ESDIS Library Number: ESDIS05161

Earth Science Data Systems (ESDS) Program, HQ SMD

Data Management Plan (DMP)
Template for Data Producers (DP)

Version 1.1

Headquarters
Washington, DC

National Aeronautics and
Space Administration

Earth Science Division (ESD) Data Management Plan (DMP) Template for Data Producers (DPs)

Each organization funded by NASA to produce, archive and/or distribute data is required to prepare a Data Management Plan (DMP) at the time of funding and to maintain the DMP as a living document by reviewing it periodically and ensuring that the plan is up-to-date. To meet these requirements for all components of the Earth Observing System Data and Information System (EOSDIS), the Earth Science Data and Information System (ESDIS) Project has developed templates tailoring the DMPs to the specific types of activities that each component is required to perform. The attached template can be used by NASA-funded Earth science Data Producers (DPs) to develop Data Management Plans (DMPs).

Change History Log

Revision	Effective Date	Description of Changes
1.0	1/31/2019	Baseline document
1.1	6/23/2020	Repaired broken links; edited for punctuation and grammar

Data Management Plan (DMP) Template for Data Producers (DPs)

Each organization funded by NASA to produce, archive and/or distribute data is required to prepare a Data Management Plan (DMP) at the time of funding and to maintain the DMP as a living document by reviewing it periodically and ensuring that the plan is up-to-date. The requirements calling for DMPs are documented in appropriate sections of NASA Procedural Requirements (NPR) 7120.5 and 7120.8. NPR 7120.5 applies to Space Flight Programs and Projects, and NPR 7120.8 applies to Research and Technology Programs and Projects. Data Management Plans are required to include information on the acquisition, processing, analysis, distribution and archiving of data. In addition, in February 2014, NASA created a plan titled "NASA Plans for Increasing Access to the Results of Scientific Research – Digital Scientific Data and Peer Reviewed Publications". This plan covers the requirements for NASA funded researchers – "…all extramural researchers receiving NASA grants, cooperative agreements, and contracts for scientific research and intramural researchers develop data management plans, as appropriate, describing how they will provide for the long-term preservation and access of scientific data in digital format." In order to meet these requirements for all components of the Earth Observing System Data and Information System (EOSDIS), the Earth Science Data and Information System (ESDIS) Project has developed templates tailoring the DMPs to the specific types of activities that each component is required to perform.

The purpose of this template is to provide the Data Producers (DPs) with guidance on the content of DMPs. DP, in the context of this template, refers to those responsible for production of data delivered to a Distributed Active Archive Center (DAAC) for archival and distribution. The data producer refers to a manager of a data production system: a Science Investigator-led Processing System (SIPS), Science Data System (SDS), Research and Analysis (R&A) or Application Program-funded Principal Investigator (PI) controlled system, or a Making Earth System data records for Use in Research Environments (MEaSUREs) data development program. The development, maintenance, and management of DMPs are the responsibility of the respective manager as specified in Table 1.

Table 1: Data Management Plan (DMP) Responsible Parties

Data Producer	*Responsible Person for DMP production and delivery*	*Delivery of DMP to for approval*
Science Investigator-led Processing System (SIPS)	SIPS Manager	ESDIS SIPS Manager
Science Data System (SDS)	Project Manager	ESDIS Deputy Project Manager for Operations
Research and Analysis (R&A) or Application Program-	Principal Investigator (PI)	ESDIS Deputy Project Manager for Operations

funded Principal Investigator (PI) controlled system		
Making Earth System data records for Use in Research Environments (MEaSUREs) data development program	Principal Investigator (PI)	ESDIS Deputy Project Manager for Operations
Airborne Investigations / Campaigns (EVS) -funded Principal Investigator (PI) controlled system	Data Manager and Project Manager	ESDIS Deputy Project Manager for Operations

The DPs DMPs are valid for five years and are to be updated as required (either as a result of a 5-year review, at the end of the mission, project or investigation, or with other necessary "off-cycle" changes). All changes to the DMP must be controlled and assigned a version number. The delivery destination of the latest 'working' copy of the DMP by the data producer is identified in the third column of Table 1. All DMPs must be coordinated with the affected NASA Earth Science Division (ESD)-assigned DAAC prior to the ESDIS' project review. Once reviewed, the DMP is approved by the ESDIS Point of Contact, and the DMP is returned for signature by the responsible manager. Once signed, the DMP is cataloged and maintained in the ESDIS Library.

