Inter-Agency Implementation and Advanced Concepts Team

Stewardship Best Practices for Documenting and Representing NASA Airborne and Field Campaign Data and Information

Version <v02r03-20220301>

Marshall Space Flight Center
Huntsville, Alabama

IMPACT
Interagency Implementation and Advanced Concepts Team

Stewardship Best Practices for Documenting and Representing NASA Airborne and Field Campaign Data and Information

Approval Page

Prepared by:

Ge Peng, Sr. Principal Research Scientist
ESSC, University of Alabama in Huntsville

Deborah K. Smith, Principal Research Scientist III
ESSC, University of Alabama in Huntsville

Stephanie M. Wingo, Research Scientist II
ESSC, University of Alabama in Huntsville

Approved by:

Rahul Ramachandran
Inter-Agency Implementation and Advanced Concepts Team Executive
NASA / Marshall Space Flight Center

Change History Log

Version	Date	Description of Changes
v01r00	09/30/2021	First version of the document for review by stakeholders
v02r00	11/15/2021	Changes made to address and implement the review edits, comments and suggestions made by Kaylin Bugbee. Sections impacted: all.
v02r01	12/21/2021	Changes made to address and implement review edits by the IMPACT team. Sections impacted: all.
v02r02	02/14/2022	Changes made to address and implement review edits by DAACs. Sections impacted: Sections 1, 3, 5, and 6.
v02r03	03/01/2022	Changes made to address and implement review edits by Rahul Ramachandran. Sections impacted: Sections 1, 5, and 6.

Table of Contents

1. EXECUTIVE SUMMARY

NASA has a long history of carrying out airborne and field campaigns, across the spectrum of Earth's ecosystems. Airborne and field campaign data products support NASA's six Earth Science focus areas: Atmospheric Composition, Weather and Atmospheric Dynamics, Climate Variability and Change, Water and Energy Cycle, Carbon Cycle and Ecosystems, and Earth Surface and Interior.[1] These data products are utilized by users from various science disciplines and for a wide range of applications.

Per requirements by NASA open data and information policies, campaign data products, metadata, and associated contextual information need to be preserved at NASA Distributed Active Archive Centers (DAACs), curated and publicly available. However, campaign data products are highly heterogeneous and cross-disciplinary, making it extremely challenging to manage and curate.

Under the direction of NASA's Earth Science Data Systems (ESDS) program, the Airborne Data Management Group (ADMG) within the Interagency Implementation and Advanced Concepts Team (IMPACT) project of NASA Marshall Space Flight Center (MSFC) has developed the following high-level stewardship best practices, recommended for DAAC personnel:

Recommendation 1: Communicate clearly and effectively with airborne and field campaign management,[2] data producers and users across multiple disciplines.

Recommendation 2: Capture and maintain a comprehensive and consistent set of critical information and documentation for every campaign.

Recommendation 3: Manage campaign data, metadata, and information as first class assets.

This document provides background, motivations, and methodology for developing the best practices along with a description of specific recommendations under each. The primary objective of developing the best practices is to promote consistency in airborne and field campaign data stewardship with a more unified data access and user experience for airborne data providers and users.

These high-level recommended practices apply to all types of airborne and field data for documenting and representing the data and information from Earth Science airborne and field campaigns that NASA supports or participates in. Additional complementary, domain specific documents are under development for specific types of airborne and field campaign data; these include but are not limited to:

- Atmospheric chemistry data,
- Campaign reports,
- Field station instrument data,
- In-situ manual observation data,
- Ground and airborne radar and lidar data,
- Facility instrument data, and

[1] https://science.nasa.gov/earth-science

[2] Field campaign management includes principal investigators (PIs) and campaign deputies or data managers who are the main point of contact with DAACs. DAACs may also interface with individual instrument scientists in the campaign who deliver campaign data products to the DAAC.

- Unmanned Aircraft System (UAS) data.

2. INTRODUCTION

Earth Science airborne and field investigations/campaigns are a key component of NASA's Earth Observing System Data and Information System (EOSDIS). In addition to supporting satellite instrument and product algorithm development, these airborne and field campaign data also support research in the six Earth Science focus areas and serve data users from various science disciplines and a wide range of applications (National Research Council 2010; National Academies of Sciences, Engineering, and Medicine 2018).

Airborne and field campaign data have historically been stored in various locations such as DAACs, NASA airborne science facilities, field archives, or even individual scientists' computer hard drives. Data products are stored in a variety of standard and non-standard formats (ESDSWG Airborne Data Working Group 2015). The airborne and field campaign data products are derived from measurements observed by many different instruments and are delivered at different data product processing levels. Measurements occur sporadically during a limited deployment timeframe (ESDSWG Airborne Data Working Group 2015; ESDS 2018) and are often managed and represented differently.

Partly due to the heterogeneity and complexity of the data collections and the distributed nature of the data storage network, airborne and field campaign data are traditionally under-curated, when compared to satellite data. Field campaign data product metadata may contain some information relevant to the campaign (such as the campaign acronym and location), but tend to lack significant contextual metadata necessary to fully convey or understand the context of the airborne and field data (ESDSWG Airborne Metadata Working Group 2017; ADMG 2020).

All of these challenges have made it difficult for data users to find, access, and efficiently utilize NASA's airborne and field campaign data. To reduce the difficulties, NASA's ESDS Program tasked the ADMG/IMPACT to develop a set of best practices documents aimed specifically at handling NASA airborne and field campaign data.

