

Data Quality Working Group Recommendations for the Data Management Plan Template for Data Producers

Status of this Memo

This document introduces a template that NASA-funded Earth science Data Producers (DPs) can use to develop Data Management Plans (DMPs). It updates the version of the DMP template for Science Investigator-led Processing Systems (SIPSs) dated September 2016 to include recommendations from the Data Quality Working Group (DQWG). The distribution of this document is unlimited.

Change Explanation

Not applicable. This Request for Comments (RFC) is not an update to an existing RFC, but a revision to a document that was produced by the Earth Science Data and Information System (ESDIS) Project in September 2016 and distributed to the SIPSs.

Copyright Notice

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Abstract

This document introduces a template that NASA-funded Earth science Data Producers (DPs) can use to develop Data Management Plans (DMPs). It updates the version of the DMP template for Science Investigator-led Processing Systems (SIPSs) dated September 2016 to include recommendations from the Data Quality Working Group (DQWG), one of NASA's Earth Science Data System Working Groups. These recommendations were developed during 2014-2016 by an analysis of use case studies that highlight issues regarding capturing, describing and conveying information about the quality of data products held at Earth Observing System Data and Information System (EOSDIS) Distributed Active Archive Centers (DAACs). The "DMP Template for DPS" incorporating the DQWG recommendations is attached to this document as Appendices B and C and is available as an independent document.

1 Introduction

This DMP Template for DPs is based on the earlier DMP Template for SIPS. It explicitly calls for a Data Quality (DQ) section and provides detailed guidance on its contents, with the goal of more complete and consistent treatment of data quality by the data producers. The updated template has a revised title, "Data Management Plan Template for Data Producers" to reflect its applicability to a broader set of data processing systems than SIPS. Appendices B and C to this RFC, to be available as a standalone document, constitute the updated template.

[The term "data producers" is used in this document to refer to Principal Investigators and science teams that develop algorithms as well as people managing/running the data production

systems (e.g., Science Investigator-led Processing Systems - SIPSs) who work closely with the algorithm developers. It is expected that the PIs/science teams and the data production system managers work out the division of responsibilities for the various tasks related to data production and document them in the DMPs.] The audience for this RFC and the "DMP Template for DPS" is data producers. The data producers may also benefit from the section describing plans for addressing data quality, as it can help them develop their plans. The audience for the DMPs generated by data producers is the NASA Headquarters Program Manager/Executive, ESDIS Project and the DAACs.

Data producers should take the recommendations of the DQWG into account while developing their Data Management Plan (DMP). To facilitate this, the DQWG guidelines regarding the contents of the DQ section of a DMP are incorporated into the DMP Template for DPS. It is to be noted that the DQ section of a DMP is not a documentation of existing data quality, but a plan of how the data producer will address and manage data quality throughout the activity lifecycle. In particular, the DQ section of the DMP indicates how the data producer creates and delivers the DQ information for their products to the DAAC. This DQ information (flags, indicators, documentation, etc.) should be delivered to the DAAC with the corresponding data products. The audience for such DQ information is data users. The DQ section should also include, when applicable, statements of expectations such as high-level mission/science requirements that pertain to DQ. Examples of such requirements are: Calibration and Validation (Cal/Val) schedule and duration; and acceptable Cal/Val performance criteria such as the targeted error budget. In rare situations, Cal/Val performance criteria may evolve as new performance metrics are recommended (e.g., by a science team/review board) and be retroactively applied to improve the calibration for data produced by certain missions/projects that are considered ground-breaking or experimental (e.g., Earth Venture missions). As such, these additional performance metrics and associated criteria should be periodically updated in the DQ section, along with a statement of rationale explaining why these unforeseen performance metrics should be considered relevant to the success of the mission/project.

The guidelines provided in this document are not intended to be directly applied to the brief DMPs submitted along with proposals in response to NASA solicitations. However, the DQWG does recognize that solicitations requesting a brief initial DMP may benefit from the inclusion of some of the basic recommended DQ information that is either already known or expected to be delivered as a result of the planned work as indicated in the proposal. It is therefore recommended that, at their discretion, specific NASA Earth Science Division (ESD) programs include language regarding the DQ information to be covered in the initial DMPs in response to the respective solicitations.

These guidelines are intended to directly apply to the more detailed DMPs that are developed by activities (missions, projects, or investigations) soon after being funded by ESD. It is recognized that not all of the information needed for the DMPs may be available during the early phases of the activities, and that the DMPs may need to be updated later in the activity lifecycle. Thus, funded Principal Investigators (PIs), at their discretion and in coordination with their ESD Program Managers, should include the DQ information that is suitable and available in early mission phases and subsequently update DMPs with more complete and current DQ information.

Such update(s) should occur before data are delivered to the DAAC at each delivery iteration and processing level.

This document is derived from the recommendations developed during 2014-2016 by the Data Quality Working Group (DQWG), which is one of NASA's Earth Science Data System Working Groups. These recommendations are the result of the analysis of 16 use cases describing users' explorations of selected data products held at the EOSDIS DAACs [1]. These use cases highlight issues and needs regarding capturing, describing and conveying information about the quality of data products to data product users so that users can access the information, understand the DQ, and properly use the data products. The members of the DQWG made recommendations for improvement, some of which apply to data producers and others to data distributors (DAACs).

1.1 Relationship to Other Documents

Recommendations from the DQWG are published in several documents.