Figure 1 depicts the flow of the DMP through this cycle. Note that the review and update of a DMP can be initiated at any time deemed necessary. The decision indicated by the rhombus 1.2 implies a requirement to review and update the DMP every 5 years (the "Yes" decision in 1.2), but the process also provides a path for an ad hoc review (the "No" decision in 1.2) and modification of the DMP.

This template supersedes the *Data Management Template for Science Investigator-led Processing System (SIPS), Revision 3, September 2016.*

Note that if some items of the information called for in this template are available through other documents or websites, they can be referenced or pointed to with a brief description included in the DMP.

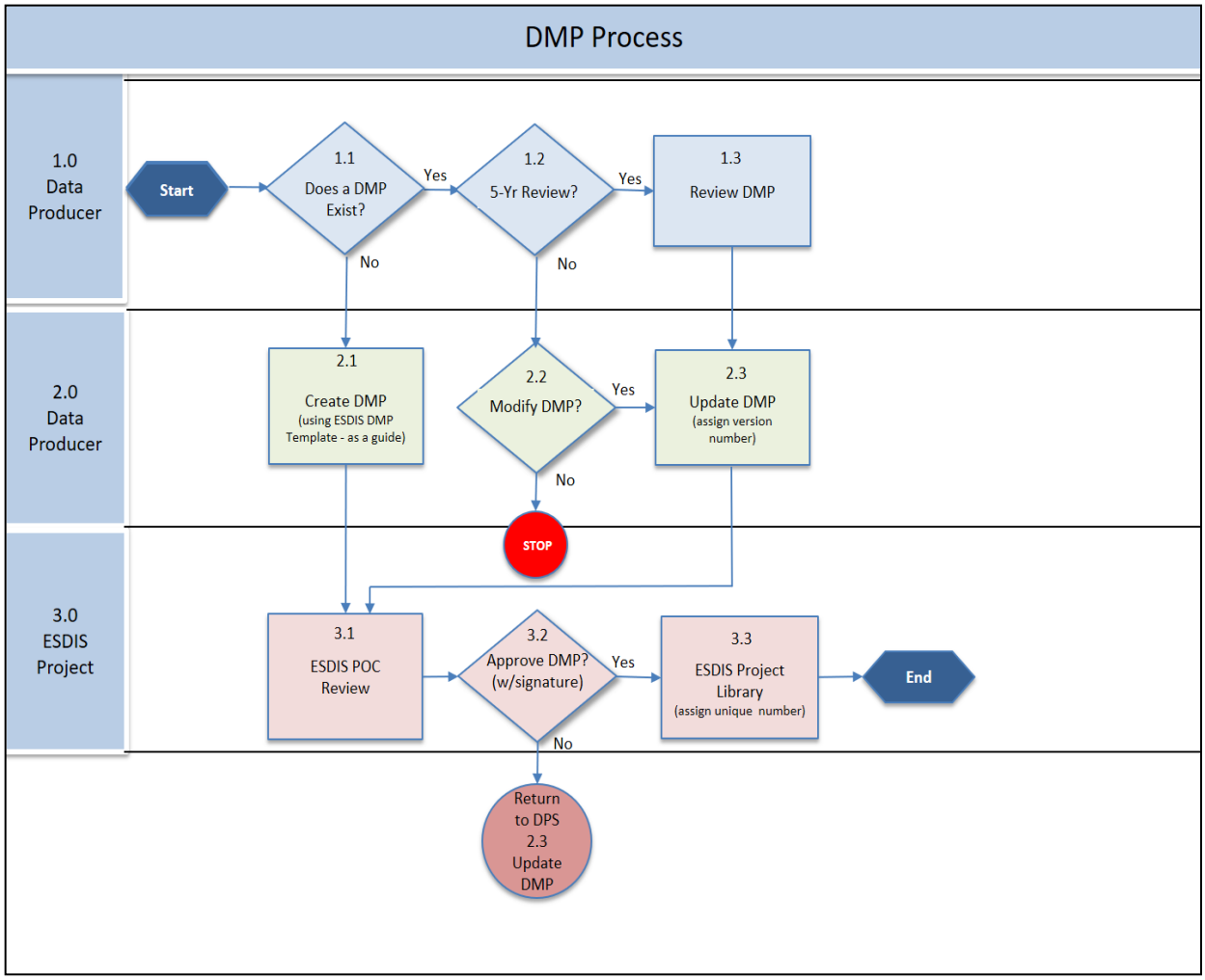

Figure 1: Data Management Plan Process

[SAMPLE Title Page]
Data Management Plan (DMP) Title

Version #
[For version # use whole #'s - update after each revision, example 1.0; 2.0]

Date
[mm/dd/yyyy]

Prepared under NASA Contract [contract number] – *only if applicable*

Approved by:
Signature: [<<SIPS>> or <<SDS>> or <<PI-controlled>> or << MEaSUREs >>]
Manager: Date: [mm/dd/yyyy]

1. Introduction

Briefly describe what the [<<SIPS>> or <<SDS>> or <<PI-controlled>> or << MEaSUREs >>] does and the types of data products it produces.

1.1 Purpose and Scope (of DMP)

1.2 Development, Maintenance and Management Responsibility for the DMP

Example 1. "The [XXX] SIPS is responsible for the development, maintenance, and management of the DMP. The [XXX] SIPS Manager, [SIPS manager name], has overall responsibility for the plan and has specific responsibility for approving any changes to the plan. All changes to the DMP will be controlled."

1.3 Contextual background of Data Processing System

When did the data processing system (i.e., code implementation used to produce the data products) come into being (e.g. 2002 to prepare for EOS Aura)? Where is the system located within its organization (e.g., in xyz Division at GSFC - could show an organization chart)? What is the time range of data processed (e.g., 2004 to present)?

2. Mission/Investigation and Instrument/Sensor Overview[1] [Repeat sections 2 and 3 for each supported mission/investigation and instrument/sensor, as appropriate]

[1] This section is not needed for MEaSUREs.

For <<SIPS>> or <<SDS>>, indicate the missions and instruments from which the data are received.

For <<PI-controlled >> producers of projects, such as Airborne or Field Campaigns, provide the in situ and remote sensing instruments and targeted measurements of each platform.

3. Data Product Summary