DEFINITIONS AND SCOPE

Unless otherwise noted, the terminology used in this document follows the formalized definitions developed by ADMG (2020) for use in NASA's Airborne and Field Investigation Inventory,[3] referred to as the Catalog of Archived Suborbital Earth Science Investigations (CASEI). For example, *field investigation or campaign* (campaign is hereafter used in this document) refers to an observational study during which individuals/ programs/agencies/institutions utilize preselected, specific sensors or sets of sensors to acquire targeted observations in a natural, non-laboratory setting in support of common, clearly defined, science or research objectives. *Deployment* refers to a previously scheduled, continuous time period during which a campaign's platforms and/or sensors are dedicated for use in support of the science objective(s). There may be one or more deployments within a single campaign. *Significant Event* refers to an event observed

[3] https://earthdata.nasa.gov/esds/impact/admg/admg-definitions

during a deployment that is notable for the occurrence of a great example of one or more specific phenomena relating to the campaign science or research objective(s).

Data stewardship is defined by the National Research Council (2007) as "activities to preserve or improve the information content, accessibility, and usability of environmental data and metadata". In this document, *stewardship* includes activities to oversee data ingestion, metadata curation, and data release with associated contextual metadata.

Best practices refer to a set of guidelines, processes, procedures or methods that are generally accepted as being correct or more effective with proven results. The primary goal of defining and sharing best practices is "to identify the best way of doing something, and then, to share that knowledge with others working on the project. The result hopefully streamlines the process and puts everyone on the same page of doing things in the best way possible" (Curtis 2019).

PURPOSE

The purpose of this document is to propose solutions to issues and challenges in handling heterogeneous and cross-disciplinary campaign data and to recommend practices for emerging stewardship requirements. The document aims to:

- Increase awareness of the need for high quality airborne and field data and information that are easily discoverable, accessible, readily usable, and interoperable across systems and tools;
- Advocate for consistency in airborne and field campaign data handling across discipline-specific DAACs;
- Provide DAACs with information and recommended specific practices for ingestion, description, and stewardship of airborne and field campaign data assigned to them; and help DAACs in providing guidance to airborne and field campaign data producers.

As the first of a series of ADMG best practice documents, this document describes high-level stewardship best practices that apply to all types of NASA airborne and field campaign data. The recommendations in this document cover aspects of data ingestion, metadata curation, and data release. It is intended for DAAC data management personnel working on airborne and field data stewardship. Data managers of airborne and field campaigns may also find some of this information applicable to their roles.

3. BACKGROUND AND MOTIVATION

Federal agencies such as NASA are mandated to comply with U.S. laws and policies on open data and data sharing (e.g., US Public Law 106-554 2001; 115–254 2018; 115-435 2019; OMB 2013; OSTP 2010; 2013). For federally funded scientific data, agencies must make plans and develop procedures to improve public access and utility of data, including utilizing machine-readable formats for enhanced interoperability across systems and tools (Federal Data Strategy Development Team 2020).

NASA has been sharing NASA-funded data since 1994, and is committed to the agency-wide open data and information policy (NASA 2014)[4] to ensure that NASA data are archived and curated for long-term accessibility and usability. NASA scientific information is freely available and openly shared (Murphy 2017; NASA 2021, aka the SPD-41 policy) to foster scientific exploration and discovery and promote open science (Ramachandran et al. 2021).[5,6] Scientific information in this policy includes publications, data, and software created in the pursuit of scientific knowledge.

These open data and open science policies led to the development of NASA procedures with specific requirements placed on airborne investigations and field campaigns of the NASA Science Mission Directorate Earth Science Division (NASA 2018; 2019).[7] Therefore, it is important for all DAACs to be aware of the various policies and requirements and be compliant with existing and emerging data stewardship requirements.

NASA has carried out many airborne and field campaigns to answer questions across the spectrum of the Earth ecosystem (Figure 1), including clouds, carbon and aerosols, and the cryosphere. Campaign data are routinely utilized in a wide range of research and applications, such as studying and monitoring air quality, monsoon variability, and Arctic ice and Greenland glacier melting (NASA 2019).

Airborne measurements often have high spatial and temporal resolutions and provide a comprehensive suite of observations of the Earth system. The data can augment satellite measurements and provide critical support for model and satellite data validation. Research, development, and validation of NASA Earth Science instruments and algorithms increasingly rely on airborne missions (National Research Council 2010). However, the airborne and campaign data curation has been historically lacking compared to the satellite data (ESDSWG Airborne Data Working Group 2015; ESDS 2018).

[4] https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy

[5] https://earthdata.nasa.gov/esds/open-science

[6] https://earthdata.nasa.gov/collaborate/open-data-services-and-software

[7] https://science.nasa.gov/earth-science/earth-science-data/airborne-mission-data-system-requirements

