Quick Start Guide to the EOSDIS Data Product Development Guide for Data Producers

Version 1.0, June 23, 2021

DOI: 10.5067/DOC/ESO/DPDG_QSG_VERSION1

Suggested Citation:

Ramapriyan, H. K., and P. J. T. Leonard. 2021. Quick Start Guide to the EOSDIS Data Product Development Guide for Data Producers version1. NASA Earth Science Data and Information System Standards Office, Publication Date 2021-06-23. https://doi.org/10.5067/DOC/ESO/DPDG_QSG_VERSION1

Status of this Memo

This memo provides information to the National Aeronautics and Space Administration (NASA) Earth Science Data Systems (ESDS) community. This memo describes a "Suggested Practice" and does not specify an ESDS standard of any kind. Distribution of this memo is unlimited.

Change Explanation

This is the original version of this document. It is a condensed version of the Data Product Development Guide for Data Producers (ESDS-RFC-041).

Copyright Notice

Abstract

The Data Product Development Guide (DPDG) for Data Producers was developed for the Earth Observing System Data and Information System (EOSDIS) by the DPDG Working Group, one of the Earth Science Data System Working Groups (ESDSWG), to aid in the development of NASA Earth Science data products. Version 1.0 of the DPDG was published in July 2020. This is a companion document, which, as a condensed version, serves as a quick start guide for novice data producers before delving into the details of the DPDG.

Table of Contents

List of Figures

1 Introduction

The Earth Science Data System Working Groups (ESDSWG) [1], under the supervision of the ESDIS Project, created a Data Product Development Guide (DPDG) for Data Producers to aid in the development of NASA Earth Science data products. The DPDG [2] is primarily intended for producers of Earth Science data products derived from remote sensing data. However, producers of related data products, including Level 0 and 1A satellite data, and airborne and ground-based data products, will also find useful guidance. The Quick Start Guide (this document) is a condensed version of the DPDG to help data producers get started with their data product development quickly, while referring to the DPDG for details as required. It is vital that data producers work closely with their assigned Distributed Active Archive Center (DAAC) to ensure compliance with its policies and data lifecycle procedures, as well as to obtain details that may not be covered in this Quick Start Guide by referring to the DPDG.

This Quick Start Guide is organized into Sections that largely mirror those of the DPDG, but which are considerably abbreviated, and, in some cases, combined. The flow chart in Figure 1 shows the various steps in data product development and delivery.