Indicate the products created from the instrument or sensor data. Include a tabular summary of the products containing, as appropriate, the following items: platform/instrument, product design (swath, gridded, etc.), data volume (to nearest GB), parameters/measurements, temporal and spatial resolution of data product. Describe in the following subsections the data acquisition, processing, quality plans, and archival processes for the products.

3.1. Data Acquisition

Briefly describe the data source and the acquisition process. Describe the product's original data source and how the data are acquired. Include all the ancillary data inputs required to generate the product. Optionally, include a data flow diagram showing the sources of data needed for product generation.

3.2. Data Processing

Briefly describe the processing activities (i.e., L0-L1A; L1A-L1B; L1B-L2, etc.). What is the process for developing and vetting ATBDs? What is the process for developing and integrating science software into the system? Briefly describe plans for reprocessing and the conditions expected to trigger reprocessing. If there are other documents that describe the processing and reprocessing activities, point to such documents and provide a summary. Indicate the data formats and metadata standards followed.

Data formats and metadata standards should follow NASA ESDIS standards as described in Standards, Requirements and References.

3.3. Data Analysis

Describe the data analysis process and the data producers involved. An example of the analysis process may include the science team’s or other users’ evaluation of the data and the feedback provided to data producers regarding the validity and quality of the data. Describe plans for supporting such activities.

3.4. Data Quality

Describe in detail the plans and processes used for assessing product quality as well as for collecting, documenting, and conveying to the DAAC the quality information of data products for which the data producer is responsible. Developing the product QA

information, disseminating it to the user community and preserving it is the joint responsibility of the data producers (including missions, instrument science teams, science data support teams, and SIPSs/SDSs) and the DAACs. The teams/organizations involved, of course, depend on the type of mission or project. The description here should include the division of responsibilities for various data quality related activities among these groups/organizations. It is essential that users of scientific data products have access to complete (to the degree to which all knowledge is available at the time) and properly articulated (i.e., correctly described for the end user to logically understand, discern and make well-informed decisions) information about the data quality, including known issues and limitations.