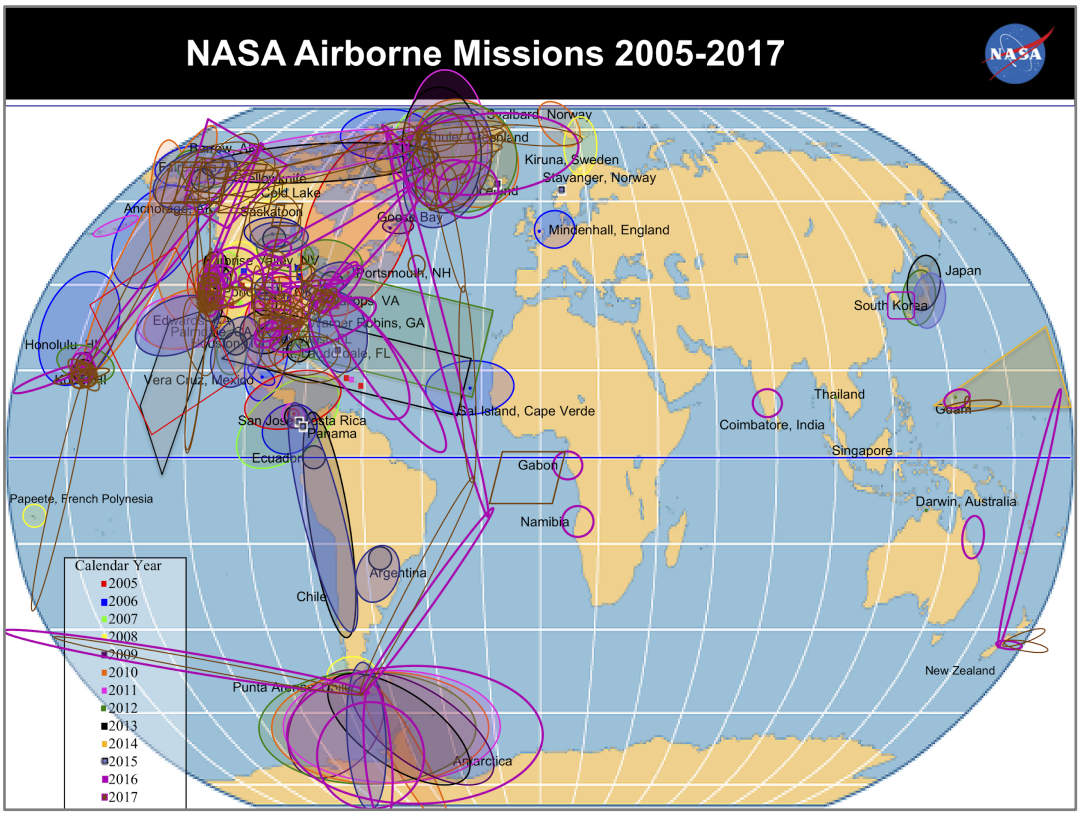

Figure 1: NASA Airborne Missions from 2005 to 2017. Source: https://airbornescience.nasa.gov/sites/default/files/documents/NASA_Airborne_Science_Mission_Map.pdf (Maps for 2018 and 2019 can be found in NASA airborne missions annual reports (NASA 2018; 2019)

A key component to enabling open science is data accessibility, which is one of the biggest challenges for airborne and field campaign data (Figure 2). To date, nearly 25% of NASA campaigns are not yet archived at a DAAC. Even after campaign data are delivered to an assigned DAAC, the process of data publication and public release can take years. In addition, data product contextual information may be lacking in many cases, making it difficult for users to easily find and effectively use the data. Best practices that ensure consistency in documenting NASA airborne data and associated information will not only help to enhance interoperability across systems and tools, but also improve the speed of data release and support of improved user experiences in finding, accessing, trusting, and (re)using these unique and valuable data products.

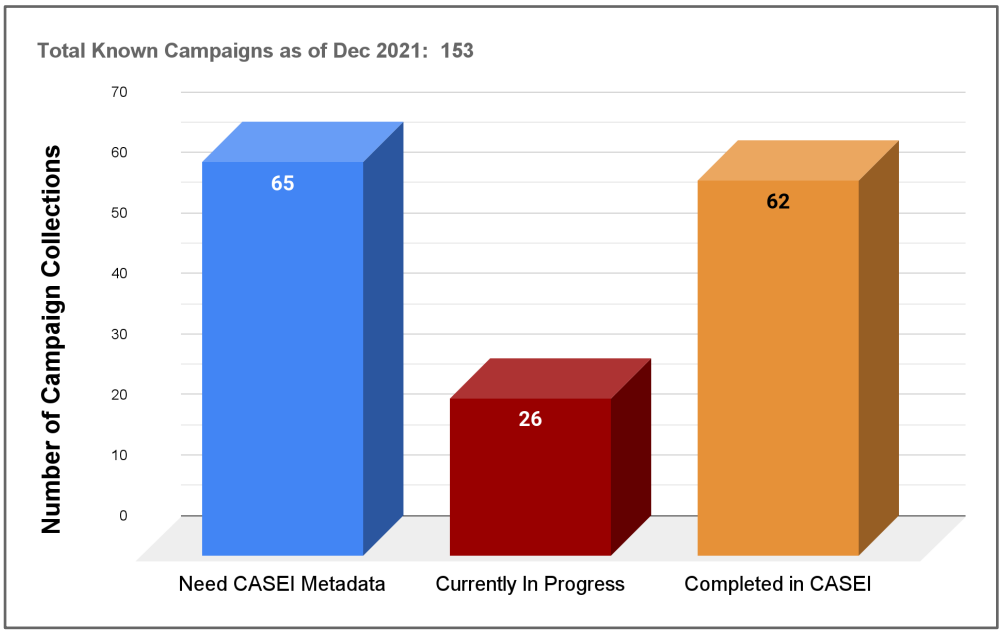

Figure 2: A breakdown of the current (December 2021) status of 153 NASA campaigns that have been identified by ADMG in the context of CASEI campaign metadata curation process. *Complete* denotes that campaign metadata have been curated and published to CASEI. The total number of campaigns evolves as more data products from historical campaigns are located or new campaigns are funded and added to the CASEI inventory. Currently, 39 campaigns from the left column, about 25% of all known campaigns, are not yet assigned to a DAAC.

NASA Earth Science data products derived from airborne and field instrument measurements are located throughout the NASA enterprise. Currently only 10 out of the 12 NASA DAACs have these types of data (see Figure 3 showing the top 6 DAACs with airborne and field campaigns). Each DAAC primarily serves one or more specific scientific discipline communities. For example, the Atmospheric Science Data Center (ASDC) provides access to data products in the Earth Science disciplines of clouds, aerosols, atmospheric composition and radiation budget,[8] while the National Snow and Ice Data Center (NSIDC) provides data and information for snow and ice processes.[9] Data management practices are developed and customized by a given DAAC to best support the particular communities it serves. It is beneficial to glean this institutional knowledge from individual DAACs to use towards developing the effective approaches to bolster efforts across all DAACs. This will promote consistency across all DAACs and provide data producers and users with a more uniform DAAC experience, thereby accelerating data availability.

