provide a unique, short name for the top-level directory: _____

If applicable, provide a short name for the CSS directory: _____

What userid should own it? _____

What group should own it? _____

If a new group is required, list the users who need to be members:

Do group members need write access permissions? (yes/no) _____

Do world read permissions need to be enabled? (yes/no) _____

High Level Computational Requirements

Parallel Computation? (yes/no) _____

Which NCCS system(s) will need to access the data? (Discover, Explore, specific ADAPT VMs):

List all licensed software that apply (IDL, MATLAB, others, none):

NCCS Data Sharing Services

Will data need to be made available to the public via the DataPortal? (yes/no) _____

If so, which services (HTTPS, GDS, TDS, ESGF, ArcGIS, FLUID, others)? _____

Disposition

When will the following data types be deleted from disk? If none of the below examples, provide a justification (e.g., sharing with other NCCS projects, need for future reprocessing):

1. Input Data (after processing, after final QC, after # of years, end of project, etc)

2. Intermediate Files (after each run, after final QC, after # of years, end of project, etc)

3. Final Products (after external archiving, after final QC, after # of years, end of project, etc)

Archive

Where will the final products be archived (NASA DAAC, PI's System, other)?

How will the final products be migrated to the archive or cloud?

Miscellaneous

Please provide URL(s) that describe the data to be stored on CSS.

What is the process for managing the data of project members who leave?

DMP Instructions

Description of the project:

- Please provide some basic information about the project like it's title, organization, expected lifetime, and the alpha-numeric computational project ID. The PI is usually the same as the computational project PI and the curator is usually the person(s) who produces the data and to whom questions about it will be directed.
- **Brief Description:** In a few sentences, what does the project accomplish.
- **Description of data files:** Describe the input, intermediate, and final data for us. It may be observational data, climate model output, restart files, in-situ data, etc.
    1. **Input Data:** Data from another project used to generate new datasets.
    2. **Intermediate Data:** Data generated during software runs that may need post analysis and/or quality checks.
        a. Not Permanent.
        b. Not to be shared publicly.
        c. Could be restart files, research results, or temporary files.
    3. **Final Data:** Data used for publications or shared with the science community and/or collaborators.
        a. Permanent.
        b. May be shared publicly.
        c. Could be multiple data formats.
        d. Could be input to other science programs/projects.
- **Workflow Diagram:** Click the image field then select "browse" to attach a diagram.

Data Storage

- Please provide information on where the project's data currently is, the total volume (GB, TB) of the data, and the rate at which you expect the data to grow.
- **Number of files/directories:** This can be estimated by running: "find /path/ | wc -l"

Access and Sharing

- Admins need to know some descriptive information and which group/user will need read and/or write access to setup the project space on the system.
- **High Level Computational Requirements:** Specify all systems or software that apply.
- **NCCS Data Sharing Services:** Please contact support@nccs.nasa.gov if you have questions about sharing public data.

Disposition

- Please indicate where final products will be archived. The NCCS is not a designated NASA public repository. If you need your data to be officially archived with a DOI, contact the Goddard DAAC, GES DISC, or another NASA DAAC.
- **Miscellaneous:**
    - The NCCS uses these URLs to provide background on datasets for other customers and to let them know who to contact for science questions.
    - Please describe who will take ownership of this data if the PI or curator leaves the project while it is active. A position title is sufficient.