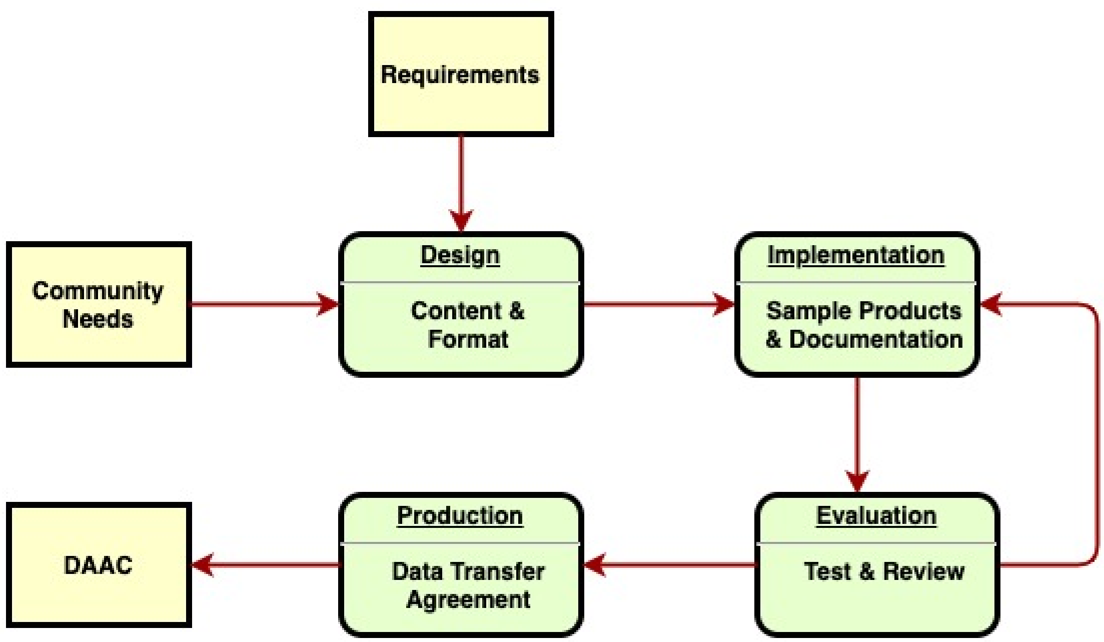

Figure 1. Flow of activities for data product development, production, and delivery. Processes are depicted in green rounded rectangles. Sources and destinations of these processes are shown as yellow rectangles.

2 Data Product Design Process

For this guide, a data product is defined as a set of data files that can contain multiple parameters, which compose a logically meaningful group of related data. This concept is equivalent to a data collection in the Common Metadata Repository (CMR) [3] and is known colloquially as a dataset.

Based on the Earth Observing System (EOS) heritage [4], "Standard" Earth Science data products should have the following common characteristics:

- Peer-reviewed algorithmic basis.
- Wide research and application utility.
- Routinely produced over spatially and temporally extensive subsets of data.
- Available whenever and wherever the appropriate input data are available.

The following paragraphs discuss steps that can be applied to achieve a good data product design.

At the beginning of the design process, the key elements are: 1) to identify the expected user communities, and 2) to understand their needs regarding data formats and metadata in addition to what is required for data search and discovery.

Once user communities are identified, there are many key questions to answer while working with the DAAC that will archive and distribute the data products (i.e., the assigned DAAC). A few examples of such questions are given below. See Section 2.1 of the DPDG for an expanded list of key questions.

- How can these data be used by the identified communities?
- Which tools and services will the community need to use the data?
- What are the prevalent data formats, data standards, and metadata standards used by the community?
- What information on data provenance (data product history and lineage) and quality will the users need for their purposes?

A data product design should address the following:

- Data format, metadata, and associated conventions.
- Identification and structure of constituent variables.
- Representation of coordinates, especially geolocation and time.
- Data chunking and internal compression.

The data producer should create sample data files to support evaluation of the data product design. Data products should be tested in coordination with the assigned DAAC using the tools and services that user communities are expected to use. Ideally, an iterative approach should be followed—supply the data product to selected representatives of the expected user communities and DAACs, make changes based on their feedback, re-test the product, and solicit additional feedback.

Soliciting external, independent evaluations will improve the usability of a data product. It is also helpful to provide a way for the user community to provide feedback on the usability and quality of the data after the products are released. Such feedback should be gathered by the assigned DAAC. The DAAC will convey the user feedback to the data producer, as appropriate. Including the data producer's name and contact information in the product metadata will aid the feedback process. See Section 2.5 in the DPDG for recommendations for establishing and conducting independent evaluations of data products.

3 Selecting a Data Product Format

While several acceptable formats are listed by the ESDIS Standards Office (ESO) [5], a highly preferred format for EOSDIS data products is netCDF-4 (network Common Data Form Version 4) that uses the Hierarchical Data Format Version 5 (HDF5) [6] data storage format [7].

Some of the advantages to using netCDF-4 are:

- Files are "self-describing", meaning they allow for inclusion of metadata that describe the contents of the file.
- Supports many data storage structures, including multidimensional arrays and raster images, and naturally accommodates hierarchical groupings of variables.
- Includes access to useful HDF5 features, such as internal data compression, and works with HDF5 tools such as HDF Product Designer (HPD) [8] and HDFView [9].
- Supported by several important software packages, programming languages and computing platforms used in Earth Science.
- Provides efficient input/output on high performance computing systems.