[8] https://earthdata.nasa.gov/eosdis/daacs/asdc

[9] https://earthdata.nasa.gov/eosdis/daacs/nsidc

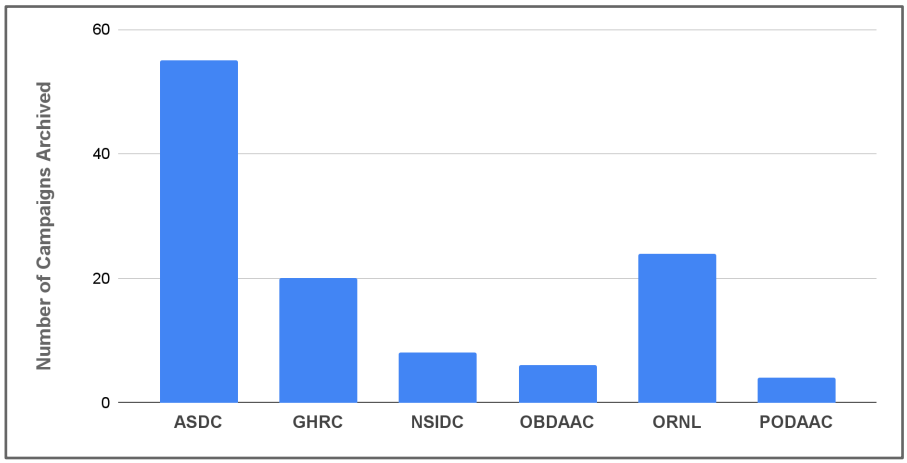

Figure 3: A breakdown of the current (December 2021) number of NASA campaigns in each of the top 6 DAACs for airborne and field campaign data. See Appendix A for DAAC abbreviations.

Due to systematically reviewing and curating metadata for NASA airborne and field campaigns, the ADMG CASEI and the IMPACT Analysis and Review of the Common Metadata Repository (ARC)[10] teams have firsthand knowledge of needed areas of improvement for airborne and field data product metadata. In addition, ADMG locates the data products of historical campaigns and curates comprehensive contextual information to describe the campaigns and data products obtained. See Figure 2 for a breakdown of the status of currently known NASA campaigns. ADMG also addresses terminology convention discrepancies, recognizing that each community has its own terms, and works to better communicate across DAACs and geophysical communities. The efforts of the ADMG CASEI and ARC teams help to both identify challenges and aid in developing solutions to effectively address issues.

Other federal Earth Science agencies, such as the U.S. Geological Survey (USGS) and the National Oceanic and Atmospheric Administration (NOAA), face similar challenges in managing their airborne and field campaign data. Stewardship best practices from these agencies and other organizations can also be beneficial to NASA's efforts.

4. METHODOLOGY

The methodology for developing the best practices presented in this document included the following activities to review and collect information in various aspects highlighted in bold:

- Review and collection of **existing practices** utilized by DAACs:
  - Literature review (e.g., GES DISC 2017; Cook et al. 2018),
  - DAAC websites review (e.g., NSIDC DAAC information for SnowEx data providers[11]; ORNL DAAC data management best practices[12] and data quality

[10] https://earthdata.nasa.gov/esds/impact/arc

[11] https://nsidc.org/data/snowex/information-snowex-data-providers

[12] https://daac.ornl.gov/datamanagement/#best_practices

review checklist[13]; PO.DAAC data management best practices, including file formats and file metadata conventions[14]),
    - Collection and analysis of feedback via a targeted survey of DAAC personnel responsible for handling airborne and field campaign data (11 responses from 8 DAACs; ADMG 2021).
- Collection and analysis of IMPACT team **feedback**:
    - ADMG CASEI content curators (ADMG 2021),
    - ARC (Analysis and Review of CMR - NASA Common Metadata Repository; ADMG 2021).
- Review of U.S. Laws and federal **mandates** (e.g., US Public Law 106-554 2001; 115-254 2018; 115-435 2019; OMB 2013; OSTP 2010; 2013).
- Review of NASA **requirements**, including:
    - NASA Earth Science data and information policy (NASA 2021)[15], procedures (e.g., DOI process[16]) and standards (e.g., data formats)[17],
    - NASA's Strategic Plan for increased access to digital scientific data (NASA 2014),
    - ESDIS Mission Requirements,[18]
    - Airborne Mission Data System Requirements.[19]
- Review of previous **user needs** surveys:
    - NASA Satellite Needs Working Group (SNWG 2016),
    - Prospective CASEI user survey (Smith et al. 2020).
- Review of data stewardship **best practices from other Earth Science entities:**
    - Literature review (e.g., WMO 2019; Ocean best practices[20]; European Facility for Airborne Research[21])
    - Review of information on other federal websites (e.g., USGS stewardship best practices - prepare data to share[22]; NOAA HATS airborne data[23]; NOAA story maps of field campaigns[24]).

To develop the ADMG recommended best practices described in the next section, the information from the above efforts were analyzed and synthesized, solutions were identified for the known challenges, and DAAC-specific best practices were integrated into a cohesive set of high-level practices.