- This document incorporates Data Quality recommendations for data producers into the DMP template for DPs.
- A separate document [2] contains a DMP template for use by DAACs to assist in conforming to the DQWG recommendations to the DAACs. Note that much of what the DAACs should address in their respective DMPs is carried over from DMPs previously drafted by data producers. It is this dependency on complete and accurate exchange of relevant dataset information between data producers and DAACs that necessitates commonalities between these two DMP templates, which is one of the core objectives and anticipated outcomes for both DMP convention specifications.
- The full list of use cases and resulting analysis is published in [1], in order to provide transparency into the origin and rationale behind most of what is being proposed in this document.
- The highest priority recommendations of the DQWG are available in [3], to encourage immediate action by members of the Earth Science data community.

This document complements earlier guidelines/templates for the development of DMPs for activities funded by NASA's ESD, which include only very brief comments on DQ. For instance, the ESD DMP Guidelines document [4] mentions that "Included with the products should be documentation for correct and independent use of the data" and indicates Data Quality Documentation among other documents to be developed. The original DMP Template for SIPS [5], which this document updates, requires a section on data quality and states "Describe plans and processes for collecting, documenting, and conveying to the DAAC information on the quality of data products for which the SIPS is responsible." However, the language does not provide sufficient details to address the issues and needs identified by the DQWG.

2 RFC Content

The “DMP Template for Data Producers” incorporating the DQWG recommendations is available from the NASA Earth Science Data Systems Program as an independent document [6].

3 References

- [1] NASA Earth Science Data Quality Working Group, Comprehensive Data Quality Recommendations for Data Producers and Distributors, ESDIS Standards Office Suggested Practice Technical Note ESDS-RFC-033, <https://earthdata.nasa.gov/user-resources/standards-and-references/recommendations-from-the-data-quality-working-group>.
- [2] NASA Earth Science Data Quality Working Group, Data Management Plan Template for DAACs, ESDIS Standards Office Suggested Practice Technical Note ESDS-RFC-031, February 2019, <https://earthdata.nasa.gov/user-resources/standards-and-references/templates-for-nasa-data-management-plans>.
- [2] NASA Earth Science Data Quality Working Group, High-Priority Data Quality Recommendations for Data Producers and Distributors, ESDIS Standards Office Suggested Practice Technical Note ESDS-RFC-034, <https://earthdata.nasa.gov/user-resources/standards-and-references/recommendations-from-the-data-quality-working-group>.
- [4] NASA SMD ESD, Guidelines for Development of a Data Management Plan (DMP), Version 1.0, January 2011, https://smd-prod.s3.amazonaws.com/science-blue/s3fs-public/atoms/files/Data_Mgmt_Plan_guidelines-20110111.pdf
- [5] NASA ESDIS Project, Data Management Plan (DMP) Template for the Science Investigator-led Processing System (SIPS), September 2016.
- [6] NASA ESDS Program, Data Management Plan Guidance, February 2019, <https://earthdata.nasa.gov/earth-science-data-systems-program/new-missions/data-management-plan-guidance>.

4 Authors

This specification was prepared by the DMP DQ Section Subgroup of the Data Quality Working Group. The DMP Template for DPs in Appendices B and C is the result of collaboration between Alfreda Hall (NASA HQ) and this subgroup. The names of the members of this subgroup are listed below:

H. K. “Rama” Ramapriyan, Science Systems and Applications, Inc. and Earth Science Data and Information System Project, NASA Goddard Space Flight Center, Tel: 805-402-7125, Hampapuram.Ramapriyan@ssaihq.com (Chair, DMP DQ Section Subgroup)

Donna Scott, National Snow and Ice Data Center (Co-Chair, DMP DQ Section Subgroup)

Ed Armstrong - Physical Oceanography DAAC, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

Charlene DiMiceli – University of Maryland

ESDS-RFC-032
Category: Convention
Updates: DMP Template for SIPS
(September 2016)

NASA ES Data Quality Working Group
February 2019
DMP Template for Data Producers

Robert Downs – Socio-Economic Data and Applications Center, Columbia University

Carolyn Gacke – Land Processes DAAC

Scott Gluck – Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

George Huffman – NASA Goddard Space Flight Center

Zhong Liu – NASA Goddard Earth Science Data and Information Service Center (GES DISC)

David Moroni - Physical Oceanography DAAC, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

Chung-Lin Shie – NASA Goddard Earth Science Data and Information Service Center (GES DISC)

Deborah Smith – Global Hydrology Resource Center DAAC

Yaxing Wei – Oak Ridge National Laboratory DAAC

Appendix A - Acronyms

Acronym/ Abbreviation	Description
Cal/Val	Calibration and Validation
DAAC	Distributed Active Archive Center
DMP	Data Management Plan
DP	Data Producer
DQ	Data Quality
DQWG	Data Quality Working Group
EOSDIS	Earth Observing System Data and Information System
ESD	Earth Science Division
ESDIS	Earth Science Data and Information System
ESDSWG	Earth Science Data System Working Groups
NASA	National Aeronautics and Space Administration
RFC	Request for Comment
SIPS	Science Investigator-led Processing System

ESDS-RFC-032
Category: Convention
Updates: DMP Template for SIPS
(September 2016)

NASA ES Data Quality Working Group
February 2019
DMP Template for Data Producers

Appendix B