NetCDF-4 is closely associated with a well-established standard called the Climate and Forecast (CF) Metadata Conventions (hereafter, the CF Conventions) [10], which specify a set of metadata that provide a definitive description of the data in each product file, including its spatial and temporal properties.

Another preferred data format is GeoTIFF (Georeferenced Tagged Image File Format), which is a georeferenced raster image that uses the public domain Tagged Image File Format (TIFF) [11], and is used extensively in the Geographic Information System (GIS) [12] communities [13]. OGC GeoTIFF Standard, Version 1.1 is an EOSDIS recommended format [14].

Another ESO-recognized format is the International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) Version 2 format [15]. Although it was primarily designed for airborne data, this format is also used for other mobile and ground-based studies.

4 Metadata

4.1 Overview

Metadata are essential for data management. The data producer shares with the assigned DAAC the responsibility of providing adequate metadata that describe the data product. Metadata requirements are stipulated by the policies of the assigned DAAC.

4.1.1 Data Product Search and Discovery

The metadata that are key for success during the discovery process include:

- Intelligible, descriptive data product names (Section 4.2).

- Precise temporal and spatial coverage (Sections 4.4 and 4.5).
- Accurate and complete list of applicable Global Change Master Directory (GCMD) Keywords [16].
- Concise but readable (including machine readable) description of the data product.

4.1.2 File Search and Retrieval

Once a user has chosen a data product to pursue, the user typically needs to find the particular files within that data product that meet their needs. Most data search engines support search by spatial and temporal criteria. Therefore, a data producer should accurately specify the spatial and temporal extents of each file.

4.1.3 Data Usage

High-quality metadata are essential for making data readable by humans and machines. The DPDG provides guidance on the types of data-product-level and the data-file-level metadata that support data usage.

4.2 Naming Data Products

The name of a data product is critical for its discovery in Earthdata Search and tools developed by the DAACs. The DAACs have rules for naming data products. Therefore, this should be a joint effort between the data producer and the assigned DAAC. Data products must be assigned both a long name, which is meant to be human readable, and a short name, which aids precise searching by keyword. The long name and the short name should be unique within EOSDIS.

There is no universal file-naming convention for NASA Earth Science data products, apart from the Dataset Interoperability Working Group (DIWG) recommendations regarding the components of file names provided in [17] (Rec. 3.8-3.11).

4.2.1 Long Name

The data product long name (called `LongName`[1] in CMR) is a scientific description of the product. It should be as brief (but as complete) as possible, to expand on the often-cryptic corresponding short name.

We provide the following recommendations regarding the formulation of data product long names:

- The data source, usually the acronym/abbreviation for the project responsible for producing the data, but can also be the instrument (e.g., HIRDLS, the High Resolution Dynamics Limb Sounder), satellite (e.g., TRMM, the Tropical Rainfall Measuring Mission), or program (e.g., MEaSUREs, Making Earth System Data Records for Use in

[1] Not to be confused with the CF `long_name` attribute for individual variables.

Research Environments); include both the instrument and satellite names to eliminate ambiguity (e.g., MODIS/Aqua).
- Science content (e.g., Aerosol, Precipitation).
- The general spatial coverage of the data (e.g., gridded, swath, orbit, *in situ*).
- The temporal resolution (e.g., orbital, daily, monthly).
- Processing level (e.g., L2 or Level 2).
- Spatial resolution (e.g., 3 km; if there are multiple resolutions in the product, then state the highest resolution).
- Version number (optional; details should be resolved with the assigned DAAC).

4.2.2 Short Name

The short name (called `ShortName` in CMR) is an abbreviated version of the product name. Alphanumeric and underscore ("_") are the only acceptable characters in short names. Data producers should contact the assigned DAAC to check if there are any additional restrictions on short names.