[13] https://daac.ornl.gov/submit/qachecklist/

[14] https://podaac.jpl.nasa.gov/PO.DAAC_DataManagementPractices

[15] https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy

[16] https://earthdata.nasa.gov/collaborate/doi-process

[17] https://earthdata.nasa.gov/esdis/eso/standards-and-references#standards-and-practices

[18] https://earthdata.nasa.gov/esdis/eso/esdis-mission-requirements

[19] https://wiki.earthdata.nasa.gov/display/library/ESDS+Document+Repository

[20] https://www.oceanbestpractices.org

[21] https://www.eufar.net/cms/good-practices/

[22] https://www1.usgs.gov/csas/training/dm-preparing/resources/PreparingDataToShare.pdf

[23] https://gml.noaa.gov/hats/airborne/hats_airborne_data.html

[24] https://storymaps.arcgis.com/stories/ce1a17c290c04f45a373af45e2b8d449

5. HIGH-LEVEL STEWARDSHIP BEST PRACTICES

To ensure quality and consistency in documenting airborne and field campaign data, it is recommended that DAACs perform the following general actions when handling NASA airborne and field campaign data.

Recommendation 1: Communicate clearly and effectively with airborne and field campaign management, data producers and users across multiple disciplines.

Specifically:

- Develop and use a uniform checklist to ensure strategic and systematic communication with each campaign team,
- Interact with and support the science team both before and during the campaign in order to:
    - Gain understanding of the scientific purpose, field activities, significant observation events, and documentation of the campaign,
    - Present on DAAC activities pertinent to the campaign science team and review roles and expectations:
        - Use plain language when communicating with the science team - avoid data stewardship jargon whenever possible and use common terminology as defined in the ADMG inventory definitions (ADMG 2020),
    - Aid in the development of and adherence with the campaign's data management plan for data producers,[25] in coordination with the campaign data manager and/or ADMG,
    - Train and assist data producers in understanding and applying NASA and DAAC file naming conventions, data file formats, keywords, metadata standards, and documentation requirements.

      NOTE: ORNL DAAC maintains a data quality review checklist[26] which is a good example of communicating the DAAC process and standards to the data producers.
    - Ensure important contextual information is obtained from the science team before team members move on to other campaign efforts or field archives are removed.
- Use and encourage campaign scientists' use of controlled vocabularies as appropriate, including terms from:
    - EOSDIS Glossary (ESDIS 2020),
    - GCMD (Global Change Master Directory) keywords[27] for projects, campaigns, platforms, and instruments,
    - Data processing levels (ESDS 2019),
    - ADMG CASEI definitions,[28]
    - NASA airborne platforms and instruments.[29]

[25] https://earthdata.nasa.gov/collaborate/new-missions/data-management-plan-guidance

[26] https://daac.ornl.gov/submit/qachecklist/

[27] https://gcmd.earthdata.nasa.gov/KeywordViewer/scheme/all/

[28] https://earthdata.nasa.gov/esds/impact/admg/admg-definitions

[29] https://airbornescience.nasa.gov/

- Develop a DAAC data management plan using the NASA DAAC DMP template [30] to clearly communicate how the DAAC will care for the data products after they are obtained from the data producers. The DAAC data management plan should specify the anticipated time frame for public release of the data. In addition, the DAAC data management plan should be:
  - Initiated at the beginning of collaboration with the science team
  - Shared with the science team to get feedback
  - Regularly reviewed and updated as needed

Recommendation 2: Capture and maintain a comprehensive set of critical information and documentation for every campaign.

Specifically:

- Be consistent in the organization of data and information across campaigns,
- Establish a discoverable and consistently arranged campaign "landing page" at the DAAC that includes and agrees with CASEI contextual metadata (see details below), contains well organized additional materials, and lists campaign data products,

  NOTE: Currently, DAACs curate data product metadata records after ingesting data products from data producers (campaign scientists). These data product metadata are incorporated into CMR - the metadata repository. ADMG curates and provides users with additional contextual campaign metadata and information links via CASEI - an airborne and field inventory. The relationship between content in CMR, CASEI, and DAAC campaign landing pages and metadata content is still evolving. Therefore, it is essential that DAACs and ADMG work together to ensure the comprehensive metadata and information needed by users are obtained and kept up-to-date in a timely manner.

  CASEI is designed to include these contextual metadata:
  - A comprehensive campaign overview that contains the:
    - Acronym and full name of the campaign,
    - Date range and description of the campaign domain,
    - Scientific objectives or campaign purpose.
  - The campaign principal investigator(s) (PIs),
  - A list of deployments, with the following accompanying information:
    - Date ranges of deployments,
    - Locations of deployments, if different,
    - Platforms utilized,
    - Instruments by platform (including picture of instrument placement, when available),
    - Variables measured using the instrument,
  - List of campaign data products and DOIs,
  - Known significant events occurring during the campaign with short descriptions and dates,
  - List of primary journal publications (e.g., a campaign overview or summary paper), if available.

[30] https://earthdata.nasa.gov/collaborate/new-missions/data-management-plan-guidance

DAACs should consider adding additional materials to the campaign landing page, such as:

- Table of field station location(s), with coordinates (lat/lon), station name(s) or number(s), site image(s), and station operational status over time,
- Moving platform tracks (e.g., from aircraft, ships/boats, vehicles, etc.), preferably shown as plots of tracks by deployment or day of operation,
- List of campaign data products and DOIs with short, informative descriptions organized by variable or relevance to science objectives, consider adding data product guide access in the table,
- Detailed description(s) of known significant events during the campaign, if available,
- Detailed list of associated journal publications, especially highlighting any calibration results or data cross-comparisons,
- Data preservation table with all information indicated by the Preservation Content Specification (PCS) document,
- Access to and description of tools for use with campaign data products.