This information will help to inform users about the potential applications of the data products and prevent data misuse. Therefore, the DMP should include details on the processes and protocols used for assessing data quality, as well as for capturing, describing, and conveying information about data quality to the NASA-designated DAACs for archiving and for distribution to users along with the data products. The recommended contents for capturing and ensuring data quality for the DMP are provided in the "Data Quality" Appendix A. It is understood that for major NASA missions, the Cal/Val period is relied upon to reveal many of the data quality issues and usage limitations of the data. Other means and methods of extracting and assessing data quality issues are also quite useful for smaller projects (i.e., those that are not bound by high-level mission science requirements as defined in a PLRA) such as MEaSUREs, which offer more flexibility in terms of duration and scope of the quality assessments. Taking into account the dedicated Cal/Val period, as well as other means of discovering data quality issues, it is important to ensure that data quality information is continually updated and disseminated in a timely manner throughout the data lifecycle.

3.5. Data Distribution and Archiving

Indicate the DAAC(s) to which the products will be sent for distribution and archiving. This section should also address:

Collection of Associated Preservation Content – include plans and processes used for collecting software, documentation, and related materials indicated in the NASA Earth Science Data Preservation Content Specification[2] to ensure that the contents are captured before the end of missions/projects that provided the data products to the DAAC. [Status of collecting such content should be maintained separately from this DMP in collaboration with the assigned DAAC(s)].

3.6. Data Product Documentation

[2] See NASA Earth Science Data Preservation Content Specification, https://cdn.earthdata.nasa.gov/conduit/upload/10607/NASA_ESD_Preservation_Spec.pdf. This document can be used as a guide to identify content to be preserved as is applicable to individual missions/investigations.

Describe plans to provide the following to support the data product release to the public: ATBDs, User Guides, File specifications, Validation status, Quality documentation (if available), and Dataset Digital Object Identifiers.

4. REFERENCES

Include a list of references cited in the above sections.

5. ACRONYMS

Include a table showing acronyms used in the document and their meanings.

Appendix A.
Data Quality

The following provides the list of recommended items to be covered by the Data Quality section of the DMP (section 3.4 in the template), with brief explanations where needed:

1. General: Describe the process for assuring data quality. Include data flows and organizations/groups involved in assuring data quality. Provide reference to Interface Control Documents between organizations that have been or will be developed. Work with the NASA-assigned DAAC(s), science teams, and/or review boards to provide data quality information through an existing community-standardized format for describing quality (e.g., GHRSST GDS-2 model for quality confidence levels, see [A-1]). If a standard format is not available or applicable for use, consider extending an existing standard that comes close to your needs. Include documentation clearly explaining the format to be used and/or consider developing a published and citable standards document through the ESDIS Standards Office (ESO). In rare instances, a data producer may opt to create their own standard format; ESO provides a means for reviewing and publishing new data standards that will become officially recognized NASA standards after publication.

2. Errors/Uncertainties: Indicate how errors/uncertainties in the input data used to produce the products will be accounted for, minimized through improved calibration (i.e., to meet the error budget constraints if required for Cal/Val), and/or propagated in higher level products. (See references [A-2] and [A-3]). Describe the level of detail at which errors/uncertainties are shown in the products to be delivered – e.g., per data value, per granule, or at the collection level.

3. Calibration/Validation (Cal/Val): Calibration/Validation is only applicable to missions and funded projects in which Cal/Val is explicitly mandated. Provide the targeted error budget that will be used to assess Cal/Val performance. Note that target error budgets are specified for missions in the Program-Level Requirements' Appendix (PLRA) and can be cited by the DMP. For PI-provided datasets, this may be more loosely defined by either the data producer or by the scientific community that functions as the primary targeted user community. PI-provided datasets that use data from a NASA mission may opt to cite the PLRA-derived error budget for the applicable mission. Provide the expected duration of Cal/Val. Also, document the conditions under which Cal/Val activities are expected to be conducted.