4.3 Versions

Data products are uniquely identified by the combination of `ShortName` and `VersionID` within the CMR. Data producers can also specify a `product_version` as an Attribute Convention for Data Discovery (ACDD) global attribute [18]. In most cases, the `product_version` and `VersionID` are identical although there may be some exceptions to this (e.g., when selected files associated with a limited reprocessing of data have a different value for the `product_version`). However, if reprocessed data files have significant differences in terms of science content, then these files should be organized into a separate data product having a different `VersionID`. Guidance for setting version numbers should be sought from the assigned DAAC.

The software version used to generate a data product is specified via the CMR attribute `PGEVersion`. If possible, it is useful to use the same version for the data product (`product_version`) as for the algorithm software used to generate the product (`PGEVersion`), to avoid confusing the data product users.

4.4 Representing Coordinates

Earth Science data files should be produced with complete information for all geospatial coordinates to ensure software application capabilities for data visualization, mapping, reprojection, and transformation. This geolocation should be encoded in an interoperable way based on the CF Conventions.

4.4.1 Latitude and Longitude

Variables representing latitude and longitude must always explicitly include the CF `units` attribute. The recommended unit of latitude is `degrees_north` (valid range: -90 to 90 degrees) and the recommended unit of longitude is `degrees_east` (valid range: -180 to 180 degrees).

Practice the following guidelines to support the widest range of software tools while avoiding storage of redundant geospatial coordinate data:

- Specify coordinate boundaries by adding the CF `bounds` attribute [19](Rec. 2.3).
- Include horizontal and vertical (as necessary) attributes.
- Store all coordinate data for a single file in coordinate variables only [17](Rec. 3.5).
- Files are required to contain the most applicable type of geospatial coordinates for the data. The decision whether to provide any additional types of geospatial coordinates is left to the data producer.
- The projection information should be designated via both the CF `grid_mapping` variable-level attribute and as OGC Well-Known Text (WKT) whenever possible [20].

4.4.2 Time

The CF Conventions represent time as an integer or a floating-point number, with the units attribute set to the time unit since an epochal date-time, represented as YYYY-MM-DDThh:mm:ss (e.g., "seconds since 1993-01-01T00:00:00Z"). Use Coordinated Universal Time (UTC) instead of local time unless there is a strong justification for not doing so.

The date-time information in the file names should adhere to the following guidelines (detailed in [17], Rec. 3.11):

- Adopt the ISO 8601 standard [21] [22] for date-time information representation.
- If describing time intervals, the start time should appear before the end time.
- Date-time fields representing the temporal extent of a file's data should appear before any other date-time field in the file name.
- All date-time fields in the file name should have the same format.

4.4.3 Vertical

Some data have a vertical dimension, and, therefore, a variable should be included to describe the vertical axis. The most commonly used values to describe vertical coordinates are layer, level, pressure, height, and depth. It is important to identify the vertical coordinates using the most common standard terminology, and to include the following information:

- `long_name`: This CF attribute can be used to describe the vertical coordinate and can also be used to clarify the CF `units` attribute.
- `positive`: This CF attribute refers to the direction of increasing coordinate values, with a valid value of either up or down.

4.5 Data Quality

It is essential that users of scientific data products have access to complete information regarding data quality, including known issues and limitations. This information will inform users about the potential applications of the data products and prevent their misuse. Therefore, data products should include metadata pointing to documentation of the data quality as well as the processes used for assessing data quality.

4.5.1 Data Product Documentation

This subsection provides guidance regarding the information that should be covered in documents considered too extensive to be contained within the data files themselves.