- Coordinate with ADMG to gather and capture the needed contextual metadata and information, should it not already exist in CASEI.
- Ensure the DAAC campaign landing page content is machine accessible, to be in line with the NASA digital strategy,[31] and to make it easier for users to integrate the information across their own tools and systems,
- Assign a DOI as a persistent campaign identifier to this campaign landing page,[32]
- Publish campaign reports (such as science reports or flight reports) as data products, using consistent terminology when naming the products and consistent formats,
- Publish any moving platform tracks as separate navigation data products for easy use by tools and services.
- Update metadata as new information comes to light or circumstances change and ensure that these updates are appropriately reflected in the DAAC campaign landing page, CASEI, and CMR.

In addition, it is beneficial to provide information on sensor calibration, data product validation, and the procedures utilized to ensure data product quality.

Recommendation 3: Manage campaign data, metadata, and information as first class assets.

Campaign data and information are equally as valuable to NASA and the public as satellite data and, therefore, should be managed according to the same high standards. These valuable data are subject to the requirements of SPD-41, the Scientific Information Policy for the Science Mission Directorate (NASA 2021), which states that all NASA-funded research data and information must be openly available and freely shared. In addition, there are three stewardship service levels (basic, standard, and comprehensive) defined in the ESDS Level of Service document that are based

[31] https://www.nasa.gov/digitalstrategy

[32] How a DOI to a campaign page should be implemented is still under discussion by the Earth Science DOI Working Group

primarily on funding and data types.[33] All NASA data must adhere to the basic service level. The requirements below pertain to most of the actions for the basic service level that apply to all NASA data.

Applying, however, service levels to airborne and field campaign data management is complicated. Generally speaking, campaign data can be categorized into six categories of airborne and field collections (listed in order from higher service level to lower service level):

- NASA Flight Campaigns (such as Earth Venture Suborbital or Operation IceBridge),
- NASA Research and Analysis campaigns typically funded through ROSES,
- Facility Instruments or Major Airborne Instruments,
- Calibration/Validation Activities/Campaigns focused on satellite instrument or algorithm development,
- NASA as a funder or significant partner of Multi-Organization regional campaigns, and
- Instrument Networks on the Earth's surface, both land and ocean.

The category indicates the service level for the scientifically important data collected. Within any one campaign, however, there will be varying levels of service applied to data products. For instance, the Earth Venture Suborbital campaigns will contain both important data that should get a comprehensive level of service and other support data, such as NOAA NEXRAD radar data or forecast model data that can have a basic level of service.

The stewardship best practices recommended below may already be utilized for NASA satellite data - and possibly for some campaign data at some DAACs. To ensure improved stewardship of all NASA campaign data and compliance with the agency's core requirements, DAACs are encouraged to continue using or adopt the following specific stewardship practices for handling campaign data:

- Adhere to NASA standards and conventions in data stewardship, particularly:
    - Follow NASA metadata standards,
        - CMR collection-level metadata[34],
        - CMR granule metadata.[35]
    - Use GCMD keywords, including those for projects, platforms, and instruments. In particular, DAACs are encouraged to:
        - Contribute to GCMD keyword maintenance (participating in reviews and updates),
        - Request new GCMD keywords as needed for campaign data publication.
    - Use the NASA Preservation Content Specifications (PCS) (Ramapriyan et al. 2022a); and the PCS Implementation Guidelines documents (Ramapriyan et al. 2022b) to collect and organize preservation materials throughout the campaign,

[33] https://earthdata.nasa.gov/collaborate/new-missions/level-of-service

[34] https://wiki.earthdata.nasa.gov/display/CMR/UMM-C+Schema+Representation

[35] https://wiki.earthdata.nasa.gov/display/CMR/UMM-G+Schema+Representation

- Advise, encourage, and assist data producers to use a machine-interoperable, NASA-approved data file formats[36] (Ramapriyan and Leonard 2020; Jelenak et al. 2019).
  - Ramapriyan and Leonard (2020) have specifically recommended the NASA Earth Science community to adopt the netCDF-4 data file format.
  - Within the context of NetCDF data file format, follow the following community metadata conventions for improved data discovery and interoperability, including:
    i. Attribute Conventions for Data Discovery (ACDD)[37],
    ii. Climate Forecast (CF) conventions[38]

    NOTE: PO.DAAC provides a list of key ACDD/CF global metadata attributes with implementation examples, which serve as a good reference.

- Ensure and maintain quality of ingested and released data and information:
  - Confirm integrity (file names, checksums, file sizes, and number of files) of data products and any supplementary files provided,
  - Confirm consistency in:
    - File and parameter naming conventions,
    - File organization.
  - Verify compliance with NASA standards and recommendations:
    - File format,
    - Date and time,
    - Metadata,
    - GCMD Keywords.
  - Periodically confirm that the referenced data and information remain accessible and readable,
  - Regularly update data and information, as needed, leveraging CASEI (e.g., campaign papers may be published a couple of years after the end of the campaign, or CASEI metadata may be updated over time)
- Ensure cross-connection of information and use of networking tools to improve information access and education for user communities:
  - Provide campaign or flight catalogs to improve summaries of activity details;
  - Ensure highly inter-connected metadata relevant to related instruments, satellites, models and other campaigns;
  - Work with the other DAACs and other organizations to provide cross-linkages to other information and data that may exist;
  - Broadly engage new communities to introduce the data products and information from a campaign or instrument;
  - Improve tools and services to ease data use.

[36] https://earthdata.nasa.gov/esdis/eso/standards-and-references#data-formats

[37] https://wiki.esipfed.org/Category:Attribute_Conventions_Dataset_Discovery

[38] https://earthdata.nasa.gov/esdis/eso/standards-and-references/climate-and-forecast-cf-metadata-conventions

6. SUMMARY AND FUTURE WORK

These high-level stewardship best practices are recommended to support a uniform reporting layer for NASA's diverse airborne and field campaign data:

Recommendation 1: Communicate clearly and effectively with airborne and field campaign management, data producers and users across multiple disciplines.
Recommendation 2: Capture and maintain a comprehensive and consistent set of high-quality information and documentation for every campaign.
Recommendation 3: Manage campaign data, metadata, and information as first class assets.