4. Ancillary datasets: Describe plans for managing, archiving and distributing the ancillary datasets that used for QA/QC, Cal/Val, error budget validation, uncertainty quantification, and uncertainty characterization.

5. Quality Flags/Indicators: Indicate how quality flags and/or indicators are defined and used in the generated products. The following are provided as general considerations regarding quality flags/indicators:
   - Define and/or create "indicators" to represent the quality of a data product from different aspects (e.g., data dropout rate of a "sea surface temperature" data product can be considered as one data uncertainty indicator).
   - Ensure that quality flags are related to quantifiable metrics that directly refer to the usefulness, validity, and suitability of the data. See examples in Table A-1 below.
   - Identify quantifiable data quality criteria such as confidence levels and the values of quality flags, which can be used as criteria for refining search queries.
   - Provide quality and measurement state information for Climate and Forecast (CF) metadata attributes: flag_values, flag_masks, and flag meanings. The choice of flag_values vs flag_mask will depend on the use case. Generally, flag_values and flag_masks are not used together, but there are very complex cases in which they may both be used, as illustrated in CF 1.7 example 3.5 (see above link). In all cases, flag_meanings is used in conjunction with either flag_masks and/or flag_values.

     Below are some definitions and actual examples from existing NASA data:
     - flag_values: a set of scalar flag values that represent a quality "state" on a pixel-by-pixel basis such as data confidence level. The meaning of the flag_value is deciphered by the flag_meanings attribute.

     Example 2. GHRSST example from GDS-2 AMSR-2 L2P Product
       - flag_meanings: 0= no data; 1= bad (near ice/land, sunglint, RFI, edge of swath, SST out of range, wind > 20m/s, bad quality); 2= bad (rain); 3= useable but diurnal estimate shows warming > 1; 4= useable but possible error, see l2p_flags bits 9-15; 5= best quality data
       - flag_values: 0b, 1b, 2b, 3b, 4b, 5b

     - flag_masks: serve to represent the expressed values of the Boolean conditions of specific bits described by the flag_meanings attribute; these are often used to describe known algorithm anomalies.
     - flag_meanings: human-readable attributes that describes, as a series of space-separated unique strings, the unique individual meanings of either the flag_values or flag_mask Boolean conditions. See Table A-1 for an example of flag_mask values and flag_meanings used in a SMAP Level 2 product produced by Remote Sensing Systems (RSS), Inc.

- Provide ancillary quality and/or uncertainty flags to facilitate detection of areas that are likely to contain spurious data (e.g., ice in unexpected places).
- Provide pixel-level uncertainty information where possible and meaningful; provide the confidence level (e.g., 95%) to indicate the statistical significance of the uncertainty information.
- Provide data quality variables and metadata along with detailed documentation on how the variables/metadata are derived and suggestions on how to interpret the variables/metadata or use them in different applications.
- Provide a definition and description of each data quality indicator, including the algorithms and datasets used to derive the quality information and a description of how each quality indicator can be used.

6. Metadata Standards and Compliance: Describe plans for developing data quality-related metadata in compliance with current standards – ISO 19157 (Geographic information – Data quality), Climate and Forecast (CF) conventions including those for flags and indicators, Attribute Convention for Dataset Discovery (ACDD), and ISO 8601 (date and time). Include plans for ensuring compliance with these standards; for example, by using an automated compliance checker such as the PO.DAAC Metadata Compliance Checker (MCC) [A-4] and the EOSDIS Common Metadata Repository (CMR) compliance checker [A-5].

7. Documentation: Indicate how quality flags/indicators will be documented and conveyed to DAACs (and users). Show how uncertainty will be documented and how information will be provided to the DAACs. Include plans for documenting known issues and caveats for users and consider leveraging DAAC resources to expedite updating and publication of this information (e.g., Forums, web/email announcements, etc.). See references [A-6] and [A-7] for examples of how known issues can be conveyed to a DAAC and shown on the DAAC website. Some of the most useful QA information for data users may be in journal articles and in various science team member presentations (at conferences or science team meetings). Those gathering QA information should consider these valuable sources of information that are fairly easy and inexpensive to collect and maintain.