1. Document the process(es) used, including data flows and organizations involved in assuring data quality. Provide reference to Interface Control Documents (ICDs), if any, between organizations that have been or will be developed.
2. Provide documentation of the Calibration/Validation (Cal/Val) approach used, including sources of data, duration of the process, the targeted uncertainty budget that was used to assess performance, and the conditions under which Cal/Val are conducted.
3. Provide a description of how quality flags or indicators are used in the product and explain their meanings. This includes documenting quality flags that are related to a quantifiable metric that directly relates to the usefulness, validity, availability, uncertainty and confidence, and suitability of the data at the pixel level.
4. A quality summary should also be documented and disseminated at the time that a new dataset version is published. The quality summary should at least be a high-level overview of strengths and limitations of the dataset and should be directly traceable and reproducible by the variables within the dataset. It should also include information such as the percent of observations that are missing or are in each quality category.
5. Provide documentation of methods used for estimating uncertainty and how uncertainty estimates are included in the data product.

4.5.2 Product Metadata

This subsection provides guidance regarding the information that should be provided in the metadata entries within the data files themselves.

1. Include the uncertainties in the delivered data. For example, these can be expressed per data (pixel) value, per file, or per data product.
2. Provide pointers (URLs or citations) to the ancillary data products that are used for quality assessments, Cal/Val, uncertainty budget validation, uncertainty. quantification, and uncertainty characterization.
3. Implement quality flags and indicators consistent with the documentation discussed above (Item #3 in Section 4.5.1).

Consider compliance with metadata standards related to data quality – International Organization for Standardization (ISO) 19157 [23], CF Conventions [10] including those for flags and indicators, ACDD [18], and ISO 8601 [21] [22]. Plan on using an automated compliance checker (see Section 6).

4.6 Global Attributes

Global attributes (i.e., those that apply to an entire file rather than to a particular group or variable in a file) improve data discoverability, documentation, and usability. Descriptions of the recommended global attributes, according to CF, Attribute Conventions for Data Discovery (ACDD) [18], and other conventions, can be found in Appendix D of the DPDG.

4.6.1 Provenance

Data provenance consists of the origins, lineage, custody, and ownership of data, and must be included in file-level and product-level metadata for transparency and reproducibility. When describing provenance, include information about the environment used to create the data product (e.g., software version, processing system, processing organization) and the context of the run (e.g., production date-time, list of input data, and ancillary files).

4.7 Variable-Level Attributes

Variables should have specific attributes attached to them, which describe the data within each variable. The recommended variable attributes given by the CF and the ACDD conventions can be found in Appendix E of the DPDG. See also the CF Metadata Template [24].

5 Data Compression, Chunking and Packing

Data compression and chunking are two storage features provided by the HDF5 library and available through the netCDF-4 API. HDF5's internal compression can reduce the space taken up by variables, especially those with many fill values or value repetition. The saved space can pay significant dividends in both storage space and transmission speed over the network. HDF5 includes a compression method referred to as "deflation", based on the common compressor gzip (itself based on the Lempel-Ziv algorithm)[2]. HDF5/netCDF-4 variables can also be "chunked," which means that each variable is stored as a sequence of separate chunks in the file. This allows read programs to decompress only the chunks needed for a read request, and not the entire variable, resulting in even greater I/O efficiencies.

[2] Deflation levels run from 1 to 9 with storage efficiency and time to compress increasing with each level. A level of 5 is often a good compromise.

An alternative way to reduce data size is to apply a scale and offset factor to the data values, allowing the values to be saved as a smaller datatype, such as a 16-byte integer. This technique, known as “packing,” is appropriate for data with a limited dynamic range.

The DIWG recommends that packing be employed only when data are stored as integers [19](Rec. 2.6).

In addition, the following command-line utilities can be used to chunk and compress files after they have been written:

- h5repack (part of the HDF5 library [6]).
- nccopy (part of the netCDF library [25]).
- ncks (part of the NCO package [26]).

These utilities are also useful in experimenting with different compression levels and chunking schemes.