The intent of this document is to provide best practices that are beneficial to DAAC personnel and lead to:

- Culture change toward improved consistency in airborne and field campaign data stewardship, and better communication among all parties involved;
- Accelerated campaign data release at DAACs;
- Improved data discoverability, usability, and utilization by various applications as well as easy integration across systems and tools (e.g., CMR);
- More unified DAAC experience for airborne data producers and users.

These recommended practices apply to all types of airborne and field campaign data. Additional documents, with details relevant for specific types of airborne and field campaign data, are in development. These include (in no particular order):

- Atmospheric chemistry data,
- Campaign reports,
- Field station instrument data,
- In-situ manual observation data,
- Ground and airborne radar and lidar data,
- Facility instrument data, and
- Unmanned Aircraft System (UAS) data.

New federal mandates on improving the sharing and (re)use of federally funded research data may be created in the future. New technology such as machine learning or emerging data storage platforms such as Earthdata cloud may bring a new designated user community. New DAAC and community stewardship best practices may be developed over time. Best practices for the stewardship of NASA funded research data should be regularly updated to accommodate future federal mandates, agency policies, and evolving community best practices. Changes will result in future updates to this document.

ADMG also welcomes feedback to further improve the clarity and comprehensiveness of this document. The contact information can be found at:
https://earthdata.nasa.gov/esds/impact/admg.

ACKNOWLEDGEMENTS

ADMG is supported by NASA through the MSFC IMPACT project under a cooperative agreement with the University of Alabama in Huntsville (NASA Grant NNM11AA01A). We thank Katrina Virts for providing information on the satellite needs survey database (SNWG 2016).

Feedback from DAAC data managers and other personnel who work with airborne and field campaign data, ADMG CASEI inventory content curators, and the ARC team members, was extremely beneficial in developing these stewardship best practice recommendations. In particular we thank Bruce Wilson, Kasey Phillips, Walt Meier, Badford Wirt, Sheri Fick, David Meyer, Geoffrey Stano, Michelle Harbin, Jack McNeils, Steve Tanner, and their colleagues for participating in the DAAC survey and providing comprehensive points of view and beneficial information. We thank ADMG CASEI inventory content curators, especially, Shelby Bagwell, Ashlyn Shirey, and Lucia Alonso Guzman for providing feedback and beneficial information based on their experience of curating metadata for airborne and field campaign data. We thank the ARC team members, especially, Jeanné le Roux, Camille Woods, Stephen McNeal, and Jenny Wood for providing feedback and beneficial information based on their experience of evaluating the metadata quality of DAAC data product metadata records in NASA's CMR.

We thank Kaylin M. Bugbee, Aaron Kaulfus, Danielle Groenen, Jeanné le Roux, and Camille Woods for reviewing the previous versions of the document as part of the IMPACT team review process. Their comments and suggestions have improved the quality of this document.

A special thanks to Kasey E. Phillips, Megan E. Buzanowicz, and Bruce Wilson for reviewing the v02r01-2021221 version of the document with constructive and beneficial comments and suggestions as part of the DAAC review process.

REFERENCES

ADMG (Airborne Data Management Group), 2020: Maintenance and User Discovery Interface Requirements - Conceptual Operations Document. Version: v2.1 January 18, 2020. *Available upon request.*

Airborne Data Management Group, 2021: Airborne and Field Campaign Data Stewardship Survey Responses and Post Survey Analysis Results. Version: v1 September 30, 2021. *Available upon request.*

Cook, R.B., Y. Wei, L.A. Hook, S.K.S Vannan, J.J. McNelis, 2018: Preserve: Protecting Data for Long-Term Use. In: F. Recknagel, W. Michener (eds). *Ecological Informatics*. Springer, Cham. https://doi.org/10.1007/978-3-319-59928-1_6

Curtis, L., 2019: The Best Ways to Document and Share Best Practices among Colleagues. Version: April 29, 1019. Accessed: February 3, 2021. Available at: https://www.mpug.com/the-best-ways-to-document-and-share-best-practices-among-colleagues

ESDIS (Earth Science Data and Information System), 2020: Glossary. Version: 28 January 2020. Available at: https://earthdata.nasa.gov/learn/user-resources/glossary

ESDS (Earth Science Data Systems), 2018: Management of Airborne Data Products: Challenges and Recommendations. Version: February 27, 2018. NASA Internal Document - Available upon request.

ESDS, 2019: Data Processing Levels. Version: 23 August 2019. Available at: https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy/data-levels

ESDSWG Airborne Data Working Group, 2015: Report of ESDSWG Airborne Data Working Group Study. Version: 10/17/2015. Available for download from the NASA EarthData Wiki site (login required).

ESDSWG Airborne Metadata Data Working Group, 2017: Final report of ESDSWG airborne metadata working group. Available for download from the NASA EarthData Wiki site (login required).

Federal Data Strategy Development Team, 2020: Federal data strategy and 2020 action plan. Available at: https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf

GES DISC (Goddard Earth Sciences Data and Information Services Center), 2017: GES DISC Data and Metadata Recommendations to Data Providers. Version: 11/04/2017. Available at: https://docserver.gesdisc.eosdis.nasa.gov/public/project/DataPub/GES_DISC_metadata_and_data_formats.pdf

Jelenak, A., Leonard, P.J.T., and the NASA ESDSWG Dataset Interoperability Working Group, 2019: Dataset Interoperability Recommendations for Earth Science: Part 2. Document ID: ESDS-RFC-036. Version: v1.1 October 2019. Available from: https://earthdata.nasa.gov/esdis/eso/standards-and-references/dataset-interoperability-recommendations-for-earth-science

Murphy, K., 2017: NASA Earth Science Data: Yours to Use, Fully and Without Restrictions. Versions: 01/17/2017. Available at: https://earthdata.nasa.gov/learn/articles/tools-and-technology-articles/nasa-data-policy.