8. Post-Delivery Actions: Describe plans for informing DAACs (and users) when datasets are compromised and/or problems are discovered. Also, describe plans for responding to data quality issues reported by DAACs (and/or users) and discuss plans for updating data quality documentation when new versions of products are generated. Consider leveraging DAAC resources for the most expedient and widest dissemination of information (e.g., Forums, web/email announcements, etc.).

Table A-1: SMAP Example from Level 2 RSS Product

Flag Mask	*Flag-Meaning*
1	no_radiometer_observations_in_cell
2	problem_with_OI
4	strong_land_contamination
8	strong_sea_ice_contamination
16	MLE_in_SSS_retrieval_algo_has_not_converged
32	sunglint
64	moonglint
128	high_reflected_galaxy
256	land_contamination
512	sea_ice_contamination
1024	high_residual_of_MLE_in_SSS_retrieval_algo
2048	low_SST
4096	high_wind
8192	no_valid_imager_wind_speed_available_within_60_minutes_of_SMAP_observation
16384	rain_flag

REFERENCES (to Appendix A)

[A-1] GHRSST Science Team (2010), The Recommended GHRSST Data Specification (GDS) 2.0, document revision 5, available from the GHRSST International Project Office, 2015, pp 123, https://www.ghrsst.org/wp-content/uploads/2016/10/GDS20r5.pdf.

[A-2] JCGM 100: 2008. Evaluation of measurement data — Guide to the expression of uncertainty in measurement.

[A-3] International Organization for Standards (ISO), 1994: Accuracy (trueness and precision) of measurement methods and results -- Part 1: General principles and definitions. ISO 5725-1:1994. https://www.iso.org/standard/11833.html

[A-4] PO.DAAC, Metadata Compliance Checker, https://podaac-uat.jpl.nasa.gov/mcc/ (accessed June 28, 2018)

[A-5] ESDIS Project, Common Metadata repository (CMR) Ingest Application Programming Interface (API) Documentation, https://cmr.earthdata.nasa.gov/ingest/site/docs/ingest/api.html (accessed June 28, 2018).

[A-6] LP DAAC, SRTMGL3S: NASA Shuttle Radar Topography Mission Global 3 arc second sub-sampled V003, https://lpdaac.usgs.gov/products/srtmgl3sv003/ (accessed June 23, 2020).

[A-7] LP DAAC, Known Issues for MODIS Vegetation Indices Version 6 Products, https://lpdaac.usgs.gov/news/known-issues-for-modis-vegetation-indices-version-6-products/, (accessed June 23, 2020).

Table A-2: Acronyms and Abbreviations Used

Acronym/ Abbreviation	*Description*
ACDD	Attribute Conventions for Data Discovery
AMSR-2	Advanced Microwave Scanning Radiometer 2
ATBD	Algorithm Theoretical Basis Document
Cal/Val	Calibration and Validation
CF	Climate and Forecast (metadata conventions)
CMR	Common Metadata Repository
DAAC	Distributed Active Archive Center
DMP	Data Management Plan
DP	Data Producer
ESD	Earth Science Division
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System
ESDIS	Earth Science Data and Information System
ESO	ESDIS Standards Office
EVS	Earth Venture Sub-orbital
GB	Gigabytes
GDS-2	GHRSST Data Specification version 2.0
GHRSST	The Group for High Resolution Sea Surface Temperature
GSFC	Goddard Space Flight Center
ISO	International Organization for Standards
MCC	Metadata Compliance Checker
MEaSUREs	Making Earth System data records for Use in Research Environments
MLE	Maximum Likelihood Estimate
NPR	NASA Procedural Requirement
PLRA	Program Level Requirements Appendix
PO.DAAC	Physical Oceanography DAAC
QA	Quality Assurance/Assessment
QC	Quality Control
PI	Principal Investigator
R&A	Research and Analysis
RFI	Radio Frequency Interference
RSS	Remote Sensing Systems, Inc.
SDS	Science Data System
SIPS	Science Investigator-led Processing System
SMAP	Soil Moisture Active and Passive
SSS	Sea Surface Salinity
SST	Sea Surface Temperature