6 Tools for Data Product Testing

Compliance and usability of a new data product should be tested via data inspection, automated compliance checkers, etc. The DPDG provides details about several tools for data product testing. A few of them are highlighted here:

- HDFView [9]: Read HDF5 and netCDF-4 files; view any data object; select “Table” from menu bar and then “export to text” or “export to binary”; create, edit, and delete content of netCDF-4 and HDF5 files.
- h5dump [6]: Dump HDF5 and netCDF-4 files to ASCII format.
- ncdump [25]: Dump netCDF-4 content to ASCII format.
- Panoply [27]: Read HDF5 and netCDF-4 files; the Array tab displays the actual data values that can further be edited in a spreadsheet; create images and maps from the variables contained in a data product.
- HDF Product Designer [8]: Design tool that can be used for checking which metadata in a given file are CF- and ACDD-compliant and which are not. Also, can add or remove metadata to make the file CF- and/or ACDD-compliant; can include ISO Metadata in an Earth Science data file.
- Metadata Compliance Checker [28]: Web-based tool and service designed by the Physical Oceanography DAAC for netCDF and HDF formats.
- CF-Convention Compliance Checker [29]: Developed by Hadley Centre for Climate Prediction and Research for netCDF formats.
- Integrated Ocean Observing System (IOOS) Compliance Checker [30].

7 Data Product Digital Object Identifiers

A Digital Object Identifier (DOI) is a unique alphanumeric character string that can be used to identify a data product for universal referencing and discoverability, as well as for proper attribution and citation. A DOI, once registered, is permanent, and can be used to locate the object to which it refers permanently. Data producers should work with the assigned DAAC to designate DOIs to their data products. The process has two steps: reservation and registration. When DOIs are reserved, data producers can start using them in the metadata while the products are generated. In particular, the DOI resolving authority and the DOI identifier must be included as CF global attributes (see [31]). DOIs are then registered in preparation for publication of the data through the assigned DAAC [3].

8 Product Delivery and Publication

DAACs generally have similar data publication processes, which include: 1) obtain the data, documentation and metadata, and related information from data producers; 2) produce CMR compliant metadata and additional documentation (e.g., User's Guide) describing the data; and 3) release the data and documentation for access by the user community. Some aspects of data publication vary across DAACs. The specifics of data delivery and publication, such as schedules, interfaces, workflow, and procedures for submitting data product updates, are best established by communications between the data producers and their assigned DAACs.

9 References

[1] ESDIS, "Earth Science Data System Working Groups," 25 February 2020. [Online]. Available: https://earthdata.nasa.gov/collaborate/esdswg. [Accessed 14 May 2021].

[2] H. K. Ramapriyan and P. J. T. Leonard, "Data Product Development Guide for Data Producers. ESDS-RFC-041v1.0.," ESDIS Project, 9 July 2020. [Online]. Available: https://doi.org/10.5067/DOC/ESO/RFC-041VERSION1. [Accessed 14 May 2021].

[3] ESDIS Project, "Common Metadata Repository," 12 March 2020. [Online]. Available: https://earthdata.nasa.gov/eosdis/science-system-description/eosdis-components/cmr. [Accessed 14 May 2021].

[4] G. Asrar and H. K. Ramapriyan, "Data and Information System for Mission to Planet Earth," *Remote Sensing Reviews,* vol. 13, pp. 1-25. https://doi.org/10.1080/02757259509532294, 1995.

[5] ESO, "Standards and Practices," ESDIS Project, 18 February 2020. [Online]. Available: https://earthdata.nasa.gov/esdis/eso/standards-and-references#standards-and-practices. [Accessed 14 May 2021].

[3] Until the DOI is registered, it can be modified or deleted (withdrawn). However, once registered, the DOI becomes permanent.

[6] ESO, “HDF5 Data Model, File Format and Library – HDF5 1.6,” ESDIS, 5 August 2019. [Online]. Available: https://earthdata.nasa.gov/esdis/eso/standards-and-references/hdf5. [Accessed 14 May 2021].

[7] ESO, “netCDF-4/HDF5 File Format,” ESDIS Project, 5 August 2019. [Online]. Available: https://earthdata.nasa.gov/esdis/eso/standards-and-references/netcdf-4hdf5-file-format. [Accessed 14 May 2021].