NASA, 2014: NASA Plan for Increasing Access to the Results of Scientific Research: Digital Scientific Data and Peer-reviewed Publications. Version: 11/21/2021. Available at: https://www.nasa.gov/sites/default/files/atoms/files/206985_2015_nasa_plan-for-web.pdf

NASA, 2018: Airborne Science Program - 2018 Annual Report. *Science Mission Directorate*. Available at: https://airbornescience.nasa.gov/sites/default/files/documents/ASP_2018_April19_web.pdf

NASA, 2019: Airborne Science Program - 2019 Annual Report. *Science Mission Directorate*. Available at: https://airbornescience.nasa.gov/sites/default/files/documents/ASP_2019_AnnualReport_0.pdf

NASA 2021: Scientific Information policy for the Science Mission Directorate. SMD Policy Document SPD-41. Version: August 4, 2021. Available at: https://science.nasa.gov/science-red/s3fs-public/atoms/files/Scientific%20Information%20policy%20SPD-41.pdf

National Academies of Sciences, Engineering, and Medicine, 2018: Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space. Washington, DC: The National Academies Press. https://doi.org/10.17226/24938

National Research Council, 2007: Environmental Data Management at NOAA: Archiving, Stewardship, and Access. Washington, DC: *The National Academies Press*. https://doi.org/10.17226/12017.

National Research Council, 2010: Revitalizing NASA's Suborbital Program: Advancing Science, Driving Innovation, and Developing Workforce. 98 pp. Washington, DC: *The National Academies Press*. https://doi.org/10.17226/12862

OMB (Office of Management and Budget), 2013: Open Data Policy – Managing Information as an Asset. Version: OMB Memorandum May 9, 2013. Available at: https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf

OSTP (Office of Science and Technology Policy), 2010: Scientific integrity. Available at: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/scientific-integrity-memo-12172010.pdf

OSTP, 2013: Increasing access to the results of federally funded scientific research. Version: OSTP Memorandum February 22, 2013. Available at: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Ramachandran, R., K. Bugbee, and K. Murphy, 2021: From open data to open science. *Earth and Space Science*, 8, e2020EA001562. https://doi.org/10.1029/2020EA001562

Ramapriyan, H. K., and P. J. T. Leonard, 2020: Data Product Development Guide (DPDG) for Data Producers version 1. *NASA Earth Science Data and Information System Standards Office*, 9 July 2020. https://doi.org/10.5067/DOC/ESO/RFC-041VERSION1

Ramapriyan, H. J. Moses, and D. Smith, 2022a: NASA Earth Science Data Preservation Content Specification. Documentation ID: 423-SPEC-001. https://doi.org/10.5067/doc/ESO/423-SPEC-001

Ramapriyan, H., J. Moses, and D. Smith, 2022b: Preservation Content Implementation Guidance. Document ID: ESDS_RFC-042. Version: 1.0, January 25, 2022. https://doi.org/10.5067/DOC/ESO/RFC-042

Smith, D. K., S. M. Wingo, C. R. Davis, K. Bugbee, R. Ramachandran, and B. Freitag, 2020: Construction of an Airborne Data Inventory for Improved Data Discoverability and Access. *The AMS Annual Meeting 2020*. Virtual. Available at: https://ntrs.nasa.gov/citations/20200000477

SNWG (Satellite Needs Working Group), 2016: 2016 survey NASA response database. NASA Internal Document - Available upon request.

US Public Law 106-554, 2001: Information Quality Act. Publ. L. 106-554. Page 101. Available online at: http://www.gpo.gov/fdsys/pkg/PLAW-106publ554/html/PLAW-106publ554.htm

US Public Law 115-254, 2018: FAA Reauthorization Act Of 2018. Title VII - Subtitle F: Geospatial Data Act. Available online at: https://www.govinfo.gov/content/pkg/PLAW-115publ254/pdf/PLAW-115publ254.pdf

US Public Law 115-435, 2019: Foundations for Evidence-Based Policymaking Act of 2018. Title II OPEN Government Data Act. Version: 14 January 2019. Available at: https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf

WMO, 2019: Manual on the High-quality Global Data Management Framework for Climate. Document ID: WMO-No. 1238. *World Meteorological Organization*. Available at: https://library.wmo.int/index.php?lvl=notice_display&id=21686

APPENDIX A. ABBREVIATIONS AND ACRONYMS

ADMG	Airborne Data Management Group
ARC	Analysis and Review of CMR
ASDC	Atmospheric Science Data Center
ASF	Alaska Satellite Facility
CASEI	Catalog of Archived Suborbital Earth Science Investigations
CMR	Common Metadata Repository
DAAC	Distributed Active Archive Centers
ESDS	Earth Science Data Systems
ESDIS	Earth Science Data and Information System
EOSDIS	Earth Observing System Data and Information System
GCMD	Global Change Master Directory
GES DISC	Goddard Earth Sciences Data and Information Services Center
GHRC	Global Hydrometeorology Resource Center
IMPACT	Interagency Implementation and Advanced Concepts Team
LP DAAC	Land Processes Distributed Active Archive Center
NSIDC	National Snow and Ice Data Center
OB DAAC	Ocean Biology Distributed Active Archive Center
ORNL	Oak Ridge National Laboratory
PO DAAC	Physical Oceanography Distributed Active Archive Center