[8] A. Jelenak, “HDF Product Designer,” 11 February 2019. [Online]. Available: https://wiki.earthdata.nasa.gov/display/HPD/HDF+Product+Designer. [Accessed 14 May 2021].

[9] The HDF Group, “HDF View,” 2006. [Online]. Available: https://www.hdfgroup.org/downloads/hdfview/. [Accessed 14 May 2021].

[10] ESDIS, “Climate and Forecast (CF) Metadata Conventions,” ESDIS Standards Office, 5 August 2019. [Online]. Available: https://earthdata.nasa.gov/esdis/eso/standards-and-references/climate-and-forecast-cf-metadata-conventions. [Accessed 14 May 2021].

[11] Adobe, “TIFF,” 2019. [Online]. Available: https://www.adobe.io/open/standards/TIFF.html. [Accessed 14 May 2021].

[12] wikipedia, “Geographic information system,” wikipedia, 16 April 2020. [Online]. Available: https://en.wikipedia.org/wiki/Geographic_information_system. [Accessed 14 May 2021].

[13] OGC, “Welcome to the Open Geospatial Consortium,” 2020. [Online]. Available: https://www.opengeospatial.org. [Accessed 14 May 2021].

[14] ESO, “GeoTIFF File Format, ESDS-RFC-040v1.1,” 16 September 2019. [Online]. Available: https://earthdata.nasa.gov/esdis/eso/standards-and-references/geotiff. [Accessed 14 May 2021].

[15] E. Northup, G. Chen, K. Aikin and C. Webster, “ICARTT File Format Standards V2.0,” January 2017. [Online]. Available: https://earthdata.nasa.gov/esdis/eso/standards-and-references/icartt-file-format. [Accessed 14 May 2021].

[16] T. Stevens, “GCMD Keyword Access,” 17 March 2021. [Online]. Available: https://wiki.earthdata.nasa.gov/display/CMR/GCMD+Keyword+Access. [Accessed 14 May 2021].

[17] DIWG, “Dataset Interoperability Recommendations for Earth Science: Part 2, 423-ESO-036v1.2,” 12 June 2020. [Online]. Available: https://earthdata.nasa.gov/files/ESDS-RFC-036v1.2.pdf. [Accessed 14 May 2021].

[18] ESIP Documentation Cluster, "Attribute Conventions for Data Discovery," 24 August 2020. [Online]. Available: http://wiki.esipfed.org/index.php/ Attribute_Convention_for_Data_Discovery. [Accessed 14 May 2021].

[19] DIWG, "Dataset Interoperability Recommendations for Earth Science, ESDS-RFC-028v1.3," ESDIS Project, July 2016. [Online]. Available: https://earthdata.nasa.gov/files/ESDS-RFC-028v1.3.pdf. [Accessed 14 May 2021].

[20] OGC, "Geographic information — Well-known text representation of coordinate reference systems," 13 August 2019. [Online]. Available: http://docs.opengeospatial.org/is/18-010r7/18-010r7.html. [Accessed 14 May 2021].

[21] ISO, "ISO 8601-1:2019 Date and time -- Representations for information interchange -- Part 1: Basic rules," February 2019. [Online]. Available: https://www.iso.org/standard/70907.html. [Accessed 14 May 2021].

[22] ISO, "ISO 8601-2:2019 Date and time -- Representations for information interchange -- Part 2: Extensions," February 2019. [Online]. Available: https://www.iso.org/standard/70908.html. [Accessed 14 May 2021].

[23] ISO, "ISO 19157:2013 Geographic Information - Data Quality," December 2013. [Online]. Available: https://www.iso.org/standard/32575.html. [Accessed 21 May 2021].

[24] D. Auty, "Template for CF (Climate-Forecast) Compliance," 31 December 2018. [Online]. Available: https://wiki.earthdata.nasa.gov/display/ESDSWG/Template+for+CF+(Climate-Forecast)+Compliance. [Accessed 14 May 2021].

[25] UCAR, "Unidata Data Services and Tools for Geoscience," 2021. [Online]. Available: https://www.unidata.ucar.edu/downloads/netcdf/. [Accessed 14 May 2021].

[26] NCO, "NCO 4.9.9 User Guide," 2021. [Online]. Available: http://nco.sourceforge.net/nco.html#Chunking. [Accessed 14 May 2021].

[27] NASA, "Panoply netCDF, HDF and GRIB Data Viewer," 13 May 2021. [Online]. Available: https://www.giss.nasa.gov/tools/panoply/. [Accessed 14 May 2021].

[28] PO DAAC, "Metadata Compliance Checker," [Online]. Available: https://podaac-tools.jpl.nasa.gov/mcc/. [Accessed 14 May 2021].

[29] Hadley Centre for Climate Prediction and Research, "CF-Convention Compliance Checker for NetCDF Format," May 2019. [Online]. Available: http://puma.nerc.ac.uk/cgi-bin/cf-checker.pl. [Accessed 14 May 2021].

[30] IOOS, "Integrated Ocean Observing System Compliance Checker," [Online]. Available: https://ioos.us/compliance-checker. [Accessed 14 May 2021].

[31] ESDIS Project, "DOI Background Information," 27 April 2020. [Online]. Available: https://wiki.earthdata.nasa.gov/display/DOIsforEOSDIS/DOI+Background+Information. [Accessed 14 May 2021].

10 Authors' Addresses

Hampapuram K. Ramapriyan
Telephone: 805.402.7125
Email: Hampapuram.Ramapriyan@ssaihq.com

Peter J. T. Leonard
Telephone: 301.352.4659
Email: peter.j.leonard@nasa.gov

11 Contributors and Editors

Edward M. Armstrong, JPL PO.DAAC/CalTech
Siri Jodha Singh Khalsa, NSIDC DAAC
Peter J.T. Leonard, ADNET & NASA/GSFC
John Moses, NASA/GSFC (Retired)
Hampapuram K. Ramapriyan, SSAI & NASA/GSFC
Tammy Walker, ORNL DAAC

Appendix A. Abbreviations and Acronyms

ACDD	Attribute Convention for Data Discovery
API	Application Programming Interface
ASCII	American Standard Code for Information Interchange
Cal/Val	Calibration/Validation
CF	Climate and Forecast Metadata Conventions
CMR	Common Metadata Repository
DAAC	Distributed Active Archive Center
DIWG	Dataset Interoperability Working Group
DOI	Digital Object Identifier
DPDG	Data Product Development Guide for Data Producers
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System
ESDIS	Project Earth Science Data and Information System Project
ESDS	Earth Science Data System (Program)
ESDSWG	Earth Science Data System Working Group
ESO	ESDIS Standards Office
GCMD	Global Change Master Directory
GeoTIFF	Georeferenced Tagged Image File Format
GIS	Geographic Information System
GSFC	NASA Goddard Space Flight Center
HDF5	Hierarchical Data Format, Version 5
HIRDLS	High Resolution Dynamics Limb Sounder
HPD	HDF Product Designer
ICARTT	International Consortium for Atmospheric Research on Transport and Transformation
ICD	Interface Control Document
ISO	International Organization for Standardization
JPL	NASA Jet Propulsion Laboratory
MEaSUREs	Making Earth System Data Records for Use in Research Environments
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NCO	NetCDF Operator
NetCDF-4	Network Common Data Form, Version 4
NSIDC	National Snow and Ice Data Center
OGC	Open Geospatial Consortium
ORNL DAAC	Oak Ridge National Laboratory DAAC (NASA)
PGE	Product Generation Executable
RFC	Request for Comments
SIPS	Science Investigator-led Processing System
TRMM	Tropical Rainfall Measuring Mission
URL	Uniform Resource Locator

UTC	Coordinated Universal Time
WKT	Well-Known Text mark-up language